Title
Spectral Estimation Techniques for time series with Long Gaps: Applications to Paleomagnetism and Geomagnetic Depth Sounding

Permalink
https://escholarship.org/uc/item/7hn0993t

Author
Smith-Boughner, Lindsay

Publication Date
2016

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Spectral Estimation Techniques for time series with Long Gaps: Applications to Paleomagnetism and Geomagnetic Depth Sounding

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Earth Sciences

by

Lindsay Smith-Boughner

Committee in charge:

Professor Catherine G. Constable, Chair
Professor Duncan C. Agnew
Professor William A. Coles
Professor Kerry Key
Professor Robert L. Parker

2016
The dissertation of Lindsay Smith-Boughner is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2016
DEDICATION

To my husband, Robert, my daughter, Amelia and all of my family and friends for their love and support.
# TABLE OF CONTENTS

Signature Page ........................................................................ iii
Dedication ............................................................................. iv
Table of Contents .................................................................... v
List of Figures .......................................................................... viii
List of Tables ............................................................................ xvi
Acknowledgements ................................................................... xvii
Vita ......................................................................................... xviii
Abstract of the Dissertation .................................................. xx

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Introduction ..................................................................... 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1 Motivation .......................................................... 1</td>
</tr>
<tr>
<td></td>
<td>1.2 Outline .................................................................... 4</td>
</tr>
<tr>
<td></td>
<td>1.3 Earth’s magnetic field ............................................. 7</td>
</tr>
<tr>
<td></td>
<td>1.4 Measuring the Magnetic Field ........................................ 10</td>
</tr>
<tr>
<td></td>
<td>1.5 Modeling the Magnetic Field ........................................ 11</td>
</tr>
<tr>
<td></td>
<td>1.6 Studying the variability in time .................................... 12</td>
</tr>
<tr>
<td></td>
<td>1.7 Geomagnetic Induction ............................................... 14</td>
</tr>
<tr>
<td></td>
<td>1.8 Estimates of Earth’s electromagnetic response .................. 17</td>
</tr>
</tbody>
</table>

Bibliography ............................................................................. 22

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Spectral Estimation for Geophysical Time Series with Inconvenient Gaps ...................................... 26</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.1 Introduction ................................................................... 27</td>
</tr>
<tr>
<td></td>
<td>2.2 Spectral Estimation ........................................................................ 30</td>
</tr>
<tr>
<td></td>
<td>2.2.1 Prolate Multi-tapers ........................................................................ 32</td>
</tr>
<tr>
<td></td>
<td>2.2.2 Minimum Bias Tapers ........................................................................ 33</td>
</tr>
<tr>
<td></td>
<td>2.2.3 Adapting Tapers for Inconvenient Gaps ........................................................................ 33</td>
</tr>
<tr>
<td></td>
<td>2.3 Evaluating a Set of Tapers ........................................................................ 34</td>
</tr>
<tr>
<td></td>
<td>2.4 Confidence Limits ........................................................................ 40</td>
</tr>
<tr>
<td></td>
<td>2.4.1 Confidence limits using empirical resampling ........................................................................ 42</td>
</tr>
<tr>
<td></td>
<td>2.5 Algorithm ........................................................................ 48</td>
</tr>
<tr>
<td></td>
<td>2.6 Applications to Geophysical Data ...................................................................................... 53</td>
</tr>
</tbody>
</table>

2.6.1 Length of Day ........................................................................ 54
Chapter 5 Satellite Induction Study using Adapted Multi-tapers

5.1 Introduction ................................................. 164
5.2 Magnetic Satellite Data ................................. 167
5.3 Processing Satellite Data ............................... 170
  5.3.1 Choice of Main Field model ..................... 170
5.4 Solving for $e_{1}^{0}(t)$ and $i_{1}^{0}(t)$ ........... 174
5.5 Data Selection Considerations ..................... 175
5.6 Estimation the response function, $G_{te}$ .......... 179
5.7 Time Sampling ........................................... 185
5.8 Estimating the Noise in the Signals ............... 189
5.9 Applying Minimum Bias Tapers with Gaps ....... 189
5.10 Comparing Response Functions from Section Averaging and MBG Tapers ......................... 194
5.11 Estimates of the C-response .......................... 194
5.12 Comparisons with Other Methods and Previous Results ........................................ 199
5.13 Discussion and Conclusions ....................... 201

Chapter 6 Conclusions ......................................... 206
6.1 Future Work ............................................. 208
# LIST OF FIGURES

| Figure 1.1: | From Turner et al. (2007): Time variations of the geomagnetic field in relation to the age of the Earth and the Universe, indicating which are of internal and which are of external origin. | 2 |
| Figure 1.2: | From Constable and Constable (2004a): Schematic of the amplitude spectrum of Earth’s magnetic field, shown on a log-log scale with frequency units of cycles per second and amplitude in $T \sqrt{Hz}$. | 6 |
| Figure 1.3: | Time average of the radial component of the magnetic field at the core-mantle boundary, shown in $\mu T$. **Left:** From GUFM, Figure from Johnson and McFadden (2007). **Right:** From CALS10K1.b, Figure from Korte et al. (2011). | 8 |
| Figure 1.4: | Figure from Olsen et al. (2010): Spatial power spectrum of the geomagnetic field at the Earth’s surface. Black dots represent the spectrum of a recent field model (Olsen et al., 2009; Maus et al., 2008). Also shown are theoretical spectra (Voorhies et al., 2002) for the core (blue) and crustal (magenta) part of the field, as well as their | 9 |
| Figure 1.5: | From Kuvshinov (2008): $|Q_n^{(r)}|$ for different periods as a function of degree n. Solid lines: $|Q_n^{(r)}|$ for a 1-D mantle conductivity model without oceans. Dashed lines: $|Q_n^{(r)}|$ for a 1-D model overlaid by a uniform ocean of 15,000 S. | 18 |
| Figure 2.1: | (a) Prolate tapers with gaps for $NW = 12$, computed for the simulated arrangement of gaps. All tapers shown have eigenvalues, $\lambda \approx 1$. (b) Minimum bias tapers with gaps computed for the simulated arrangement of gaps, 10 shown. (c) Indicator function used. | 35 |
| Figure 2.2: | The energy concentration of the adapted tapers across 14500 time samples is shown for 15 PRG tapers with $NW = 12$ (blue solid line) and 15 MBG tapers (green solid line), 15 standard tapers PR tapers with $NW = 12$ (blue broken line) and MB tapers (green broken line). | 36 |
| Figure 2.3: | Average transfer function, $H_{f,K}$ with horizontal red lines denoting the energy level of the set of 15 tapers at half-bandwidth frequency, $W/2$. (a) the prolate tapers with gaps compared to the prolate tapers (both with $NW = 12$). A set of 15 PRG tapers has a half-bandwidth of 0.8 mHz. (b) the minimum bias tapers with gaps (solid lines). | 38 |
| Figure 2.4: | Comparison of the average energy concentration of the MB, MBG, PRG and PR tapers outside of the frequency band of $[0, \frac{2NW(1.2)}{NF}]$ as a function of the number of tapers used to produce the spectral estimate. This frequency band is 20% larger than the averaging. | 44 |
| Figure 2.5: | One of 145 replications of the synthetic process, whose spectrum is shown as the red line in Figure 2.6. | 44 |
Figure 2.6: The spectra of the synthetic data produced from 18 minimum bias tapers on the complete data set (bright green line) and 18 tapers adapted for gaps (yellow line), corresponding to the indicator function in the bottom of Figure 2.1. The data are synthesized to . . . . 45

Figure 2.7: (a) Original indicator function of synthetic data used to demonstrate algorithm. (b) Histogram of the length of breaks from regular sampling of synthetic data. (c) Indicator function of data after using AR(3) filter to interpolate gaps less than 7 sample points in duration. 50

Figure 2.8: The change in the LOD from 1962 to 2009 with simulated gaps, indicated by zero values. . . . . . . . . . . . . . . . . . . . 55

Figure 2.9: Power spectral density from change in the LOD, computed using 15 PRG tapers with data missing (blue line) and using 15 PR tapers (red line) on the complete record. This is compared to a spectrum created using 15 MBG tapers (green). Jackknife 95% confidence . . 56

Figure 2.10: Power spectral density from change in the LOD plotted on a log-log scale. The spectral estimate from the complete record using prolate tapers (red line) is compared to a section average estimate (green line) computed from 4 sections each windowed using several . . . 57

Figure 2.11: (a) The first 10 PRG (NW=12) and MBG tapers are shown for an indicator function with 5 data segments and 11% of the record missing (see Figure 1 (a,b) for legend). (b) The time domain energy concentration of 15 of the tapers show in panel (a). (c) A comparison . 60

Figure 2.12: Time series of the external dipolar fluctuations in the magnetic field of the Earth from Kuvshinov and Olsen (2006) with data removed to simulate gaps. . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

Figure 2.13: Power spectral density from the model of external dipolar fluctuations in the magnetic field from Kuvshinov and Olsen (2006) computed using both 20 MBG tapers (green line) and 20 PRG (blue line) tapers with NW=12. Missing data was simulated in the . . . . . . . . . . 63

Figure 2.14: (a) Time series of paleointensity measurements of the DSDP Leg 73, Core 522 after interpolation using AR(3) Filter on gaps less than 100 000 years. (b) Indicator function for this time series. . . . 66

Figure 2.15: Spectral estimate of the core 522 data from 22.7-34.7 Ma computed from 8 PRG tapers (black line, with blue confidence intervals), with a frequency resolution of 1.3 c/Myr. The spectral estimate from 29.4-34.7 is shown in blue, with orange blue confidence intervals ... 67

Figure 2.16: Prolate tapers with gaps for $k = 1$ and $NW = 12$, computed for the simulated arrangement of gaps in Figure 2.1c for 580 (solid line), 1450 (dotted line) and 2900 (dashed line) sample points then interpolated using cubic splines. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74
Figure 3.1: a) The geomagnetic polarity time scale based on Cande and Kent (1995) (with additional chron from Lowrie and Kent (2004)) is shown at the top of the figure with cryptochrons indicated by the grey arrows above. This illustrates the high reversal rate in the recent...

Figure 3.2: The prolate tapers with gaps for a time-bandwidth product (NW) of 8, shown are the set of 12 tapers used to analyze the entire 12 My calibrated paleointensity time series, from 22.74-34.77 Ma. The eigenvalues, $\lambda$, indicate the amount of energy contained within the...

Figure 3.3: Comparison of the average transfer functions, $H_k(f)$ of the prolate tapers with gaps used in analysis. The half-bandwidth, $\frac{\pi}{2}$ for a set of $K$ tapers is determined by the mid-point of the steepest drop in amplitude. a) Average transfer function of the tapers shown in...

Figure 3.4: Spectral estimate of the entire 12 My record produced from a set of prolate tapers with 5 gaps of 0.01 to 0.64 million years in length and a time bandwidth product of 8; spectrum has a frequency resolution of $1.4Myr^{-1}$. There is a broad peak at 8 cycles per million years...

Figure 3.5: Spectral estimate of the calibrated paleointensity record from 22.74-28.76 Ma. The estimate was produced by averaging estimates from 6 prolate tapers with gaps (PRG) with a time bandwidth product of 3.5 resulting in a frequency resolution of $1.48Myr^{-1}$. The...

Figure 3.6: Spectral estimate produced from 6 prolate tapers with NW=3.5 for a complete data set, yielding a frequency resolution of $1.32Myr^{-1}$. A spectral estimate from the PADM2M model for 0-2Ma is plotted (dotted line) for comparison.

Figure 3.7: Gaps, similar to the gaps in the younger Oligocene data are imposed on the older Oligocene data (from 29.4 to 34.7 Ma) to test compare the behaviour of PR and PRG tapers. The indicator function, $I(t)$ is 0 when there is a gap in the data at time $t$ and 1 otherwise.

Figure 3.8: Spectrum of data from 29.40-34.77 Ma from using 6 PR tapers on the complete record (blue line with 95% confidence region shaded in light blue) and 6 PRG tapers (red line with 95% confidence region shaded in light red) with gaps imposed to match gaps in 22.74-.

Figure 3.9: Top: Close up of the spectral estimates from the entire 12 My record plotted on a log-log scale (black line with 95% confidence region shaded in grey), the older portion (blue line with 95% confidence region shaded in light blue) and the more recent portion.

Figure 4.1: Indicator function showing availability of data for an arrangement with 5% missing in three pieces, $I_1 = 0$ when no data is available.

Figure 4.2: Sum of the absolute value squared of a set of tapering functions multiplied by the number of sample points (10461 points with data spread over 10953 sampling intervals) and the indicator function shown in Figure 4.1.
Figure 4.3: Comparison of the average transfer function for a set of MB tapers (broken lines) and MBG (solid lines). The indicator function of the MBG tapers is shown in Figure 4.1. Shown with a sampling frequency, $f_s = 6$ cycles/day.

Figure 4.4: Time series of a model of external (top panel) and internal (bottom panel) dipolar fluctuations of the Earth’s magnetic field in response to changes in the ring current from Kuvshinov and Olsen (2006) shown in black. Synthetic data simulated from Dst and the method.

Figure 4.5: Estimated transfer function between input, $x(t) = e^{0_1(t)}$ and output, $y(t) = i^{0_1(t)}$ from Kuvshinov and Olsen (2006) (shown in black). The transfer function is computed using 100 MB tapers. The transfer function between the synthetic time series using 100 MB tapers.

Figure 4.6: The estimate of the synthetic, $x_d$ (blue line) and original data, $x(t)$ (black line) of the external dipole coefficient time series are both computed with 100 minimum bias tapers. The energy level of the time series differs at 1 cycle per day and beyond 1.6 cycles per day.

Figure 4.7: The estimate of the synthetic, $y_d$ (blue line) and original data, $y(t)$ (black line) of the internal dipole coefficient time series are both computed with 100 minimum bias tapers. The energy level of the time series differs at 1 cycle per day and beyond 1.6 cycles per day.

Figure 4.8: Estimate of the mean-squared coherence between the synthetic data mimicking the model of internal and external dipolar fluctuations of the Earth’s magnetic field (blue lines) and compared to the data from Kuvshinov and Olsen (2006) (black lines). The estimate was.

Figure 4.9: Q-Q plot comparing the average empirical distribution of 1000 bootstrap resamplings from the transfer function, estimated using 100 minimum bias tapers, between the data sets from Kuvshinov and Olsen (2006) (shown in black on Figure 4.5) to the theoretical.

Figure 4.10: Estimate of the transfer function with $x_d$ as the input and $y_d$ the output. The transfer function from the synthetic data is shown in black. 5% noise in $y_d$, the output is shown in red and the estimate with 5% noise in $x_d$, the input is shown in blue. Here the.

Figure 4.11: Spectral estimate of $x_d$ with 95% confidence intervals, shown in red is the synthetic time series with dark red confidence intervals. Adding 5% noise to the synthetic time series yields the blue spectral estimate with dark blue confidence intervals. The spectral estimate.

Figure 4.12: Spectral estimate of the $y_d$ with 95% confidence intervals, shown in red is the synthetic time series with light grey confidence intervals. Adding 5% noise to the synthetic time series yields the blue spectral estimate with dark grey confidence intervals. The spectral estimate.
Figure 4.13: Estimate of the transfer function between $y_d$ and $x_d$ with input and output reversed when 5% relative noise has been added to $x_d$-the output (blue line with grey confidence intervals). The transfer function estimate from 100 MB tapers with 5% relative noise has been .

Figure 4.14: Estimate of the transfer function between $y_d$ and $x_d$ when 10% relative noise has been added to $x_d$ and 5% relative noise has been added to $y_d$, $\psi = 2$ (shown in blue with grey confidence intervals). The amplitude of transfer function has been .

Figure 4.15: Comparison of the mean squared coherence between $y_d$ and $x_d$ under various noise conditions. Shown in blue is the coherence when 10% relative noise has been added to $x_d$ and 5% relative noise has been added to $y_d$, $\psi = 2$. For $\psi = 0.5$, 5% relative noise has been .

Figure 4.16: Comparison of the response function estimate using 100 tapers for the internal dipole coefficient model from the synthetic data series (black line). With breaks in sampling corresponding to Figure 4.1, the transfer function estimate from 100 MBG tapers .

Figure 4.17: Comparison of power spectral density estimates using 100 MBG tapers (blue lines) to 100 MB tapers (red lines). Top: Synthetic model series of $y_d$, shown in .

Figure 4.18: Comparison of the response function estimate using 100 tapers for the internal dipole coefficient model from the synthetic time series (black line) and from the synthetic series with 5% noise added to $x_d$ (red line). With breaks in sampling corresponding to .

Figure 4.19: Comparison of the transfer function estimates from 100 tapers with 10% noise added to $x_d$, the input and with 5% noise added to $y_d$, the output (red line). The model is shown in black. Creating breaks in the sampling of both noisy time series (with $\psi = 2$) corresponding .

Figure 4.20: Comparison of the coherence estimate estimate using 100 tapers for the internal dipole coefficient model from a complete data set (black line) and from a complete data set with 10% noise added to $x_d$ and 5% noise added to $y_d$ (red line with grey confidence intervals).

Figure 4.21: Indicator function showing availability of data, $I_t = 0$ when no data is available. Out of a record spanning 10954 sample intervals, only 6824 data records are available, 38% of the record is missing .

Figure 4.22: Sum of the absolute value squared of a set of tapering functions multiplied by the number of sample points (10954) and the indicator function shown in Figure 4.21. The blue line is from 100 minimum bias tapers compared to 100 MBG tapers (green line).

Figure 4.23: Comparison of the average transfer function for a set of minimum bias tapers (broken lines) and minimum bias tapers with breaks in sampling (see Figure 4.21). Shown with a sampling frequency, $f_s = 6$ cycles/day.
Figure 4.24: Comparison of the average transfer function for a set of minimum bias tapers (broken lines) and minimum bias tapers used in section averaging, with 7 tapers used on the last segment of data and 10 tapers applied to the remainder for a total of 57 tapers. Shown with

Figure 4.25: Comparison of the response function estimate from data with 5% noise in the input (see Figure 4.21 for the indicator function).

Figure 4.26: Minimum bias tapers with gaps (K=1-5) for the indicator function shown in Figure 4.1.

Figure 4.27: Minimum bias tapers with gaps (K=6-10) for the indicator function shown in Figure 4.1.

Figure 4.28: Minimum bias tapers with gaps (K=11-15) for the indicator function shown in Figure 4.1.

Figure 4.29: Minimum bias tapers with gaps (K=16-20) for the indicator function shown in Figure 4.1.

Figure 4.30: Minimum bias tapers with gaps (K=21-25) for the indicator function shown in Figure 4.1.

Figure 4.31: Minimum bias tapers with gaps (K=1-5) for the indicator function shown in Figure 4.21.

Figure 4.32: Minimum bias tapers with gaps (K=6-10) for the indicator function shown in Figure 4.21.

Figure 4.33: Minimum bias tapers with gaps (K=11-15) for the indicator function shown in Figure 4.21.

Figure 4.34: Minimum bias tapers with gaps (K=16-20) for the indicator function shown in Figure 4.21.

Figure 4.35: Minimum bias tapers with gaps (K=21-25) for the indicator function shown in Figure 4.21.

Figure 5.1: From Olsen et al. (2010), a sketch of the near-Earth electromagnetic environment.

Figure 5.2: From Olsen et al. (2010), showing the orbital path of a satellite with the orbital inclination shown as a white arrow near the equator.

Figure 5.3: The magnetic local time of each satellite pass of the available high quality data. Data points with local times between 18h and 6h are shown in red.

Figure 5.4: Changes in the magnetic field model of CHAOS-4 between 2010 and 2001. Top panel: Percent relative change in intensity, $F$ at the surface of the Earth. Bottom panel: Relative difference (from the 2001 model) in the radial field at the surface of the Earth (in $\mu T$).

Figure 5.5: The empirical smoothed probability distributions of values of $e^{0}_1(t)$ from 2001 to 2009 when models of maximum spherical harmonic degrees of 50 (solid blue), 65 (broken red), 85 (dotted green) and 100 (dashed green) are used to remove the main field.
Figure 5.6: The empirical smoothed probability distributions of values of $i_1^0(t)$ from 2001 to 2009 when models of maximum spherical harmonic degrees of 50 (solid blue), 65 (broken red), 85 (dotted green) and 100 (dashed green). ................................. 172

Figure 5.7: Test of the rotation into magnetic dipole coordinates on a model of DGRF 2010 from $g_1^0, g_1^1, h_1^1$. The model coefficients for the dipole coordinate system are shown in green with geocentric coordinates shown in blue ...................................... 173

Figure 5.8: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of 18h to 23h. The equation and standard deviation of robust least squares line fits are shown on each plot .............................................. 176

Figure 5.9: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of 24h to 5h. The equation and standard deviation of robust least squares line fits are shown on each plot .............................................. 177

Figure 5.10: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of, from top left to bottom right: 6h, 7h, 10h, 11h, 16h, 17h. The equation and standard deviation of robust least squares line fits are shown on each plot. ........ 178

Figure 5.11: Time series of data segments used for pilot estimate of the transfer function. Slightly different data segments were tested. The first test included the large magnetic storm in late October 2003, shown in green. A shorter data segment which omitted the storm is shown in ...... 180

Figure 5.12: Spectral estimate of $e_1^0(t)$ (dark green) and $i_1^0(t)$ (dark blue) from the data segments shown in blue in Figure 5.11 shown with 95% jackknifed confidence intervals. The spectral estimates from the longer time series from Figure 5.11 is shown in light green. .... 182

Figure 5.13: Coherence between $i_1^0(t)$ and $e_1^0(t)$ from the time series (shown in blue in Figure 5.11) using 38 minimum bias tapers is shown in black with 95% confidence intervals. The coherence from 50 MB tapers applied to the longer time segment (from Figure 5.11) containing ...... 183

Figure 5.14: Response function estimate between $i_1^0(t)$ and $e_1^0(t)$ from the time series (shown in blue in Figure 5.11) using 38 minimum bias tapers is shown in blue with 95% confidence intervals. The response function estimate from 50 MB tapers applied to the longer time segment .... 184

Figure 5.15: The time between satellite passes with local times within 18h-6h on a logarithmic scale. The average sampling rate is shown as a red line at 0.0637 days ......................... 186

Figure 5.16: Filling break in sampling 3 days long (47 samples) in October 2004 using scaled predictions. The Akima spline interpolation for $e_1^0(t)$ ($i_1^0(t)$) is shown with dark green (light green) stars. The time series of scaled predictions of $e_D (i_D)$ is shown with magenta (red) .... 188
Figure 5.17: Response function estimate of the November 2003 - October 2004 data (same as the black line in Figure 5.14) from 42 minimum bias tapers is adjusted using equation 5.13 and shown in red. This adjustment lies entirely within the confidence intervals of the original.

Figure 5.18: Indicator function showing the breaks in sampling when breaks in sampling less than 100 samples long are interpolated using scaled predictions.

Figure 5.19: Average empirical transfer function of the minimum bias tapers with gaps (solid lines) when a threshold value, $T_2$ of 100 is used. These are compared to the average transfer function from a set of standard minimum bias tapers (broken lines).

Figure 5.20: Response function with noise correction, comparing MBG tapers with and without noise correction (shown in blue without the noise correction with 95% confidence intervals in light blue). Noise correction (red line) is only in amplitude.

Figure 5.21: Average empirical transfer function of the standard minimum bias tapers used for section averaging are shown in solid lines. These are compared to the average transfer function from the equivalent set of MBG tapers (broken lines).

Figure 5.22: Response function estimate from MBG tapers for $K=50$ (shown in black with grey confidence intervals) and from optimized section averaging with $K=51$ (shown in blue).

Figure 5.23: C-response estimate from 51 MBG tapers, the real portion is shown as a solid blue line and the imaginary portion is shown with a broken blue line. The noise correction is applied to this and shown in red. A section averaging approach using 15 MB tapers is shown in green.

Figure 5.24: (Top): Mean-squared coherence (MSC) shown between $i_{1}(t)$ and $e_{0}(t)$ plotted in blue for the 51 MBG tapers with a red line indicating the estimated relative noise ratio, $\psi$. The MSC estimated from 50 MB tapers using section averaging is shown in green. (Bottom): Jackknifed standard error estimated from the MBG.

Figure 5.25: C-response estimate from 15 MB tapers (applied using section averaging) shown in red (after the noise correction is applied) along with 95% confidence intervals for the uncorrected estimate are shown in grey. These are compared to (Olsen, 1999) (shown as purple stars).
### LIST OF TABLES

| Table 2.1: | Optimization problems which define the tapers. | 34 |
| Table 2.2: | Results of synthetic tests of 95% confidence intervals from prolate and minimum bias tapers with and without large breaks in sampling. | 46 |
| Table 2.3: | Algorithm for applying multi-tapers adapted for breaks in sampling. | 48 |
| Table 2.4: | Parameters to alter to improve time and frequency-domain properties of adapted tapers. | 54 |
| Table 5.1: | High precision magnetic satellites, adapted from Olsen and Stolle (2012); Olsen (1999); Kuvshinov (2008). Note*: Swarm mission consists of three satellites, one with a higher altitude orbit and two lower satellites with 1° separation in longitude. | 168 |
| Table 5.2: | Coefficients of the best fitting line relating $e_0^0(t)$ and $i_1^0(t)$ as a function of local time. | 175 |
ACKNOWLEDGEMENTS

I would like to thank my advisor, Catherine Constable. She has been a patient teacher. I have learned a great deal about science, research and professionalism from her.

I would like to thank my family for their support.

My former officemate, labmate and co-author Leah Ziegler has been a great sounding board for ideas and guidance.

Many of the students and postdocs who have passed through IGPP during my time here have been fantastic colleagues from whom I have learned a great deal. I’d especially like to thank Karen Weitemeyer and Jill Pearse.

I would like to thank Robert Parker for his helpful comments and suggestions on this work; Duncan Agnew for his useful suggestions and assistance; Kerry Key and William Coles for their time and input on my research.

Chapter 2, in full is a reprint of the material as it appears in Geophysical Journal International 2012, Smith-Boughner, Lindsay; Constable, Catherine. Spectral estimation for geophysical time-series with inconvenient gaps. 190: 1404-1422. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in Physics of the Earth and Planetary Interiors 2011. Smith-Boughner, Lindsay; Ziegler, Leah and Constable, Catherine. Changing spectrum of geomagnetic intensity variations in a fragmented 12My sediment record from the Oligocene, Volume 188, Issues 3-4, October 2011, Pages 260-269. The dissertation author was the primary investigator and author of this paper.

I would like to thank Alexey Kuvshinov for providing data used in Chapters 1, 3 & 4. This work was supported by NASA Earth Science Student Fellowship NNX08AU89H and National Science Foundation Grants EAR 0337712, EAR 0809709.
VITA

2006 B.Sc in Engineering
Queen’s University at Kingston

2016 Ph.D., Earth Sciences,
Scripps Institution of Oceanography,
University of California, San Diego

PUBLICATIONS


PRESENTATIONS

Smith, L.T., Constable, C. G. The inconvenient gaps: Timeseries analysis for incomplete data. Seafloor Electromagnetic Methods Consortium Meeting, La Jolla, CA, USA, 2010


**AWARDS AND HONORS**

2008-2011 NASA Earth and Space Science Student Fellowship,

2006 Keyser Fellowship, Queen’s University

2002-2006 Aiming for the Top Tuition Scholarship, Queen’s University

2002-2006 Albert H.Ruddell Award, Queen’s University
ABSTRACT OF THE DISSERTATION

Spectral Estimation Techniques for time series with Long Gaps: Applications to Paleomagnetism and Geomagnetic Depth Sounding

by

Lindsay Smith-Boughner

Doctor of Philosophy in Earth Sciences

University of California, San Diego, 2016

Professor Catherine G. Constable, Chair

Many Earth systems cannot be studied directly. One cannot measure the velocities of convecting fluid in the Earth’s core but can measure the magnetic field generated by these motions on the surface. Examining how the magnetic field changes over long periods of time, using power spectral density estimation provides insight into the dynamics driving the system. The changes in the magnetic field can also be used to study Earth properties - variations in magnetic fields outside of Earth like the ring-current induce currents to flow in the Earth, generating magnetic fields. Estimating the transfer function between the external changes and the induced response characterizes the electromagnetic response of the Earth. From this response inferences can be made about the electrical conductivity of the Earth.
However, these types of time series, and many others have long breaks in the record with no samples available and limit the analysis. Standard methods require interpolation or section averaging, with associated problems of introducing bias or reducing the frequency resolution.

Extending the methods of who adapt a set of orthogonal multi-tapers to compensate for breaks in sampling- an algorithm and software package for applying these techniques is developed. Methods of empirically estimating the average transfer function of a set of tapers and confidence intervals are also tested. These methods are extended for cross-spectral, coherence and transfer function estimation in the presence of noise.

With these methods, new analysis of a highly interrupted ocean sediment core from the Oligocene (?) reveals a quasi-periodic signal in the calibrated paleointensity time series at $2.5 \text{My}^{-1}$. The power in the magnetic field during this period appears to be dominated by reversal rate processes with less overall power than the early Oligocene. Previous analysis of the early Oligocene by detected a signal near $8 \text{My}^{-1}$. These results suggest that a strong magnetic field inhibits reversals and has more variability in shorter term field changes.

Using over 9 years of data from the CHAMP low-Earth orbiting magnetic satellite and the techniques developed here, more robust estimates of the electromagnetic response of the Earth can be made. The tapers adapted for gaps provide flexibility to study the effects of local time, storm conditions on Earth’s 1-D electromagnetic response as well as providing robust estimates of the C-response at longer periods than previous satellite studies.
Chapter 1

Introduction

1.1 Motivation

Many systems cannot be directly measured. Information about the system must be inferred from indirect sources. In the Earth, the velocity of the convective, conductive fluid of the Earth’s core cannot be directly measured, but one can measure a portion of the magnetic field generated from these flows. The poloidal component of the core magnetic field, after it passes through the partially conductive mantle, is measured at the surface of the Earth. The toroidal component of the core magnetic field does not exist outside of the conductive Earth.

The magnetic field generated by the Earth’s core varies on extremely long time scales. The field we measure at the surface of the Earth today is composed of the slow variations originating in the core and the magnetic storms and daily ionospheric variations caused by the sun in addition to the magnetic field of the crust (see Hulot et al. (2010) for a full review). The timescales of all these processes measured at the surface of the Earth span more than 18 orders of magnitude, as shown in Figure 1.1. To study the dynamics of the core, we cannot measure the inputs to this system but we can measure the outputs. Studying a long record of the variations in magnetic field intensity can provide insight into the core dynamics governing magnetic reversals and excursions (Constable et al., 1998; Constable and Johnson, 2005).

This is the case with many physical systems. Measurements of the natural variations of these systems are studied in either time or frequency. Estimating the amount
of power in the system as a function of frequency can provide insight into the various sources responsible for these changes. From these estimates, parametric models of the system can be made and the physics governing these changes can be probed.

For some physical systems, the input to the system can be measured as well as the output. Examining the relationship between these two signals in either the time or frequency domain provides an estimate of the influences the system has on the amplitude and phase of any input signal. From the time-domain response or the frequency domain transfer functions, parametric models connected to the physical properties of the system can be constructed.

Electromagnetic methods use external changes in the magnetic field and the internal magnetic fields induced by these changes to infer the electromagnetic response of the Earth (see Constable (2007) for a full review). Knowing the frequency-domain transfer function between the induced internal magnetic fields generated in the conductive regions of the Earth and the source fields responsible for that induction is very important. The inductive response is related to the electrical conductivity of the silicate mantle (Bailey, 1970; Parker, 1971; Weidelt, 1972). Response function estimates can be inverted (Parker and Whaler, 1981; Constable et al., 1987; Medin et al., 2007; Kelbert et al., 2008) to estimate the electrical conductivity of the Earth and from this, the com-

Figure 1.1: From Turner et al. (2007): Time variations of the geomagnetic field in relation to the age of the Earth and the Universe, indicating which are of internal and which are of external origin.
position of the mantle can be studied. Knowledge of the electrical conductivity of the mantle can also contribute to a more accurate picture of the rapidly changing magnetic fields on the core-mantle boundary (e.g. Pinheiro and Jackson (2008)).

Accurately estimating the characteristics of these systems, whether from just the output or from the response of the system to a measured input signal, requires a statistically robust method for estimating the power and cross-spectral power of a time series. Directly estimating spectra from the data using multi-taper methods can minimize the broadband bias (Thomson, 1982) or the local bias (Riedel and Sidorenko, 1995) in the spectral or cross-spectral estimate. These methods multiply a time series by a set of orthogonal data tapers, then estimate the power or cross-spectral power of the Fourier transform, producing a nearly statistically independent set of estimates with minimal bias. Averaging these singly-tapered estimates creates a robust estimate of the power while reducing the variance. Both cross-spectral and power spectral estimation typically require a complete series of discrete evenly sampled measurements.

Many observational time series are missing data samples. With measurements of the long-term variability (timescales of millions of years) of the Earth’s magnetic field, some of the longest records available are missing significant segments of data (Hartl et al., 1993). On timescales of hours to years, observations of the changes in the external magnetic field and the induced response fields also have breaks in the record. When such records come from low-Earth orbiting magnetic satellites, the orbital paths of the satellites travel through regions where the structure of the external magnetic fields is very complex. Any fields induced will be complicated as well and difficult to model.

Breaks in the continuity of time series limit the resolution of the analysis. Short breaks in sampling can be interpolated while long segments of missing data have usually required splitting the record into separate pieces. Each data segment is analyzed separately and the results combined. This reduces the amount of information that can be inferred about the system, especially at low frequencies. The length of the record analyzed controls the spectral resolution i.e. the maximum number of "frequency bins" used in the analysis (for these discrete measurements, a single frequency is associated with each bin.) With a longer time series, there are more bins available - a smaller range of frequencies being represented by each value.
Robust methods of estimating the power and cross-spectral density from time-series with long breaks in sampling would be extremely useful across a variety of physical sciences. The aim of this thesis is to develop and test such techniques and use them to study two important geophysical problems. The power spectral density techniques developed are used to study extremely long-term variations in the intensity of the magnetic field. The techniques developed for estimating the cross-spectral power from time series with missing data are used in geomagnetic depth sounding to investigate the global one-dimensional response function of the Earth at long periods.

### 1.2 Outline

The first step towards developing robust cross-spectral methods for data with long interruptions in sampling is to tackle the simpler problem of spectral estimation. Breaks in the continuity of time series are incredibly common, especially in the physical sciences. An algorithm for estimating the power spectral density of intermittently sampled time series would have a variety of applications. Once these methods have been validated, we pursue the more complicated issue of transfer function estimation.


In Chapter 2, I extend these techniques for both prolate (Thomson, 1982) and minimum-bias (Riedel and Sidorenko, 1995) multi-tapers adapted for intermittent sampling and include empirical estimates of the average frequency domain characteristics of a set of adapted multi-tapers. Methods of computing reliable confidence limits of spectra at all frequencies, computed with both standard and adapted multi-tapers are also investigated. An algorithm for estimating the power spectral density of a time series with breaks in sampling using judicious interpolation and adapted multi-tapers is presented. These techniques are tested on three geophysical data sets; a record of the change in the
length of day spanning 50 years, a 5 year long model of the external dipolar changes to the Earth’s magnetic field (Kuvshinov and Olsen, 2006) and a paleomagnetic record of relative paleointensity spanning nearly 12 million years (Tauxe and Hartl, 1997).

Using the techniques developed above, the power spectral density of an extremely fragmented deep sea sediment core from 22.74–34.77 million years ago (Ma) is analyzed in Chapter 3. This time series from the Oligocene is one of the longest quasi-continuous records available. Several long breaks in sampling, particularly one of nearly 1.5 My (from 27.4 Ma-29 Ma) limited previous spectral analysis by Constable et al. (1998) to only the older portion. Using the methods of Ziegler et al. (2011) the record is calibrated to absolute field intensity and spectral estimates from prolate tapers adapted for gaps (PRG) in sampling are computed for the entire 12 My, only the younger half and then only the older half. The robustness of the spectral features and any possible bias from the PRG tapers are investigated.

In Chapter 4, I extend the application of multi-tapers adapted for long breaks in sampling to cross-spectral, coherence and transfer function estimation. These techniques are validated using synthetic examples. Data are removed from the synthetic time series to create sampling breaks allowing tests of the confidence intervals and properties of the amplitude and phase response of the transfer functions from adapted tapers.

Chapter 5 outlines the application of these techniques to estimating the global 1-D electromagnetic response of the Earth to changes in the ring-current. Vector measurements of the magnetic field computed by CHAMP from 2001-2010 are used. Models of the external dipolar fluctuations, $e_0^E(t)$, and the change in the internal dipole field, $i_0^I(t)$, are estimated from high quality, nightside measurements. The resulting time series have several long breaks in sampling. The adapted minimum bias tapers from Chapter 4 are used to estimate the transfer function between the time series of model coefficients. Several response function estimates are produced using only data near local midnight and only data from magnetically quiet times. These various response function estimates are transformed to C-response values and compared to results from previous satellite studies.

The remainder of this introductory chapter is devoted to background information on the two geophysical problems of interest, namely long term variations in the intensity
of the Earth’s magnetic field and estimating the global one-dimensional electromagnetic response of the Earth. Background is given on the sources of Earth’s magnetic field and the variety of measurements used to study it. I also include brief discussions of how the spatial structure of the magnetic field is modeled and of the temporal variations in the primary magnetic field generated in the core. Geomagnetic induction is discussed from first principles, followed by a review of the methods used to estimate the global electromagnetic response of the Earth from satellite measurements.
1.3 Earth’s magnetic field

Most of the magnetic field measured near Earth is dipolar, like that of a bar magnetic tilted relative to the rotation axis, but with significant non-axial dipole contributions and a great deal of smaller scale structure. This field is primarily generated by the electrically conductive, convecting fluid in the core, 2900 km below the surface of the Earth. In addition to the dynamo field, Earth’s magnetic field includes contributions from the lithospheric magnetic field and ionospheric, magnetospheric and oceanic current systems. These current systems are quite dynamic, inducing magnetic field changes on very short timescales while the core field changes much more slowly, with the largest changes coming from reversals of the dipole field (see Figure 1.2). Because changes in the Earth’s magnetic field occur over a wide variety of temporal and spatial scales, many different types of data are used to study it. Characterizing the magnetic field in time and space and examining how the power of the field changes as a function of wavelength gives insight into the turbulent, non-linear processes driving the magnetic field. I provide a brief description of these changes and how they are measured; a thorough review can be found in Hulot et al. (2010).

Reversals of the dipole field occur from time to time and the interval between polarity reversals varies widely. Early statistical descriptions of the likelihood of a magnetic field reversal at any point in time were related to average reversal rate and considered independent of long the magnetic field has been a particular polarity; with these assumptions magnetic reversals were modeled as a Poisson process (Cox, 1968).

In reality, magnetic field reversals are typically preceded by a decrease in field intensity (Ziegler and Constable, 2011) and followed by a slow rebound in field strength. The time between magnetic reversals vary from only tens of thousands of years (cryptochrons) to more than 20 million years (a phenomenon referred to as a superchron). Short-term (a few thousand years) dips in the latitude (lower than 45°) of the dipole axis worldwide, known as geomagnetic excursions, occur sporadically. An excursion is difficult to identify in paleomagnetic records because the departures of the virtual geomagnetic pole must be seen in several locations to confirm a global feature, rather than a change in the local non-dipole field.

In addition to the dipolar field, the Earth’s magnetic field has dynamic, smaller
scale structure; models of this non-dipolar structure only span the last 10 000 years. Using paleomagnetic data from a variety of sources, Korte et al. (2011) modeled time variations in the vector magnetic field over the past 10 000 years. The average radial magnetic field component at the core-mantle boundary (CMB) from CALS10k.1 reveals some recurring spatial structures, especially at high latitudes. These persistent flux lobes have existed for most of the last 5000 years (Korte and Holme, 2010). Similar spatial structures at the CMB, are seen in the GUFM model of Jackson et al. (2000) from historical and modern data. These models are only able to represent features of the magnetic field that have wavelengths larger than 10 000 km (CALS10K) and 3000 km (GUFM).

The dipole magnetic field contains most of the power of the magnetic field seen at the surface of the Earth. Smaller scale contributions have much less power, when averaged over the surface of the Earth. The power from a spherical harmonic with degree $l$ and order $m$ distributed over a surface of radius, $r$ is defined using the Schmidt normalization in equation 1.1.

$$R_l(r) = (l + 1)(a/r)^{2l+4} \sum_{m=0}^{l} \left\{ [g_l^m]^2 + [h_l^m]^2 \right\}$$

(1.1)

In this equation, $g_l^m, h_l^m$ are coefficients which model the primary magnetic field of the Earth at one point in time using measurements of the magnetic field and $a$ is the radius of the Earth. Figure 1.4 shows these contributions as a function of wavelength and spherical harmonic degree from recent magnetic field models (Olsen et al., 2009; Maus et al.,
Figure 1.4: Figure from Olsen et al. (2010): Spatial power spectrum of the geomagnetic field at the Earth’s surface. Black dots represent the spectrum of a recent field model (Olsen et al., 2009; Maus et al., 2008). Also shown are theoretical spectra (Voorhies et al., 2002) for the core (blue) and crustal (magenta) part of the field, as well as their superposition (red curve)

This spatial power spectrum, $R_n(r)$, is also known as the Lowes-Mauersberger spectrum and represents the poloidal contributions and the crustal contributions to the magnetic field and depending on how the field modeling is done, may be contaminated by external field and induced contributions at higher spherical harmonic degrees.

Most of the large scale structure (below spherical harmonic degree 15) of the magnetic field measured at the surface comes from the core, overwhelming the large scale magnetic field of the crust. Above spherical harmonic degree 15, the magnetic fields generated in the crust and lithosphere (Maus et al., 2008) mask any smaller scale core field structure. These magnetic fields are generated primarily from remanent magnetization locked in the oceanic and continental crust as it cooled.

The magnetic field measured at the surface of the Earth also includes contributions from sources outside of the solid Earth including the ionosphere and magnetosphere (Langel and Estes, 1985). The largest magnetospheric contribution comes from the ring-current, a band of charged particles at a distance of 2-6 Earth radii ($R_e$), generating a predominantly westward current. This current generates a largely dipolar magnetic field which is parallel to the horizontal magnetic field at the equator. This magnetic field induces currents to flow in the conductive regions of the Earth. Changes in the
ring-current are characterized using various magnetic indices. The disturbed storm time index or Dst (Sugiura, 1964) measures the strength of the equatorial ring current and associated induced currents. This is done by measuring the change in the horizontal magnetic field near the equator at four low latitude magnetic observatories distributed around the equator.

The largest ionospheric magnetic field is induced by currents flowing in the daytime ionosphere, known as the daily solar-quiet signal, Sq. Despite the name, these currents are also enhanced by magnetic storms. Daily heating of the atmosphere from the sun energizes the charged particles in the ionosphere generating these currents. Changes in this magnetic field also induce currents to flow in the conductive regions of the Earth (Kuvshinov et al., 2007). Induction in the ocean from Sq and the ring current generate small magnetic fields, adding to those generated from oceanic and tidal currents.

1.4 Measuring the Magnetic Field

The strength of the magnetic field can also be inferred from measurements of the magnetization stored in igneous rocks. The magnetic minerals in these rocks preferentially aligned with the ambient magnetic field as they cooled, locking in the magnetization induced by the ancient magnetic field. Using the Thellier-Thellier method (Thellier, 1951) on samples of volcanic or igneous rock the intensity of the ancient magnetic field can be measured. If the original orientation and paleolatitude of the lava flow as it cooled is known, then estimates of the declination and inclination of this stored field can also be recovered and used to estimate the virtual geomagnetic pole. This VGP represents the position of the magnetic dipole axis which fits the data at the given paleolatitude of the lava flow. The theory and practice of measuring paleointensities are well described in Tauxe and Yamazaki (2007). However, these measurements represent the magnetic field at only one place, often with a large uncertainty in the time. Absolute paleointensities are also estimated from fired archeological materials and can be quite accurately dated.

Measurements of the changes in the direction and relative intensity of the paleomagnetic field are also recorded in sediments. The magnetic minerals trapped in the
sediments align with the ambient magnetic field as they slowly dewater and compact. Because magnetization contained in a sample of sedimentary material depends on the type and concentration of magnetic minerals in the sediment, and the sedimentary environment, the absolute magnetic field intensity cannot be recovered - the theory is still a subject of active research. However, measurements of the relative paleointensity in the sediment can provide a time series of the relative changes in the magnetic field. Because sediments accumulate quickly and somewhat steadily, the relative paleointensities and field directions from sediments form a time series of the paleofield with some records spanning more than 10 million years. For example, the nearly 12 million year time series from Hartl et al. (1995) is re-analyzed in Chapter 2 using the techniques developed in this thesis.

1.5 Modeling the Magnetic Field

In a region free of current sources, the spatial structure of the magnetic field can be modeled as a potential field (for a full review of the mathematical properties involved in field modeling see Sabaka et al. (2010)). Using Helmholtz’s theorem, the magnetic field, $B$ can be written as the gradient of a scalar potential, $V$ and the curl of a vector potential, $A$, as:

$$B = -\nabla V + \nabla \times A \tag{1.2}$$

The divergence of the magnetic field is zero because no magnetic monopoles have been found. Therefore, in a region free of currents, this simplifies to Laplace’s equation:

$$\nabla^2 V = 0 \tag{1.3}$$

The magnetic field must satisfy: $B = -\nabla V$. From Laplace’s equation, $V$ can be expressed as the sum of an internal magnetic field with coefficients $g_l^m, h_l^m$ and an external
magnetic field with coefficients $q^m_l, s^m_l$:

$$V(r, \theta, \phi) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} \left\{ \left( \frac{a}{r} \right)^{l+1} \left( g^m_l \cos m\phi + h^m_l \sin m\phi \right) + \left( \frac{r}{a} \right)^{l} \left( q^m_l \cos m\phi + s^m_l \sin m\phi \right) \right\} P^m_l (\cos \theta),$$

where $a$ is the radius of the Earth and $P^m_l (\cos \theta)$ is Schmidt quasi-normalized associated Legendre function.

### 1.6 Studying the variability in time

Time series of magnetic field measurements are vital to understanding the magnetic field of the Earth. Because of the huge range of time scales over which the magnetic field changes, shown in Figure 1.1, snap shots of the current magnetic field only provide insight into the short term dynamics of the magnetic field. Most of the power in the magnetic field is in the largest scale structure- dipole and quadropole contributions, seen Figure 1.4. The amplitude of the magnetic field as a function of frequency, or period, shown in Figure 1.2 from Constable and Constable (2004a), shows the largest variations occur as very low frequency or long period changes.

Studying the power in a signal as a function of frequency, power spectral density analysis allows one to investigate the physical systems which drive the magnetic field. This can be used to test core dynamo models and examine the rates of reversals, cryptochrons, excursions, secular variation, and the energy added to the magnetic field by the solar wind and magnetic storms. Understanding how the power is distributed provides insight into different types of magnetic field behavior - why the field does not reverse during superchrons and why does the magnetic field reverse often at other times. Viewing the magnetic field in these terms allows us to investigate the physics behind the changes and test our models. For example, looking at the field this way, one can examine how well the power in the magnetic field matches the power spectral density of a Poisson distribution, controlled by the average reversal rate.

Because long, high resolution absolute or relative paleointensity time series are difficult to obtain, little is understood of what drives changes in magnetic field intensity
over millions of years or why the reversal rate changes so dramatically. Constable et al. (1998) studied one of the longest time series available (Hartl et al., 1993), a record of relative paleointensity from 22-34 million years ago and found much of the variance in the during 30-34 million years ago was concentrated near frequencies of 8 cycles per million years. However, this unique record was marred by long breaks in the time series when no high quality measurements of the paleointensity were available.

Breaks in the time series occur frequently and are particularly troubling in geomagnetic and paleomagnetic research. Data sets are difficult to obtain and often unique. Some paleomagnetic samples do not provide meaningful results because the sample is too disturbed for sampling (sediments), or show evidence of being altered by paleomagnetic experiments, or have been overprinted by a more recent magnetic field (described in Tauxe and Yamazaki (2007)). Time series measurements of the modern magnetic field also have breaks in sampling. This could be due to a failure in the instrument that makes the measurement. Sometimes samples are removed from a time series because of suspected contamination from noise, geomagnetic storms, induced currents, or other sources of non-stationarity.

Gaps in a time series record occur in many physical sciences. Estimating the power spectral density (PSD) usually requires Fourier transforming an equally spaced time series. The length of the time series controls the maximum frequency resolution obtainable from the time series. When breaks occur the time series must be analyzed using section averaging or interpolated. Section averaging restricts the information available about the power in the signal at long periods. If an interpolated piece does not have the same spectral properties as the signal of interest, the power spectral estimate will be biased.

Methods for increasing the resolution and reducing the uncertainty of spectral estimates with inconvenient breaks in sampling can improve our knowledge of the long period characteristics of the system under investigation, in this case, the magnetic field. With such techniques, long paleomagnetic records with breaks in sampling can be re-analyzed.
1.7 Geomagnetic Induction

As mentioned above, changes in the magnetic field external to the Earth induce currents to flow in the conductive regions of the Earth. In turn, these currents generate a magnetic field. Because the external contributions to the magnetic field of the Earth change much more rapidly than the main field (Figure 1.1) the inducing and induced magnetic field can be separated from the primary magnetic field. With this separation, the electric and magnetic fields generated by inductive effects can be studied in one, two or three dimensions (see Constable (2007) for a review.)

Electromagnetic methods take advantage of these connections between the magnetic and electric fields at the surface of the Earth to investigate the electrical conductivity of the crust, lithosphere and mantle. Knowing the electrical conductivity of the crust and mantle can constrain: porosity, melt fraction, water content and composition. Combined with seismic methods, electromagnetic surveys can provide estimates of the local, regional and global structure of the Earth.

Most electromagnetic methods study the relationship between the source and induced fields in the frequency domain using electric and magnetic diffusion equations and some assumptions about the fields involved.

The magnetotelluric method uses a time series of measurements of orthogonal electric and magnetic fields to estimate the inductive response of the Earth as a function of frequency. Several assumptions are made about the electric and magnetic fields that simplify the governing equations. A more complete description of these derivations can be found in Constable (2007) and the references therein.

Using Faraday’s law:

\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \]

where \( \mathbf{B} \) is the magnetic field in Tesla, \( \mathbf{E} \) the electric field in V/m. It is assumed that there is very little polarization of the crust or mantle and that measurements are being made in a region free of current sources. With these assumptions, Ampere’s Law simplifies to:

\[ \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \]
where \( \mathbf{J} \) is the displacement current. Using the fact that \( \nabla \cdot \nabla \times \mathbf{A} = 0 \), yields:

\[
\nabla \cdot \mathbf{J} = 0. \tag{1.8}
\]

Because there are no magnetic monopoles:

\[
\nabla \cdot \mathbf{B} = 0. \tag{1.9}
\]

Taking the curl of Faraday’s law yields gives the relationship:

\[
\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \left( \frac{\partial \mathbf{B}}{\partial t} \right) \tag{1.10}
\]

where \( \mu_0 \) is the permittivity of free-space. If it is assumed that the magnetic field variations are of the form \( \frac{\partial \mathbf{B}}{\partial t} = -i\omega \mathbf{B} \). Substituting this into equation 1.10 and applying a vector identity yields:

\[
-\nabla^2 \mathbf{E} + \nabla(\nabla \cdot \mathbf{E}) = -\nabla \times (i\omega \mathbf{B}) \tag{1.11}
\]

Simplifying and substituting in Ampere’s law and Ohm’s law, \( \mathbf{J} = \sigma \mathbf{E} \) (\( \mathbf{J} \) is the current density in \( \text{A/m}^2 \), \( \sigma \) the conductivity in Siemens/m):

\[
\nabla^2 \mathbf{E} = \mu_0 \sigma \omega i \mathbf{E} \tag{1.12}
\]

With the same assumptions, the magnetic field can be expressed as:

\[
\nabla^2 \mathbf{B} = \mu_0 \sigma \frac{\partial \mathbf{B}}{\partial t}. \tag{1.13}
\]

To gain some intuition into how these fields interact, consider an ideal simple geometry: an infinite half space with uniform conductivity, \( \sigma_0 \) and an insulator above the half-space. A plane wave magnetic field \( \mathbf{B}(\omega) \), at frequency \( \omega \), would induce an electric field, \( \mathbf{E}(\omega) \). No currents flow above the half-space - it is an insulator and the magnetic field is continuous. For this infinite half-space, \( B_z \) and \( B_y \) must be zero at the
surface. Using this boundary condition, equation 1.13 is solved by:

\[ B(z) = \hat{x}e^{-z(1+i)/z_0} \left[ \cos(\omega t) - i \sin(\omega t) \right], \]  

(1.14)

where \( z_0 \) is defined as:

\[ z_0 = \sqrt{\frac{2}{\mu_0 \omega \sigma_0}} = \sqrt{\frac{1}{\sigma_0 \mu f}}. \]  

(1.15)

This quantity, \( z_0 \) is known as the skin-depth. It is the depth at which electromagnetic waves of a given frequency are attenuated by a factor of \( 1/e \) as they penetrate a uniform half space with conductivity \( \sigma \). The skin-depth depends on the conductivity of the half-space and the frequency at which the signal changes, this can be used to estimate the inductive length scale of a signal and the depth to which the signal is likely to penetrate.

In a one-dimensional layered Earth, the inductive characteristics of the Earth at frequency \( \omega \) can be expressed as the ratio of the Fourier transforms of the electric and magnetic field at \( \omega \).

\[ C(\omega) = \frac{-E_y(\omega)}{i\omega B_x(\omega)} \]  

(1.16)

known as the C-response (Schmucker, 1970). It is also referred to as the admittance, as it is a complex quantity that represents how easily current may flow for a signal at a frequency of \( \omega \). For the infinite uniform half-space described above this simplifies to:

\[ C(\omega) = \frac{z_0}{\sqrt{2}} \]  

(1.17)

The C-response is a useful quantity. Different definitions of the C-response are used by various electromagnetic methods, discussed below. From any definition, the C-response at several frequencies can be used to interpret data. The C-response can be inverted for models of conductivity at depth. If the C-response is known precisely at all frequencies, then only one continuous conductivity model will fit the data (Bailey, 1970). In reality, C-response estimates have some uncertainty associated and are only available for a finite number of frequencies (Parker, 1971). Regularized inversion can be used to select an appropriate conductivity model (Parker, 1980).

The depth at which conductivity estimates can be made depends on the lowest frequency estimates of the C-response available. The magnetotelluric method typically
focuses on estimating conductivity structures within the upper 500 km of the mantle, assuming a plane wave source field.

If the source field of interest has more complex structure or the conductivity varies laterally, the horizontal spatial gradient method may be used (Schmucker, 1970). This method uses this form of the C-response:

\[ C(\omega) = \frac{-B_r(\omega)}{\nabla_{\tau} \cdot \mathbf{B}_\tau(\omega)} \]  

(1.18)

where \( \nabla_{\tau} \cdot \mathbf{B}_\tau \) represents the horizontal part of the divergence. This method requires a dense model of the time-varying spatial structure of the magnetic field. To accurately estimate the spatial gradients of the magnetic field, a set of closely spaced measurements of the horizontal \( B_\theta \) and vertical \( B_r \) is needed. These measurements are fit to a series of analytical functions; these functions are differentiated to estimate the spatial gradient of the magnetic field (Olsen, 1998).

### 1.8 Estimates of Earth’s electromagnetic response

Estimates of the conductivity of the Earth on a global scale use one of the ring-current-measured by Dst, the Sq current system, or the polar or equatorial electrojet as the inducing source field. Each source field has different source structure and slightly different variability with frequency. An extensive review of these considerations and methods can be found in Kuvshinov (2012).

Using the potential field description from section 1.5, the internal and external portions of the magnetic field can be modeled. The slowly-varying, internal field can be separated from the rapidly varying external components of the magnetic field. Once separated, a time-series of the complex internal and external coefficients at each spherical harmonic, \( Y_l^m(\theta, \phi) \) can be estimated as shown in Equation 1.19.

\[ V(r, \theta, \phi, t) = \sum_{l=1}^{L} \sum_{m=-l}^{l} \left\{ \left( \frac{a}{r} \right)^{l+1} i_l^m(t) + \left( \frac{r}{a} \right)^{l} e_l^m(t) \right\} Y_l^m(\theta, \phi) \]  

(1.19)

In this equation, \( e_l^m(t) \) \( (i_l^m(t)) \) represents the fully normalized, time-varying coefficient
The main attraction of the finite-difference (FD) approach is its relatively simple numerical implementation, especially when compared to other approaches, as well as its effectiveness. One then works with a sparse matrix after discretization of the second-order differential equation (either with respect to electric or magnetic fields) deduced from Maxwell's equations (1). The drawback is that the convergence of an iterative solution of the discretized form of the equation is not guaranteed even after pre-conditioning. Note that iterative solution is the only choice for large scale problems. The large computational volume needed for the decay (or stabilization) of EM fields at the boundaries is another disadvantage.

In the finite-element (FE) method, the whole modeling volume is decomposed into elementary volumes (such as prisms, tetrahedrons or volumes of more complex shapes) that specify the geometry of the conductivity model. Accordingly, the electric field (or its potentials) is decomposed via some basic functions and the coefficients of decomposition are sought using the Galerkin method. As with the FD approach, one again arrives at a sparse system of linear equations. The drawbacks of the FD approach generally also apply to the FE approach. The main attraction of the FE solutions is that they are believed to be better able to account for the geometry of the conducting anomalies than other approaches. This attraction is counterbalanced by a nontrivial and usually time-consuming construction of the finite elements themselves.

In spectral methods the quantities of interest are expanded in spherical harmonics in order to represent their lateral variations. Spectral solutions work well with rather smooth

![Figure 1.5: From Kuvshinov (2008): \(|Q_n^{(r)}|\) for different periods as a function of degree n. Solid lines: \(|Q_n^{(r)}|\) for a 1-D mantle conductivity model without oceans. Dashed lines: \(|Q_n^{(r)}|\) for a 1-D model overlaid by a uniform ocean of 15,000 S. For comparison, the dotted line in the figure shows an upper limit, 1, for \(|Q_n^{(r)}|\) which corresponds to a perfectly conducting Earth.](image)

of the uniform external magnetic field (internal dipolar) magnetic field. From these time series, the frequency-domain response function of spherical harmonic degree \(l\) is estimated as follows:

\[
Q_l(\omega) = \frac{I_l^m(\omega)}{E_l^m(\omega)}
\]  

(1.20)

Examining \(Q_l(\omega)\) at several different periods shows the relative amount of the power of the source field induced in the Earth, shown in Figure 1.5 from Kuvshinov (2008). The strongest and largest inductive signal occurs at long-periods, produced by variations in the ring current, are attributed to variation in the magnetic field generated by the ring current. Shorter period and smaller spatial scale inductive fields are generated by Sq- the solar quiet current system, related to solar heating of the ionosphere on the dayside of the magnetosphere. The equatorial and polar electrojets induce magnetic fields with very short spatial scales and much less relative power. Because of these differences, each source field assumption requires a slightly different approach.

The magnetic field associated with the solar-quiet current system has complex structure and requires several spherical harmonics to model appropriately. Because of
the complex structure associated with the source and induced fields, the horizontal spatial gradient method described in Equation 1.18 is used.

Using the ring-current as the source field and the induced response provides simple geometry and a strong signal. The ring-current creates a magnetic field, usually in the opposite direction to the Earth’s magnetic field, usually approximated by $P_0^1$ structure and induces a response that also has $P_1^0$ structure. The external and internal field are estimated, in geomagnetic coordinates, from the rapidly varying portion of the magnetic field. The response function of the Earth to the variations in the ring current at a frequency $\omega$ is given by:

$$Q(\omega) = \frac{i_0^1(\omega)}{e_0^0(\omega)}. \quad (1.21)$$

This admittance can be transformed from a spherical Earth to a flat-Earth C-response as used in magnetotellurics (Weidelt, 1972),

$$C(\omega) = \frac{a}{2} \frac{1 - 2Q_1(\omega)}{1 + Q_1(\omega)}. \quad (1.22)$$

The global depth sounding technique also uses a $P_0^1$ source field and assumes a 1-D conductivity and uses this form of the C-response:

$$C(\omega) = -\frac{a\tan \theta_d}{2} \frac{B_r(\omega)}{B_\theta(\omega)}, \quad (1.23)$$

where $\theta_d$ is the geomagnetic co-latitude. Because of this similarity, the potential method is also referred to as geomagnetic depth sounding.

Estimates of the local response function use data from magnetic observatories (Olsen, 1998) or inferred from voltages induced in oceanic cables (Utada et al., 2003). However, because magnetic observatories are sparsely distributed only over the continents, global response function estimates from these data are biased towards continental values. Truly global estimates of the response function of the Earth require data from low-Earth orbiting magnetic satellites (Constable and Constable, 2004b; Kuvshinov and Olsen, 2006; Velímský and Martinec, 2006).

Low-Earth orbiting satellites have near uniform coverage of the Earth. These satellites travel at 7-8 km/s in polar orbits and measure the time-varying magnetic field.
This makes it difficult to separate the spatial structure of the field from the time variations. These magnetic satellites can precess through all local times over the course of a few months. At some times, both the ascending and descending orbital paths travel through regions influenced by several different external magnetic fields. When using the variations in the ring-current to study the inductive response of the Earth, including data from regions directly influenced by the dayside ionospheric current systems in addition to the ring current introduces bias in the response function estimate (Balasis and Egbert, 2006).

Previous studies of the global 1-D response from satellite data were limited by the length of the available satellite time series. With 5 months of vector measurements from the MAGSAT satellite as it traveled through the dawn and dusk, Constable and Constable (2004b) used multi-taper techniques to estimate the response function to periods of 100 days. The CHAMP, Ørsted and SAC-C satellite missions greatly extended the length of time series available. Using 5 years of data from three magnetic satellites, each with slightly different orbits, Kuvshinov and Olsen (2006) estimated the response of the Earth out to periods of 145 days using section averaging.

Frequency domain methods of estimating the electromagnetic response function of the Earth use cross-spectral techniques to estimate the transfer function between the induced magnetic fields and the proposed source. The length of the available time series reduces statistical reliability of the long-period characteristics of the response. However, in order to estimate the conductivity of the deep Earth, the source signal must have power at very long periods.

From the skin-depth relationships (equation 1.15) developed for infinite half-spaces, a rough estimate of the required period of the signal can be made. If the lowermost mantle has a conductivity of 10 S/m then in order for a signal to penetrate below 2000 km before decaying by $1/e$ the period of the inducing signal should be several years or longer.

Now that over 10 years of data from CHAMP are available, estimates of the global 1-D response function of the Earth can be made at longer periods than previous estimates. However, when selecting high quality data from the nightside orbital passes, breaks occur in the time series. Depending on how magnetic local times are selected to
minimize the influence of the spatial variations in the ring current, these breaks can last a few weeks.

A method for computing the response function of the Earth using time series with large breaks in sampling and accurate empirical estimates of uncertainty should produce a response function estimate of the Earth with lower, more accurate uncertainties and more information about the low frequency characteristics. These methods allow one to estimate the long period response of the Earth using data from a narrow range of local times. The influence of the asymmetric structure of the ring-current magnetic field on the response function can be studied. Response functions from long times series of appropriately selected data can be used to extend the depths to which we can model the conductivity of the deep Earth and work towards a more detailed understanding of the deep Earth.
Bibliography


Chapter 2

Spectral Estimation for Geophysical Time Series with Inconvenient Gaps

Abstract

The power of spectral estimation as a tool for studying geophysical processes is often limited by short records or breaks in available time series. Direct spectral estimation using multi-taper techniques designed to reduce variance and minimize leakage can help alleviate the first problem. For records with gaps, systematic interpolation or averaging of multi-taper spectra derived from record fragments may prove adequate in some cases, but can be cumbersome to implement. Alternatively, multi-tapers can be modified for use in direct spectral estimation with intermittently sampled data. However, their performance has not been adequately studied. We investigate reliability and resolution of techniques that adapt prolate and minimum bias multi-tapers to accommodate the longest breaks in sampling, comparing the tapering functions (referred to as PRG or MBG tapers) with the standard prolate and minimum bias tapers used for complete data series, and with the section-averaging approach. Using a synthetic data set, we test both jackknife and bootstrap methods to calculate confidence intervals for PRG and MBG multi-taper spectral estimates and find the jackknife is both more accurate and faster to compute. To implement these techniques for a variety of data sets, we provide an algorithm that allows the user to balance judicious interpolation against the use of suitably adapted tapers, providing empirical measures of both bias and frequency resolution for candidate sets of tapers. These techniques are tested on diverse geophysical data
sets: a record of change in the length of day, a model of the external dipole part of the geomagnetic field produced by the magnetospheric ring current, and a 12 My long irregularly sampled relative geomagnetic paleointensity record with pernicious gaps. We conclude that both PRG and MBG tapers generally perform as well as, or better than, an optimised form of the commonly used section averaging approach. The greatest improvements seem to occur when the gap structure creates data segments of very unequal lengths. Ease of computation and more robust behavior can make MBG tapers a better choice than PRG except when very fine-scale frequency resolution is required. These techniques could readily be applied for cross-spectral and transfer function estimation and are a useful addition to the geophysical toolbox.

2.1 Introduction

Power spectral density estimation is used to determine how the variance in a signal is distributed across the frequency domain, often enabling inferences about the nature of underlying physical processes. Spectral techniques have been used with great success in many physical sciences applications, especially in geophysics. Their use is widespread in seismology, for example, in analyzing the frequency content of seismic waveforms to identify source parameters of earthquakes (Prieto et al., 2004), and to identify the normal modes of oscillation of the earth which provide information about the structure of the deep Earth (Ritzwoller et al. (1986)). Similar methods find applications in helioseismology (e.g. Fodor and Stark (2000); Christensen-Dalsgaard (2002)). In geomagnetism, induction problems are naturally analyzed in the frequency domain, because of the frequency dependence of electromagnetic skin depth (e.g. Kuvshinov and Olsen (2006); Chave et al. (1987); Chave and Thomson (2004); Egbert and Booker (1986)). Similarly, internal and external geomagnetic field variations are often separated on the basis of their frequency content which depends on the various physical processes involved (Olsen, 2007). On longer timescales changes in the geomagnetic dipole moment are separated into paleosecular variation (Guyodo and Valet, 1999), excursional and reversal processes and long term changes in reversal rate (Constable and Johnson, 2005; Ziegler and Constable, 2011).
Spectral methods are often limited in resolution by the length of time series available, restricting inferences about physical processes at long periods or long spatial wavelengths. Missing data or breaks in the record are a common problem and can be a serious limitation on the length of the time series. Breaks in sampling can occur for a variety of reasons, common examples being the failure of a data recorder, or contamination by sporadic noise sources. These gaps in the time series often occur at random, resulting in variable lengths for both data segments and the breaks in sampling. For many geophysical data series, the missing samples cannot be recovered or remeasured, because the original experiments were unique or the observations were very difficult to acquire. Care must be taken when considering the method of spectral estimation used.

There are both direct and indirect methods for estimating the power spectral density of a time series. Indirect methods generally work with the Fourier transform of the autocorrelation function (Blackman and Tukey, 1958) or suppose that a parametric model describes a stochastic process (Priestley, 1981), and can be inefficient or require unwarranted assumptions about the properties of the time series. We therefore choose to use direct spectral estimation, in particular multi-taper methods (Percival and Walden (1998); Thomson (1982); Riedel and Sidorenko (1995)), which are based on modifications of the periodogram technique, and are designed to minimise bias and reduce variance in the resulting estimates.

Direct spectral methods generally require a continuous, evenly sampled time series, something that is not always available, and several strategies have evolved for dealing with breaks or uneven sampling in the data record. When only a few samples are missing, breaks are often filled by careful interpolation, designed to introduce minimal bias into the spectral estimate. For long breaks in a time series, filling these inconvenient gaps becomes problematic, requiring possibly unjustified assumptions about the process under investigation, and generating potentially unquantifiable bias. An alternative approach known as section averaging splits the time series into several continuous pieces which are analysed separately and then averaged to produce the final spectral estimate. The shortened series reduce frequency resolution and may limit understanding of long period processes.

For time series with periodic components, several iterative methods exist for es-
imating the spectrum from a series with long breaks in sampling. For example, the CLEAN algorithm of Roberts et al. (1987) estimates both the periodic components in a spectrum and the bias from sampling using an iterative non-linear deconvolution. This approach can estimate both the bias associated with a finite sampling and the effects of irregular sampling or breaks in the data record and remove their signature from the spectral estimate. Another iterative approach to breaks in sampling, the GAPES method of Wang et al. (2005) estimates the spectrum and interpolates gaps in the series using a least squares estimator based on an adaptive-filter bank. These methods are computationally intensive and appear to work best for time series that are well described by periodic components with gaussian noise.

A general method is provided by the Lomb periodogram (Lomb, 1976). This allows for arbitrarily spaced samples, estimating the power spectrum directly from a least-squares fit to sinusoids at specified Fourier frequencies. In its basic form it suffers from similar bias issues to the raw periodogram. A more strategic approach suggested by Bronez (1988) works from first principles to minimise bias by simultaneously solving for the Fourier coefficients and optimal windowing functions (also known as data tapers) for an incompletely sampled time series. For a complete and evenly sampled time series the solution yields Thomson’s (1982) result, the prolate tapers (Slepian, 1978) as the optimal windowing functions. Fodor and Stark (2000) specialised the approach of Bronez (1988) by reframing the optimisation problems used by Thomson (1982) and Riedel and Sidorenko (1995) for multi-taper spectral estimation to explicitly solve for a set of windowing functions which are forced to be zero wherever there is a gap in the time series.

The primary focus of our work is to assess the usefulness of the strategies outlined by Fodor and Stark (2000) for increasing the frequency resolution of spectral estimates of broken time series and providing more insight into the long period behaviour of the physical processes being studied. Along with improved resolution, we require reliable estimates of the uncertainty associated with spectral estimates from intermittently sampled time series. We re-evaluate and build on the techniques of Fodor and Stark (2000) who presented tests on a single helioseismological example. We provide a robust method for adapting a set of tapering functions to the samples available. It
can be run on a personal computer. We develop enhanced tools to assess the effectiveness of this technique: the average transfer function of a set of tapers is examined to evaluate the frequency response and resolution; using synthetic data we compare bootstrap method and jack-knife methods (Thomson and Chave, 1991) for finding confidence limits, which are particularly important for geophysical problems and must be reliable at all frequencies; ultimately we quantify the uncertainty in the spectral estimate with jackknife confidence limits. Using the tools outlined above, we outline a strategy for analysing irregularly sampled time series and test the algorithm on synthetic data.

These techniques for spectral estimation for a time series with inconvenient gaps are applied to three geophysical data sets, chosen because the underlying physical processes (changes in length of day, external magnetic field variations, and very long period geomagnetic paleointensity variations) each have distinct spectral amplitudes, dynamic range, and shapes. Data are removed from the complete time series to simulate gaps in measurements of the change in the length of day from 1962-2009 and from the geomagnetic time series (derived from satellite data) that is related to the external dipolar component of the Earth’s magnetic field (Kuvshinov and Olsen, 2006). The properties of spectral estimates from standard multi-tapers and intermittently sampled tapers are compared for these data sets. We also illustrate this technique using a paleomagnetic record of field strength (Tauxe and Hartl, 1997) recorded in a marine sediment core. The record spans 12 million years and has many gaps, the longest of which is over 600,000 years. This time series was previously subjected to limited spectral analysis by Constable et al. (1998). A more detailed study of these paleointensity data using our new techniques is described elsewhere (Smith-Boughner et al., 2011).

### 2.2 Spectral Estimation

The simplest direct spectral estimate, the periodogram is provided by the appropriately normalised square of the absolute value of the Fourier coefficients of the time series (Schuster, 1898; Press et al., 2007). The periodogram has the effect of tapering an equally spaced time series \(d, \text{ sampled at } t = 0, 1, \ldots, N - 1\) by a box-car function in the time domain which corresponds to a convolution by a \(sinc\) function in the frequency do-
main and introduces substantial bias. Another problem with the periodogram estimate is its high variance which can be reduced by section averaging. The bias can be reduced if a suitable tapering function, \( w_t \), (e.g., Hamming window, cosine taper, or prolate taper) is applied to the data in the time domain. The power spectral density (PSD), \( S_f \), is then defined at frequency \( f \) as

\[
S_f = \frac{1}{N} \left| \sum_{t=0}^{N-1} w_t d_t e^{-i2\pi ft} \right|^2 \quad (2.1)
\]

Section averaging can also be used when there are large breaks in data sampling, and involves splitting a record of length \( N \) at the breakpoints, yielding \( M \) sections with lengths \( L_i, i = 1, \ldots, M \). A spectral estimate is produced from each piece and the results are averaged. The maximum spectral resolution of each piece is \( 1/\Delta t L_i \) where \( \Delta t \) is the uniform time interval between samples. To reduce the bias associated with a periodogram estimate, each section can be windowed to produce a singly-tapered spectral estimate, so that each section has resolution \( W/\Delta t L_i \) where \( W \) is the full bandwidth of the window \((w_t)\) used. Averaging these independent singly-tapered spectral estimates from multiple sections of the data provides a spectral estimate with lower variance and less bias than the periodogram. It is also possible to use section averaging on multi-tapered spectral estimates, with each section chosen to have similar frequency resolution.

In multiple taper spectral estimation, the entire data series is windowed in the time-domain by a series of \( K \) orthogonal tapers, \( v_{t,k} \). Each windowed time series is then Fourier transformed. From this set, a robust spectral estimate is produced by taking a linear combination, \( a_k \) of the \( K \) singly tapered linearly independent estimates:

\[
S_f = \frac{1}{K} \sum_{k=1}^{K} a_k \left| \sum_{t=0}^{N-1} v_{t,k} d_t e^{-i2\pi ft} \right|^2 \quad (2.2)
\]

(The individual estimates are approximately statistically uncorrelated because of the orthogonality of the tapers). Depending upon the application, various functions can be used to taper data, each set the solution (or an approximate solution) to an optimization problem. The resulting spectral estimates can have much lower variance and a higher frequency resolution than section averaging.
2.2.1 Prolate Multi-tapers

Thomson (1982) sought a set of orthogonal windows with their energy concentrated within a desired frequency band \((-\omega, \omega)\) by maximizing the integral:

\[
\int_{-\omega}^{\omega} \frac{\sin N\pi(f - f')}{\sin \pi(f - f')} V_k(f') df' = \lambda_k V_k(f).
\]  

where the eigenfunction solutions of this integral in the frequency domain, \(v_k(f)\), are known as the prolate spheroidal wave functions.

For discrete problems, the bandwidth of interest is controlled by the choice of \(NW\), known as the time-bandwidth product. This parameter is the product of the number of samples \(N\) and the desired frequency resolution \(W = 2\omega\). These orthogonal functions \(v_{t,k}\) are calculated in the time-domain by solving the discretized eigenvalue problem:

\[
\sum_{s=0}^{N-1} \frac{\sin(2\pi\omega(s - t))}{\pi(s - t)} v_s = \lambda v_t
\]  

The eigenvector solutions to the above equation are known as the prolate spheroidal sequences (PR tapers) or the Slepian functions (Slepian, 1978). The corresponding eigenvalue, \(\lambda\) of each eigenfunction indicates the amount of energy that is concentrated within the desired frequency band. At most \(2NW\) tapers have a large amount of energy within the desired frequency band, and typically no more than \(1.6NW\) are used in the final spectral estimate. The high energy concentration creates a filtering function with an exceptionally steep roll-off outside of the bandwidth of interest. This property can be exploited further by using the set of eigenvalues as the weights \(a_k\) in Equation 2.2. Tapering using prolate functions is particularly useful for spectra with a large dynamic range.
2.2.2 Minimum Bias Tapers

Another set of orthogonal tapers exists, which minimise the amount of local bias in the resulting spectral estimate. The minimum bias tapers of Riedel and Sidorenko (1995) minimize the integral:

$$\int_{-1/2}^{1/2} f^2 |U_k(f)|^2 df,$$

with the eigenfunctions $U_k(f)$. Approximating this integral in the time domain, the discrete minimum bias tapers $v_t$ are the solution to the eigenvalue-eigenfunction problem:

$$\sum_{s=0}^{N-1} \frac{(-1)^{t-s+1}}{2\pi^2 (t-s)^2} v_s = \lambda v_t$$

(2.6)

The set of eigenfunction solutions to the minimum bias optimization problem can be further approximated by a set of sinusoids:

$$v_{t,k} = \sqrt{\frac{2}{N+1}} \sin \left( \frac{\pi kt}{N+1} \right)$$

(2.7)

For the minimum bias (MB) and the sinusoidal approximation, each taper $v_{t,k}$, has most of its energy concentrated within the interval $\left[ -\frac{k+1}{2(N+1)}, -\frac{k-1}{2(N-1)} \right] \cup \left[ \frac{k-1}{2(N-1)}, \frac{k+1}{2(N+1)} \right]$. The bandwidth of the spectral estimate in Equation 2.2 is controlled by the number of tapers in the set.

2.2.3 Adapting Tapers for Inconvenient Gaps

For incomplete sampling, both of the multi-taper optimization problems discussed above can be adapted using an indicator function, $I_t$, from Fodor and Stark (2000) given as,

$$I_t = 1, \quad \text{if} \quad d_t \neq 0$$

$$0, \quad \text{if} \quad d_t = 0$$

(2.8)
Table 2.1: Optimization problems which define the tapers.

<table>
<thead>
<tr>
<th></th>
<th>Prolate</th>
<th>Minimum Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>[ \sum_{s=0}^{N-1} \frac{\sin(2\pi \omega(s-t))}{\pi(s-t)} v_s = \lambda v_t ]</td>
<td>[ \sum_{s=0}^{N-1} \frac{(-1)^{t-s+1}}{2\pi^2(t-s)^2} v_s = \lambda v_t ]</td>
</tr>
<tr>
<td>Adapted</td>
<td>[ \sum_{s=0}^{N-1} I_s I_t \frac{\sin(2\pi \omega(s-t))}{\pi(s-t)} \bar{v}_s = I_t \lambda \bar{v}_t ]</td>
<td>[ \sum_{s=0}^{N-1} I_s I_t \frac{(-1)^{t-s+1}}{2\pi^2(t-s)^2} \bar{v}_s = I_t \lambda \bar{v}_t ]</td>
</tr>
</tbody>
</table>

where \( d_t \) is set to zero if a sample is unavailable at time \( t \). The indicator function is incorporated into the optimization problems which define the multi-taper set of interest as outlined in Table 2.1. Using the indicator function to adapt the optimization problem restricts the tapers to be zero when there is a gap in the time series. The solution of these adapted optimization problems to provide minimum bias tapers with gaps (MBG tapers) and prolate tapers with gaps (PRG tapers) that correspond to the available samples of a particular time-series poses some computational difficulties leading us to use suitable approximate solutions for long time series. Details of how the solution to the optimization problems for PRG and MBG tapers can be found in a numerically stable and efficient fashion are provided in the appendix.

### 2.3 Evaluating a Set of Tapers

In Section 2.5 we develop an algorithm for applying PRG and MBG tapers to real data. Before describing the algorithm, we outline some general properties of the adapted tapers in both the time and frequency domains, initially focusing on a single example of an indicator function. The breaks for this particular function vary in length from 230 to 1169 samples, and amount to 14.5% of the total 14500 sample record. The function used is shown in Figure 2.1, along with its corresponding PRG and MBG tapers.

The tapers drop smoothly to zero where there are no samples, as would be expected near the ends of complete time series. For the PRG tapers, as the energy concentration deteriorates (for \( k > 19 \), \( \lambda \) is less than 0.9 for these tapers), the tapers increase...
Figure 2.1: (a) Prolate tapers with gaps for $NW = 12$, computed for the simulated arrangement of gaps. All tapers shown have eigenvalues, $\lambda \approx 1$. (b) Minimum bias tapers with gaps computed for the simulated arrangement of gaps, 10 shown. (c) Indicator function used.
Figure 2.2: The energy concentration of the adapted tapers across 14500 time samples is shown for 15 PRG tapers with \( NW = 12 \) (blue solid line) and 15 MBG tapers (green solid line). For comparison, 15 standard tapers for a complete data set are plotted in broken lines, PR tapers with \( NW = 12 \) (blue broken line) and MB tapers (green broken line)

in amplitude near the breaks in sampling, resulting in increased ringing in the frequency domain. Each of the MBG tapers samples only one segment of the available data and gives zero weight to data elsewhere. From Figure 2.1 it seems that in this case we might expect the gap-adapted minimum bias tapers to produce something rather like a section-averaged spectral estimate, with multiple orthogonal tapers used within most sections, but self-adapting to provide comparable frequency resolution for the multi-taper spectral estimate within each segment.

We can examine the data energy distribution, which reflects data usage in the spectral estimate by plotting \( \sum_{k=1}^{K} |v_{t,k}|^2 \), shown in Figure 2.2 for the PRG and MBG tapers and their standard counterparts. As more tapers are added the energy distribution becomes more uniform across the sampled time interval, but spectral leakage will also become a greater problem. It is notable that the PR tapers appear to underutilize data near each end of the record, while all of the PRG, MB, and MBG seem less susceptible
to this problem. This somewhat counterintuitive result arises because we have limited
the number of PR tapers used to 15, whereas a typical PR spectrum would use at least
1.6NW. In the PRG case using tapers with NW < k < 2NW gave poor control over
spectral leakage, further discussed below.

The frequency domain properties of both the individual tapering functions and
the cumulative impact of the entire set of K tapers provide a useful complement to the
energy concentration. They provide empirical measures of the filtering bandwidth, the
slope of the fall-off in energy and the level of out-of-band signal rejection. We examine
the frequency domain properties of a set of multi-tapers empirically by considering the
average transfer function of a set of tapers as used by Percival and Walden (1998),

\[ H_{f,K} = \frac{1}{K} \sum_{k=1}^{K} \left| \frac{1}{N} \sum_{t=0}^{N-1} v_{t,k} e^{-i2\pi f t} \right|^2 \] (2.9)

The average transfer functions of the standard prolate and minimum bias tapers
are plotted in Figure 2.3(a,b), where we have assumed 1 second sampling. From this
plot of the response for positive frequencies (the negative frequency response is a mir-
ror image), we can estimate the half-bandwidth and filter roll-off of the combined set
of tapers. These plots also allow us to estimate the energy level of the tapers at high
frequencies; this controls the out-of-band signal rejection. For standard tapers, these
quantities can be easily predicted from theory but for the adapted tapers they vary de-
pending on the indicator function of the data. The empirically determined bandwidth
and level of out-of-band rejection provided by a particular set of tapers can be used as a
diagnostic for selecting the appropriate number of tapers. To be conservative in estimat-
ing the properties of these adapted tapers, the bandwidth of a set of tapers is determined
visually from a log-linear plot as the mid-point of the steepest decrease in the energy
of the average transfer function and is shown in Figure 2.3 for a set of 15 tapers as a
horizontal red line in panels (a),(b) and for a set of 17 tapers (explained in Section 6.1),
is a blue line in panel (d).

The average transfer function for the set of PRG tapers with NW = 12 (from
Figure 2.1(a)) is shown in Figure 2.3(a) and compared to the equivalent PR tapers. For
these prolate-based tapers, the bandwidth of the set is a function of the time-bandwidth
Figure 2.3: Average transfer function, $H_{f,K}$ with horizontal red lines denoting the energy level of the set of 15 tapers at half-bandwidth frequency, $W/2$. (a) the prolate tapers with gaps compared to the prolate tapers (both with $NW = 12$). A set of 15 PRG tapers has a half-bandwidth of 0.8 mHz. (b) the minimum bias tapers with gaps (solid lines) compared to the standard minimum bias tapers ( 15 MBG tapers have a half-bandwidth of 0.9mHz). (c) The PRG ($NW = 12$) are compared to the MBG. (d) The prolate tapers for complete data (broken lines) compared to the average transfer function used for section averaging (solid lines). The section average scheme using prolate windows has a half-bandwidth of 1 mHz and a much higher level of energy at frequencies higher than $W/2$. 
parameter, $NW$. With a time-bandwidth product of 12, these tapers should have a half-bandwidth of $12/N_f$, roughly 0.8 mHz for a data series 14500 samples long. If a different bandwidth is required, the tapers must be recomputed for the new time-bandwidth parameter. In our example, the prolate tapers with gaps have a slightly larger half-bandwidth of about 0.9 mHz for a set of 15 tapers. We studied other arrangements of incomplete data and found that the bandwidth of the PRG tapers was typically 20-30% larger than predicted by the time-bandwidth parameter, $NW$. The PRG tapers also have very good energy concentration and fall-off at the same rate as the PR tapers but their out-of-band energy level is many orders of magnitude higher than the prolate tapers. As more tapers are added to the set, this out-of-band energy level increases, greatly reducing the protection against broad-band spectral leakage generally provided by prolate tapers.

For a set of minimum bias tapers, the bandwidth is determined strictly by the number of tapers used, and this is also true for the MBG tapers. Their average transfer function for positive frequencies (adapted for the example indicator function in Figure 2.1(c)) is shown in Figure 2.3(b). For this example, and many other arrangements of gaps we examined, the MBG tapers have a rate of fall-off identical to the standard minimum bias tapers. However, the MBG tapers have the smooth decay of the MB tapers with a slight ripple in the amplitude of the fall-off, this decreases as more tapers are added to the set. In our example, the MBG have a half-bandwidth which is 20-30% larger than the standard minimum bias tapers. The bandwidth of the MBG tapers increases as the time series becomes more fragmented.

To compare the properties of the MBG and PRG tapers, we examine the average transfer function of the PRG and MBG tapers in Figure 2.3(c). We find that the PRG tapers have a much steeper fall-off than the MBG tapers when fewer than $NW$ tapers are used, giving them a smaller averaging bandwidth for small numbers of tapers. When more than $NW$ tapers are used, the fall-off and averaging bandwidth of the MBG tapers are very similar to the PRG tapers but above $3mHz$ (for the tapers in Figure 2.3(c)) the MBG have a lower level of energy. In fact, for our example, at frequencies greater than $f_N^*/4$, the average transfer function of the MBG tapers is four orders of magnitude lower than the PRG tapers. Here $f_N$ is the Nyquist frequency, which lies far outside Figure
2.3. For this example with $NW=12$, there is a steep increase in the out-of-band energy level of the average transfer function shown in Figure 2.3 when more than 12 tapers are in the set of interest. In our tests with several different values of $NW$, we found a sharp decrease in the protection against spectral leakage offered by the PRG tapers when more than $NW$ of the tapers are used in the set.

Figure 2.3(d) also shows a comparison of PR tapers with the section averaging approach demonstrating the greater spectral leakage and loss of frequency resolution. Comparisons between PRG and section averaging are discussed further in the context of the length of day data in Section 2.6.1.

To quantify the differences between energy concentrations of the PRG and MBG tapers, we determine the average amount of the taper’s energy, in the frequency domain, that is outside our desired averaging band and therefore contributes to spectral leakage. Selecting a target averaging bandwidth of 0.0019 Hz, using numerical integration and a slow Fourier transform we compute the energy contained within 0.002-0.50 Hz. Our results, plotted in Figure 2.4 indicate that as more MBG tapers are used to produce a spectral estimate there is a steady decrease in energy concentration, resulting in worse frequency resolution. For the PRG tapers, when more than $NW$ tapers are used there is a steep increase in broadband energy leakage to the 0.002-0.5Hz band although this does not increase the computed averaging bandwidth. These results indicate that, although small numbers of PRG tapers have the best energy concentration, when a large number of tapers is required for variance reduction the MBG tapers will provide better energy concentration and protection against spectral leakage and hence less bias in the resulting spectral estimate.

2.4 Confidence Limits

The confidence limits of a power spectrum estimate a range within which the ‘true value’ of the power spectral density is likely to lie with a given probability. Accurate confidence limits for spectra require knowledge of the probability distribution of the spectral estimates. For multiple taper spectral estimates, the spectrum is computed as a linear combination of a set of linearly-independent estimates, each produced from
a single taper. The singly tapered spectral estimates each follow a scaled chi-squared ($\chi^2$) distribution (Percival and Walden, 1998). If prolate tapers or minimum bias tapers are used, then the mean of these spectral estimates roughly follow a $\chi^2$ distribution with $K - 1$ degrees of freedom, where $K$ is the number of tapers used. For a sufficiently large number of tapers, this can be further approximated as a Gaussian random variable. 

Under these assumptions, we can estimate with 95% confidence that for white spectra, the value of the spectral estimate is within $2\sigma$ of the mean of the singly-tapered estimates, where $\sigma^2 = 2S_f^2/\nu$, and $\nu$ is the equivalent number of degrees of freedom (usually $K - 1$). When the spectral estimate has complicated structure (e.g. narrow spectral peaks) or is significantly coloured, the assumption that the singly-tapered spectral estimates follow a $\chi^2$ distribution fails. More accurate methods for calculating confidence intervals are needed.

Parametric methods are typically unreliable when used to determine uncertainties for spectral estimates. The manifestation of this is usually that a $1 - 2\alpha$ level of confidence is $1 - 2\beta$ with $\beta$ significantly larger than the stated value of $\alpha$. Various strategies can be used to address this issue.

We determined empirically (from bootstrap resampling) that for each time-series with samples missing, the statistical distributions of the single PRG or MBG tapered spectral estimates have a different shape from the distributions from PR and MB tapers. As the arrangement of gaps and continuous data changes, the shapes of these distributions also change. For any given time series with gaps, the distribution of the mean of a set of spectral estimates produced from PRG or MBG tapers is strongly influenced by the number of tapers in the set.

Because of the above properties, we seek an empirical method for computing confidence intervals. Thomson and Chave (1991) suggested using empirical jackknife resampling and a logarithmic transformation of the sampled variable to estimate confidence intervals for coloured spectra. This jackknife method was compared to a suite of bootstrap-based methods by Fodor and Stark (2000), who examined the statistical coverage at one frequency of interest and found that a computationally intensive bootstrap method gave the best coverage. However, we are interested in the accuracy of the confidence intervals across all frequencies and investigate which method provides the
most accurate confidence intervals for a wide variety of PRG and MBG tapers. Two promising methods are the bias-accelerated and corrected bootstrap percentile method and the jackknife technique of Thomson and Chave (1991).

2.4.1 Confidence limits using empirical resampling

A jackknife uses an empirical resampling of the statistic of interest, \( \theta \), created by removing one (or more) samples at a time. The jackknife estimate of \( \theta \) is the mean of the set of estimates of \( \theta_{(i)} \), where the \( i^{th} \) data point has been removed from the set. Thomson and Chave (1991) used a logarithmic transformation of the singly tapered spectral estimates \( S_k \) and approximated \( \left( \ln \hat{S}_{(i)} - \ln \hat{S}_{(\cdot)} \right) / \tilde{\sigma} \) as a \( t_{N-1} \) distribution. To use this approximation, the variance, \( \tilde{\sigma}_f^2 \) of the set of singly tapered estimates at \( f \) is estimated using a delete-one jackknife:

\[
\tilde{\sigma}^2_f = \text{var}\{\ln \hat{S}_f\} = \frac{N-1}{N} \sum_{i=1}^{N} (\ln \hat{S}_{f,(i)} - \ln \hat{S}_{f,(\cdot)})^2
\]  

(2.10)

This yields confidence intervals:

\[
\hat{S}_f e^{-t_{N-1}(1-\alpha/2)\tilde{\sigma}_f} < S_f \leq \hat{S}_f e^{t_{N-1}(1-\alpha/2)\tilde{\sigma}_f}
\]

(2.11)

The width of these confidence intervals depends on the spectral estimate and is exponentially related to \( \tilde{\sigma} \), the variance of \( \ln S_f \).

A bootstrap is a random resampling of the \( K \) spectra by drawing each datum \( x_i \) from the set at random, with replacement to generate a new sampling of our data set. This resampling operation is performed \( B \) times to form \( B \) sets of pseudo-data. From this, a more robust value of the statistic of interest, \( \theta \), can be estimated.

Confidence intervals of our statistic \( \theta \) are determined from the empirical \( \alpha \) and \( 1 - \alpha \) percentiles of the mean of the spectral estimates. A very large number of bootstrap resamplings, typically 1000, are usually required to accurately estimate the \( \alpha \) and \( 1 - \alpha \) empirical percentiles of \( \theta \). This is known as the percentile method; it is a first-order accurate method, that is the deviation of the empirical \( \alpha \) and \( 1 - \alpha \) percentiles from their true value heads to zero at a rate of \( 1/K \) as \( K \), the number of sample points, increases.
For our applications $K$ is the number of singly-tapered spectral estimates. Several more accurate methods exist (see Davison and Hinkley (1997) for further details).

For our application, we consider the "bias-corrected and accelerated" bootstrap (BCA) percentile method outlined by Davison and Hinkley (1997); it is one of the faster bootstrap methods used in Fodor and Stark (2000). Through the use of a bias-correction parameter $\hat{z}_0$ and acceleration parameter $\hat{a}$ to adjust the empirical percentiles, the BCA method converges at a rate of $1/\sqrt{K}$. The bias-correction parameter is calculated from the set of $B$ estimates of $\theta$ as:

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b) < \hat{\theta}\}}{B}\right)$$

(2.12)

where $\#\{\hat{\theta}^*(b) < \hat{\theta}\}$ is the number of estimates of $\theta$ which are less than the mean and $\Phi^{-1}$ is the inverse of the standard normal distribution. The acceleration parameter, which is similar to a measure of skewness, is estimated from:

$$\hat{a} = \frac{\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}_i)^3}{6(\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta}_i)^2)^{3/2}}$$

(2.13)

where $\hat{\theta}_i$ is the delete-one jackknife estimate of $\theta$ with the $i^{th}$ datum removed and $\hat{\theta}_{(i)}$ is the mean of these jackknifed estimates $\hat{\theta}_i$.

The accuracy of the bootstrapped BCA and jackknife confidence intervals for the PRG and MBG tapers are explored using many replications of a process from a manufactured spectrum. The spectrum is produced by combining a set of smooth, continuous functions in the time domain to form a time series $x_t$ whose power spectral density has a large dynamic range, broad peak, and steep fall-off in energy at low frequencies. The spectrum of this series is plotted in red in Figure 2.6. From this synthetic power spectral density $S_f$, a set $\{y_{t,j}\}_{j=1}^J$ of different replications of this process is generated using phase noise:

$$y_{t,j} = \mathcal{F}^{-1}[\sqrt{S_f}e^{-i2\pi p_{f,j}}]$$

where $p_{f,j}$ is a uniformly distributed random variable and $\mathcal{F}$, $\mathcal{F}^{-1}$ denote the Fourier and inverse Fourier transforms respectively. A sample time series is given in Figure 2.5.

Using 145 different realizations of our synthetic spectrum $S_f$, sets of singly-
**Figure 2.4:** Comparison of the average energy concentration of the MB, MBG, PRG and PR tapers outside of the frequency band of $[0, \frac{2NW(1.2)}{N_f}]$ as a function of the number of tapers used to produce the spectral estimate. This frequency band is 20% larger than the averaging bandwidth for standard prolate tapers.

**Figure 2.5:** One of 145 replications of the synthetic process, whose spectrum is shown as the red line in Figure 2.6
Figure 2.6: The spectra of the synthetic data produced from 18 minimum bias tapers on the complete data set (bright green line) and 18 tapers adapted for gaps (yellow line), corresponding to the indicator function in the bottom of Figure 2.1. The data are synthesized to correspond to the spectra plotted in red. Nominal 95% confidence intervals are shown for both the jackknife (black filled region) and BCA methods (gray region). The jackknife confidence intervals shown obtain 95% coverage of the synthetic spectra while the BCA only achieves 91%.
Table 2.2: Results of synthetic tests of 95% confidence intervals from prolate and minimum bias tapers with and without large breaks in sampling.

<table>
<thead>
<tr>
<th># of Tapers</th>
<th>Prolate</th>
<th>Minimum Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PR</td>
<td>PRG</td>
</tr>
<tr>
<td></td>
<td>BCA</td>
<td>JK</td>
</tr>
<tr>
<td>Coverage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=6</td>
<td>82.8</td>
<td>95.5</td>
</tr>
<tr>
<td>K=10</td>
<td>88.7</td>
<td>95.2</td>
</tr>
<tr>
<td>K=14</td>
<td>91.3</td>
<td>95.5</td>
</tr>
<tr>
<td>K=18</td>
<td>92.4</td>
<td>95.5</td>
</tr>
<tr>
<td>Width</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=6</td>
<td>1.71</td>
<td>6.20</td>
</tr>
<tr>
<td>K=10</td>
<td>1.48</td>
<td>2.38</td>
</tr>
<tr>
<td>K=14</td>
<td>1.28</td>
<td>1.69</td>
</tr>
<tr>
<td>K=18</td>
<td>1.13</td>
<td>1.36</td>
</tr>
</tbody>
</table>

tapered spectral estimates of $y_t$ were created to test the performance of the BCA and jackknife methods for computing confidence intervals for the standard and gap-adapted prolate and minimum bias tapers. From each replication, prolate and minimum bias tapers were used to create a set of singly-tapered spectral estimates. Then, using 6, 10, 14 or 18 tapers, approximate 95% confidence intervals were calculated from an unweighted mean of the single taper estimates using the BCA and jackknife methods described above. Although Equation 2.2 would normally imply the use of a weighted mean here, in this context we found it produced confidence intervals with inadequate coverage.

To test the PRG and MBG tapers, data were removed from the time-series in the arrangement shown in Figure 2.1c and singly-tapered spectral estimates produced. Again, confidence intervals were computed using 6, 10, 14 or 18 tapers. Typical BCA and jackknife confidence intervals for one realization are shown in Figure 2.6 for minimum bias tapers with gaps and would be nearly identical for prolate tapers with gaps.

The results for 145 replications are compared using two measures, the average normalized width of the confidence interval and the coverage of the 95% confidence interval. The normalized width for a spectral estimate from K tapers is computed at each frequency by subtracting the upper and lower confidence intervals and dividing by the spectral estimate computed from the mean of the K singly-tapered spectral estimates (for PR and PRG tapers the mean is weighted by the eigenvalues as specified in Equa-
tion 2.2). The coverage is given as the percentage of frequencies where the confidence intervals contain the ‘true’ spectrum of the process, plotted as the red line in Figure 2.6. These results are documented for BCA and jackknife confidence intervals in Table 2.2.

From these synthetic tests we see that the BCA confidence intervals are approximate, the actual coverage is always less than our desired 95%, increasing from 83-85% to 91-93% as more tapers are added. For the jackknife method, the coverage is between 95-96% regardless of the number or type of tapers used to compute the spectra. Further experiments revealed that when 90% coverage was requested the actual coverage of the jackknifed confidence intervals was roughly 92-93% and therefore broader than necessary while the BCA confidence intervals were slightly too narrow to obtain the desired coverage, achieving 90% coverage only when 18 PR, PRG, or MB tapers were used; the coverage was at most 87% for the MBG tapers.

As more tapers are added to the set, the width of the jackknife confidence intervals shrink dramatically while the BCA confidence intervals reduce in width only slightly as more tapers are used. For both the BCA and jackknifed confidence intervals, the normalized width of the interval for adapted tapers (PRG, MBG) are smaller than those corresponding to the standard tapers. This difference gets smaller as more tapers are used.

Jackknife confidence intervals provide more accurate and consistent coverage. Additionally, jackknife resampling requires far less computation than the bootstrap. These synthetic examples with 2049 frequencies were computed in 6 seconds, while the BCA computation required 220 seconds on a 2.66 GHz Intel Core Duo 2 processor with 4 GB of RAM, 30 times longer to compute. However, jackknife estimates can be problematic when the estimator is not smooth like $\tilde{\sigma}$ (Davison and Hinkley, 1997). Because variance in the log of the spectra is estimated for the jackknife confidence intervals, they can be excessively large when the distribution of Equation 4.26 is highly skewed giving more coverage than requested. These problems were most likely to occur when the spectra had a large dynamic range and were therefore prone to spectra leakage. An example of this can be seen with the length of day data discussed in section 2.6. However, this issue can be overcome with the use of prewhitening or by using more tapers to compute the spectral estimate.
Table 2.3: Algorithm for applying multi-tapers adapted for breaks in sampling.

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1. | Examine Data  
   (a) Find location of missing samples.  
   (b) Generate histogram of the length of breaks in sampling. |
| 2. | Prepare Data  
   If $\Delta t$ is non-uniform  
   i. Use Akima spline interpolation to create regular sampling rate $\Delta \tau$.  
   If $\text{Range}(\Delta t)$ is large  
   i. Select $T_i$  
   ii. For all breaks in sampling less than $T_i$ in length, use $AR(p)$ filter to interpolate data. |
| 3. | Compute Tapers  
   (a) Decide on MBG and/or PRG tapers.  
   i. For MBG tapers: select $K_{set}$, the number of tapers to compute.  
   ii. For PRG tapers: select NW, the time-bandwidth product of the desired tapers. |
| 4. | Evaluate Tapers  
   (a) Plot $H_{f,k}$ for several values of $K$. Evaluate:  
   i. Bandwidth  
   ii. Fall-off of filter just outside of pass-band frequencies  
   iii. Energy level at high frequencies  
   (b) Plot $\Sigma_{k=1}^{K} |v_{t,k}|^2$ for several values of $K$. Evaluate:  
   i. Weighting of tapers across all available data segments. |
| 5. | Choose final value of $K$, create spectral estimate and compute confidence intervals. Consider:  
   Dynamic range of spectral estimate  
   Width of confidence intervals  
   Indications of spectral leakage |

Now that we have provided a framework for evaluating and understanding the properties of these adapted tapers, we outline an algorithm for applying them to a time series with long interruptions in sampling.

### 2.5 Algorithm

Our strategy is to compute tapers which compensate for the longest gaps within a record to improve the recovery of long period characteristics of the power spectrum. Real data often have many irregular breaks in sampling. The duration and location of
these breaks within the time series are unique, therefore a specific set of adapted tapers must be computed for each time series. However, as we saw with the PRG tapers in Figure 2.1, there is a trade-off between the tapering properties (minimizing local or broadband bias) sought in the optimisation problem and the length of the continuous segments of data. Many short data segments detract from the desired low-pass filtering properties and this can result in uneven weighting of the data segments in the time domain. Very uneven weighting of the time series will result in a biased characterisation of the average spectral properties of the time series. Although this problem can be alleviated by using interpolation to eliminate small breaks in sampling, this approach will introduce bias at frequencies inversely proportional to the length of the break being filled.

To balance these considerations, we develop an algorithm to apply adapted multi-tapers and interpolation to a data set with interrupted sampling. This algorithm is outlined in Table 2.3 and we discuss certain aspects of its application below.

Based on the histogram of the length of the breaks in sampling and the indicator function of the time series, the user selects an interpolation threshold, \( T_i \), in terms of the number of samples missing, where \( i \) is the iteration number. All breaks in sampling which are shorter than \( T_i \) are filled by applying the same auto-regressive (AR) filter to the segment of data immediately preceding and following a break in sampling to create forward and backward predictions. The break in sampling is filled by averaging these two predictions. All gaps shorter than \( T_i \) are filled using the same AR filter and the indicator function is updated to include the interpolated segments. For the time series we studied, the AR filter coefficients are computed from the whole time series using a modified-covariance technique (Marple, 1991) and the order of the filter is chosen through trial and error. After this interpolation, all breaks remaining in the sampling are considered as “gaps” when the adapted tapers are calculated.
Figure 2.7: (a) Original indicator function of synthetic data used to demonstrate algorithm. (b) Histogram of the length of breaks from regular sampling of synthetic data. (c) Indicator function of data after using AR(3) filter to interpolate gaps less than 7 sample points in duration ($T_i = 7$). (d) Sum of the absolute value squared for 8 MBG tapering functions in the time-domain for 2 different levels of interpolation (corresponding to plots c and e). (e) Indicator function of time series after interpolating gaps less than 10 samples long ($T_i = 10$). (f) As plot d) for 8 PRG tapers with $NW = 4$. (g) Average transfer function of PRG with $NW = 4$ tapers for 2 different levels of interpolation. (h) As plot g) for MBG tapers.
Once the adapted tapers are calculated for the current indicator function, plots of $\sum_{k=1}^{K} |v_{i,k}|^2$ and $H_{f,K}$ are examined for several values of $K$, the number of tapers used for the spectral estimate. If the set of tapers does not have the desired properties in the time or frequency domains (e.g. frequency resolution, filter fall-off, even weighting of data segments in the time domain), parameters are adjusted and new tapers computed. Table 2.4 outlines how and when to change the parameters of the algorithm to improve the desired property. For most data sets, if the tapers are not satisfactory, the threshold for interpolation must be increased and tapers recalculated using the new indicator function.

To demonstrate how the algorithm works, we apply it to a time series of 580 samples (with 1 second between samples) from the synthetic data set used in Section 2.4 and remove data to synthesize gaps in sampling. The indicator function describing the availability of data is shown in Figure 2.7(a) and a histogram of the time between samples is shown in Figure 2.7(b). To apply the algorithm, we first choose $T_1 = 7$ as the threshold value and use an AR filter to interpolate any gaps shorter than $T_1$. After the interpolation, the new indicator function has 6 gaps in sampling (shown in Figure 2.7(c)). Both prolate and minimum bias tapers are calculated for the current indicator function.

The time-domain weighting and average transfer functions for the prolate tapers with $NW = 4$ and the minimum bias tapers are shown as solid lines in Figure 2.7(d) and (h) (for MBG) and in (f) and (g) (for PRG tapers). For this amount of interpolation, the MBG tapers give zero weight to two data segments, between samples 216-248 and 501-514 (Figure 2.7(d)). From the average transfer function in Figure 2.7(h), we see that the half-bandwidth of these MBG tapers ranges from 0.012-0.02 Hz for sets of 2-8 tapers. The prolate tapers adapted for these 6 gaps have a half-bandwidth of 0.08 Hz (Figure 2.7(g)). These tapering functions are adequate but, looking at the histogram of the sampling breaks in Figure 2.7(b), we see that increasing the threshold for interpolation slightly will eliminate several gaps and increase the lengths of the continuous segments of data. The small increase in bias due to interpolation should be outweighed by the improvement in the tapers.

To determine the amount of improvement in the filtering properties of the tapers, we set $T_2=10$ and interpolate the time series again, resulting in 3 breaks in sampling.
This indicator function is shown in Figure 2.7(e) and is similar to that shown in Figure 2.1. New adapted tapers are computed; the dashed-lines in Figure 2.7 (d) and (h) (for MBG) and (f) and (g) (for PRG tapers) of Figure 2.7 show the time-domain weighting and average transfer function of these new tapers. The MBG tapers from $T_2$ have a much better distribution of energy across the time domain than the MBG tapers from $T_1$; for $T_1$ the very short segments of continuous data are essentially ignored. In the frequency domain, the new MBG tapers have a smaller averaging bandwidth but a similar fall-off in energy. At frequencies above the pass-band, the energy level of the tapers is lower than the previous iteration for the MBG tapers. The decrease in energy at high frequency (and in susceptibility to spectral leakage) is much more pronounced for the PRG tapers. The new PRG tapers also have a steeper fall-off in energy near 8 mHz; this should increase the ability of the PRG tapers to resolve closely spaced spectral lines, a particular strength of the prolate tapers. The slight increase in bias from interpolation has greatly improved the tapering functions.

Once a satisfactory set of tapers is calculated, it can be applied to the data to estimate the power spectral density and confidence intervals. This could reveal an inappropriate choice of $K$ (the number of tapers used), the interpolation scheme or type of multi-taper used, in which case another iteration of the algorithm could be required. To determine how robust the various features of the spectrum are, estimates can be made using a different number of tapering functions to assess how different frequency resolutions alter the spectrum. As another check, the tapers from the previous iteration can be used to verify that the increased interpolation does not unduly alter the spectral properties at the frequencies of interest.

### 2.6 Applications to Geophysical Data

The PRG and MBG tapers are tested on three interesting geophysical data sets. For the length of day and external magnetic field data sets, samples are removed, then a spectrum is produced and compared to a spectral estimate from the complete time series. The techniques are also tested on a record of paleointensity measurements from a sediment core spanning 12 Myr, but with many breaks in sampling, the longest of
Table 2.4: Parameters to alter to improve time and frequency-domain properties of adapted tapers.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Time domain weighting is erratic.</td>
<td>Increase $T_i$ to raise level of interpolation.</td>
</tr>
<tr>
<td>2. Frequency resolution is too low.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>If using PRG tapers, decrease NW</td>
</tr>
<tr>
<td></td>
<td>Increase $T_i$ to raise level of interpolation to create longer segments of continuous data</td>
</tr>
<tr>
<td>3. Filter fall-off in frequency domain is too shallow.</td>
<td>Increase $T_i$ to raise level of interpolation.</td>
</tr>
<tr>
<td>4. Large amount of energy at high-frequencies</td>
<td>Increase $T_i$ to raise level of interpolation.</td>
</tr>
</tbody>
</table>

which is over 600 kyr. These techniques have been used elsewhere (Smith-Boughner et al., 2011) to improve the analyses of the long period characteristics in this record.

2.6.1 Length of Day

A time series of 14500 daily samples (fs=1 cycle/day) of the change in the length of day from the International Earth Orientation Reference System (data from http://hpiers.obspm.fr/, June, 2010), from 1962 to 2009 is used to test the capabilities of tapers adapted for gaps. A full review of the broad range of contributions to changes in length of day is given by Gross (2007), who notes the largest change as due to the fortnightly tide. Many of these variations occur at low frequencies as high amplitude, narrow band signals seen in pilot spectral estimates. These signals contribute to the large dynamic range seen in pilot spectral estimates and (although in practice these tidal signals would be removed prior to most spectral analyses) they provide a good test of the frequency resolution, local bias properties, and protection against spectral leakage offered by the PRG and MBG tapers.

To compare the spectral properties of the various multi-tapers, we conducted 3 tests with various gap arrangements: the arrangement of Figure 1, equal length data segments, and a more variable gap structure.

Initially we discuss removing 14.5% of the available data record (times series is shown in Figure 2.8) again using the arrangement shown in Figure 2.1(c). Power spectral estimates for this intermittently sampled record are generated from 15 PRG
Figure 2.8: The change in the LOD from 1962 to 2009 with simulated gaps, indicated by zero values.

tapers with $NW = 12$ (from Figure 2.1(a), and shown as the blue spectrum in Figure 2.9) and 15 MBG tapers (green line). These estimates are compared to an estimate from 15 PR tapers used on the complete record (red line). All three spectra are plotted in Figure 2.9 along with 95% jackknife confidence limits produced from the 15 PRG tapers (gray shaded region.) Inset in Figure 2.9 is a close up of the low frequency response. The resolution of the PRG tapers is $1.7 \times 10^{-3}$ cycles per day (cpd), $1.6 \times 10^{-3}$ cpd for the PR tapers and $1.8 \times 10^{-3}$ cpd for the MBG tapers (the average of these is shown on the inset Figure near $f=0.04$ cpd).

Comparing the PR and PRG spectra reveals that the amplitudes are very similar at low frequencies with both spectra resolving nearly the same peak widths and shape for nearly all of the narrow band signals. At intermediate frequencies, the PRG spectrum has less energy than the PR spectrum while at higher frequencies, the PRG spectrum has more energy as a result of spectral leakage. As we see in the average transfer functions shown in Figure 2.3(a), we expect the PRG tapers to have similar resolution to the PR but less protection against spectral leakage. When more tapers are used to estimate the spectrum, the PRG estimate becomes increasingly prone to spectral leakage at high
Figure 2.9: Power spectral density from change in the LOD, computed using 15 PRG tapers with data missing (blue line) and using 15 PR tapers (red line) on the complete record. This is compared to a spectrum created using 15 MBG tapers (green). Jackknife 95% confidence limits for the spectrum created with 15 PRG tapers are shown (gray shaded region). The inset shows a close up of the low frequency region. All spectra shown have a frequency resolution of approximately 0.002 cpd (shown as the small blue bar in the inset plot near 0.04 cpd.)

frequencies; with 18 PRG tapers, the spectrum flattens out at frequencies above 0.28 cpd to a power level of $10^{-10}$.

The spectrum estimated from MBG tapers (green line in Figure 2.9) also closely matches the estimate from the complete data set at low frequencies but does not resolve the base of many narrowband signals as well as the PRG tapers. At intermediate frequencies, the MBG spectra matches the PRG spectra, reflecting the similarities in their average transfer function at low frequencies (red lines in panel (c) of Figure 2.3.) At high frequencies, the MBG tapers have less energy, because as we noted in Figure 2.4, the estimate from MBG tapers has less spectral leakage than the PRG estimate. Additionally, the MBG confidence intervals have a normalized width of 3.07, while the PRG confidence intervals (gray region in Figure 2.9) have a width of 2.14. The narrower intervals indicate that there is less variance in the spectral estimate and therefore it is more
Figure 2.10: Power spectral density from change in the LOD plotted on a log-log scale. The spectral estimate from the complete record using prolate tapers (red line) is compared to a section average estimate (green line) computed from 4 sections each windowed using several prolate multi-tapers (method described in text). The spectral estimate from the incomplete data using the PRG tapers is plotted in blue.
accurate.

To further test the benefits of adapted tapers, we compare spectral estimates from PRG tapers to one where breaks in sampling are addressed using section averaging. Section averaging of the raw periodogram reduces variance in a spectral estimate, and to address the issue of bias in the raw periodogram, we use prolate tapers on each of the four data segments shown in Figure 2.1(c) to make a fair comparison to our method. We choose prolate tapers because the frequency resolution is controlled by the time-bandwidth product, rather than the number of tapers (the case for minimum bias tapers.) Each of the 4 segments is zero padded to be 8192 samples long. The time-bandwidth product (NW) of the tapers for each segment is chosen to produce the same frequency resolution in each resulting spectral estimate. Each segment is tapered using 80% of the available tapers (the segments are windowed using 6,4,4 and 3 tapers with NW=4,2.5,2.5 and 2, respectively) and combined. The average of these 17 singly tapered spectral estimates is plotted in Figure 2.10 (green line) and their average transfer function is shown in as a light blue line in Figure 2.3(d) and has a resolution of roughly $2.4 \times 10^{-3}$ cpd. The peak at the lowest frequencies has a width of roughly 1.6 mHz, this corresponds to $NW_i/\Delta t L_i$ where $NW_i=4$ and $L_i=4791$, the length of longest data section. This peak broadens to a width of 2.4 mHz which is roughly to be expected from applying prolate tapers with $NW = 2$ to the shortest segment, 1473 samples long (Figure 2.3(d) shows only positive frequencies). The broader shape of this transfer function increases the amount of spectral leakage, particularly at high frequencies. Narrow peaks in the spectral estimate, particularly when they are closely spaced are not well resolved because of this shape. Jackknifed confidence intervals for this spectrum have a normalized width of 3.09 and are wider than confidence intervals of the PRG spectra (shaded gray region in Figure 2.9).

For this data set with 14.5% of the record missing, the narrower confidence intervals and finer frequency resolution of the PRG tapers provide the best spectral estimate, the increased spectral leakage from the PRG tapers could be reduced through the use of prewhitening. However, for this availability of data, if we ignore the shortest segment of data when using section averaging, the frequency resolution would be match the PRG tapers but the estimate would have larger confidence intervals.
To compare section averaging and adapted tapers under conditions that might be considered ideal for section averaging, other arrangements of missing data were tested. With 10% of the record missing in 5 equal, evenly distributed pieces, section averaging using prolate tapers (or minimum bias) provided better frequency resolution than PRG tapers. MBG tapers gave comparable frequency resolution to section averaging and were less prone to spectral leakage but had wider confidence intervals.

We examine the performance of the adapted tapers on a more variable gap structure with 10% of the record missing, leaving 5 segments of data each of different lengths. We compute PRG and MBG tapers, the first ten of each are shown in Figure 2.11(a) along with the indicator function. Here again, each MBG taper only windows one data segment and the tapers are extremely similar to the tapers used for section averaging with minimum bias tapers. The time-domain energy concentration of these tapers (Figure 2.11(b)) reveals that while the PRG tapers are more complex in the time-domain they provide more even weighting that the MBG tapers and the amplitude of the ripple is similar to that from standard prolate tapers. When only 15 tapers are used, both families of multi-tapers ignore the shortest data segment. The average transfer functions of the PRG and MBG tapers are very similar to those from our earlier example with 14.5% of the record missing, shown in the two leftmost plots of Figure 2.11(c). We compare these to the average transfer function from using the section averaging scheme described previously (two rightmost plots in Figure 2.11(c)). Using prolate tapers for section averaging can provide a frequency resolution similar to PRG tapers when the shortest segment of data is ignored (solid blue line in plot) but much less protection against spectral leakage and wider confidence intervals. For minimum bias tapers with gaps, the frequency response is extremely similar to what would be provided from section averaging, particularly if the last data segment is ignored (solid blue line in plot).
Figure 2.11: (a) The first 10 PRG (NW=12) and MBG tapers are shown for an indicator function with 5 data segments and 11% of the record missing (see Figure 1 (a,b) for legend). (b) The time domain energy concentration of 15 of the tapers show in panel (a). (c) A comparison of the average transfer functions for the indicator function shown in from panel (a) to the standard tapers (left two plots) and to section averaging (right two plots). (d) Spectral estimate of the change in the LOD from 15 PR tapers for NW=12 applied to the complete record (red line), the 15 PRG tapers from panel (a) applied to the LOD data with data missing. The spectral estimate from a section average estimate using PR tapers is plotted in green.
Figure 2.12: Time series of the external dipolar fluctuations in the magnetic field of the Earth from Kuvshinov and Olsen (2006) with data removed to simulate gaps.

In Figure 2.11(d), we compare the spectral estimates using various prolate-based methods. The spectrum estimated using section averaging (green line) has more spectral leakage and distorts the shape of the narrow band peaks, particularly at low frequencies. In contrast, PRG tapers (blue line) have slightly less spectral leakage and resolve nearly the same narrow band features that the prolate tapers do when the complete record is available (red line). Further experiments were conducted using a slightly different indicator function, one where the break in sampling was shifted further down the record, leaving one very long data segment, three segments of similar length and one short segment. In this case, the MBG tapers had more spectral leakage than section averaging.

### 2.6.2 External dipole magnetic field contribution, from 2001-2005

A model of 14500 estimates of the external dipolar fluctuations in the magnetic field of the Earth, sampled once every three hours from Kuvshinov and Olsen (2006) is used to test our techniques (time series shown in Figure 2.12). These data are of some
Figure 2.13: Power spectral density from the model of external dipolar fluctuations in the magnetic field from Kuvshinov and Olsen (2006) computed using both 20 MBG tapers (green line) and 20 PRG (blue line) tapers with NW=12. Missing data was simulated in the arrangement shown at the bottom of Figure 2.1. These estimates are compared to a spectrum from 20 MB tapers (red line) applied to the complete record. The MBG (PRG) tapers provide a frequency resolution of $8.3 \times 10^{-3}$ c/day, $(7.8 \times 10^{-3}$ c/day) and the MB tapers have a frequency resolution of $6.7 \times 10^{-3}$ c/day. Jackknife 95% Confidence intervals are shown in grey.

interest because the technique for adapting tapers with gaps could be helpful for solving the problem of low frequency transfer function estimation for global geomagnetic depth sounding from satellite measurements. Although not a problem for this model time series, the CHAOS and Ørsted satellites from which it is derived have breaks in the sampling of the magnetic field of several months in length. To simulate such gaps, data are removed from the model in the arrangement indicated in Figure 2.1(c). The model time series is down-sampled to once every three hours to match the number of data available for the LOD record, allowing us to use the same tapers adapted for gaps.

The spectral estimate of the external dipolar contributions to the Earth’s magnetic field does not have a large dynamic range or fine spectral lines. These properties
suggest that minimum bias tapers would provide the most accurate spectral estimate because they are designed to minimize the local bias in a spectral estimate, rather than broadband bias. In Figure 2.13, we compare the spectral estimates produced from the complete record using MB tapers (red line) to the estimate produced using the MBG tapers on incomplete data (green line). Approximate 95% confidence intervals (gray shaded region) are computed using the jackknife method. The MBG tapers provide a frequency resolution of $8.3 \times 10^{-3}$ c/day while the MB tapers have a frequency resolution of $6.7 \times 10^{-3}$ c/day. This slight difference in frequency resolution leads to some minor differences between the spectral estimates from the MB and MBG tapers (shown in inset of Figure 2.13). The spectrum from the MBG tapers has a slightly lower amplitude and broader peak than the complete MB spectrum. The confidence limits on the spectral estimate from the MBG tapers do not always contain the MB spectrum (shown in inset of Figure 2.13). Note, however, that we expect 96% coverage for the MB spectrum and about 1% less for the MBG, so this does not seem problematic (see Table 2.4).

For comparison we also include an estimate from 20 PRG tapers with $NW = 12$, (blue line in Figure 2.13) with a frequency resolution of $8.8 \times 10^{-3}$ c/day. Because the frequency resolutions, shown in Figure 2.3(c), of a set of 20 PRG (solid light blue line) and MBG (broken light blue line) tapers are similar and the dynamic range in the spectrum is relatively low, the two spectral estimates are very similar for this data set. However, a large number of MBG tapers has better energy concentration than a comparable set of PRG tapers, so the PRG spectrum is more susceptible to spectral leakage than the MBG. Near the Nyquist frequency spectral leakage causes the PRG spectrum to deviate from the MB and MBG. This also yields narrower confidence intervals for the MBG tapers, a normalized width of 1.44, compared to 1.22 from the MBG tapers.

### 2.6.3 A 12 My Record of Paleointensity

For our final example we examine a relative paleointensity record (Tauxe and Hartl, 1997) of 2238 samples from a sediment core taken from Site 522, Leg 73 of the Deep Sea Drilling Program. This core spans 12 million years from 22.7-34.7 Ma, during the Oligocene period and remains one of the longest relative paleointensity records.
available making it a good candidate for recovering the spectrum of long-period variations in geomagnetic field strength. However, there are numerous interruptions in the temporal sampling. The longest gap occurs at 29.4 Ma and lasts nearly 600 000 years, while between 22.7 and 28.8 Ma there are several gaps ranging in length from 40 000 to 120 000 years. The temporal sampling rate is also somewhat variable, because of uneven depth sampling and variations in the rate of sediment accumulation. Previous work by Constable et al. (1998) used Akima spline interpolation and multi-taper spectral estimation to evaluate the power spectrum of various subsets of the record (thereby avoiding excessive interpolation), but the resulting spectra are far from optimal for the fragmented later part of the record. When the algorithm is applied to this data, with one sample every 4000 years, spanning 3000 sample intervals, we use an AR(3) filter and Akima spline interpolation to fill all gaps shorter than 100 000 years (most of these gaps are 10 000 - 65 000 years in length).

Even with interpolation, the numerous interruptions in sampling create many short data segments in the younger portion of the record (see the indicator function (shown in Figure 2.14a). These short data segments require the flexibility of the prolate tapers with gaps to adequately weight each segment. The minimum bias tapers with gaps computed for this record do not window the shortest segments of data and have broader frequency resolution than a comparable number of prolate tapers with gaps. Previous spectral estimates by Constable et al. (1998) showed a modest dynamic range, so spectral leakage is not a pressing concern.

We compare the spectral estimate of the entire record computed using PRG tapers to the spectral estimate from PRG tapers of the incompletely sampled section from 22.7- 28.8 Ma and PR tapers for the complete portion of the record from 29.5-34.6 Ma (Figure 2.15.) The PRG and PR tapers are used in this analysis to produce spectral estimates of the three data sections of roughly the same frequency resolution by selecting different values of NW for each section. As there is low dynamic range in this spectrum, there should be little problem with spectral leakage. The confidence limits of each spectrum were computed using the jackknife technique.

The plot of all three spectra (Figure 2.15) and their confidence intervals reveals more energy in 29.5-34.6 Ma portion of the record than in the time span of 22.7-28.8
Figure 2.14: (a) Time series of paleointensity measurements of the DSDP Leg 73, Core 522 after interpolation using AR(3) Filter on gaps less than 100 000 years. (b) Indicator function for this time series.
Figure 2.15: Spectral estimate of the core 522 data from 22.7-34.7 Ma computed from 8 PRG tapers (black line, with blue confidence intervals), with a frequency resolution of 1.3 c/Myr. The spectral estimate from 29.4-34.7 is shown in blue, with orange blue confidence intervals has a frequency resolution of 1.3 c/Myr. Spectral estimate from 22.7-28.7 is shown in red with pale pink confidence intervals. The estimate has a frequency resolution of 1.4 c/Myr. All confidence intervals shown are jackknife 95% intervals.
Ma at most frequencies. These results suggest that the behaviour of the magnetic field changed during this 12 Myr period. We also see a peak at a frequency of $8 \text{Myr}^{-1}$ which was reported in Constable et al. (1998). These results are considered in detail by Smith-Boughner et al. (2011).

### 2.7 Discussion and Conclusions

Our goal in this work was to evaluate the utility of the MBG and PRG adapted multi-tapers when applied to a variety of geophysical data sets and develop tools to evaluate their performance and reliability. We have applied the basic techniques of Fodor and Stark (2000) for calculating a set of multi-tapers for a time series with data missing with some modifications to the numerical algorithms. Prolate tapers with gaps (PRG) are calculated using numerical integration while minimum bias tapers with gaps (MBG) are computed using a QR decomposition performed by a series of Householder transformations. Confidence limits for the spectra are calculated accurately and quickly using the jackknife method of Thomson and Chave (1991).

Both the MBG and PRG tapers provide better spectral estimates than section averaging for most data sets. However, the adapted taper method does have some limitations. When there are many breaks in sampling spanning a variety of lengths, a threshold must be chosen and all gaps shorter than this must be interpolated. Often, this requires some trial-and-error to balance the trade-off between bias due to interpolation and bias from deteriorating filter characteristics. With each different level of interpolation, a new set of tapers must be calculated, which is quite costly when using PRG tapers. The algorithm we outlined allows the user to study the effects of interpolation and the filtering properties of a particular set of adapted tapers and choose the best approach. In evaluating the prolate and minimum bias tapers adapted for gaps, we ignored any potential benefits obtainable from using adaptive-weighting schemes in order to consider the ‘worst-case’ performance of the tapering functions.

Prolate tapers, which are designed to optimise the amount of energy concentrated within a chosen bandwidth, maintain a large amount of energy within a narrow band when computed for an intermittently sampled time series and can provide fine-
scale frequency resolution not available from MBG tapers or section averaging for most applications. With a small number of tapers, the adapted prolate tapers will generally provide smaller confidence intervals than either minimum bias tapers with gaps or section averaging. However for a set of more than \( NW \) tapers, PRG tapers have more energy at high frequency than the same number of MBG tapers. As a result, the minimum bias tapers with gaps are better tapering functions than the PRG tapers when a large number of tapers is required, especially when there is a large dynamic range in the spectrum (and no prewhitening is used).

There are also significant computational benefits to using MBG tapers rather than PRG tapers. A set of MBG tapers need only be calculated once for a given arrangement of data and the computation is faster and more numerically stable. One set of MBG tapers can provide different averaging bandwidths by changing the number of tapers used to compute the spectral estimate. With the PRG tapers, if the resulting bandwidth is not appropriate, a different value of \( NW \) must be chosen and the tapers recalculated. Each adapted minimum bias taper often windows only one segment of data, strongly resembling section averaging. In many of these cases the frequency resolution and protection against spectral leakage are quite similar to cleverly tailored section averaging, as we found in our test on the length of day data.

Both PRG and MBG tapers were applied to analyse a record of changes in the length of day from 1962 to 2009 where data had been removed from the record in various different arrangements. Spectral estimates from the complete data set were compared to adapted tapers and to estimates produced from multi-taper section averaging of the interrupted record. The LOD time series allowed us to study how the MBG and PRG tapers performed on a data set with a very large dynamic range with high amplitude, narrow band features. When the time series consists of equal length data segments with identical breaks in between, our experiments indicated that the resulting MBG tapers essentially performed an optimized section averaging. The two other tests conducted with the LOD data indicated that when the interrupted record has continuous data segments of various lengths, the PRG tapers give slightly better frequency resolution and more accurate recovery of the shape of spectral features, with much more certainty than section averaging. Our example here is purely illustrative. In practice, this data set would be
filtered to remove tidal signals and a prewhitening filter could be used to reduce dynamic range. For intermittently sampled data, the prewhitening would increase the length of any gaps by twice the length of the filter used, requiring new adapted tapers.

The analysis of the DSDP sediment core from the Oligocene was conducted using PRG tapers because they are more suitable for short gaps and provide the best possible frequency resolution. Results provide enhanced frequency resolution for the fragmented part of the record and unambiguously show different behaviour in the power spectral densities of the two time periods analysed (Smith-Boughner et al., 2011).

Our experiences so far suggest that we have a robust algorithm for evaluating the PRG and MBG spectral methods, and that it could be applied to a broad range of geophysical time series. We conclude that when a large number of tapering functions is required, our method of adapting minimum bias tapers to accommodate breaks in sampling will provide better frequency resolution, protection against spectral leakage and usually provides narrower confidence intervals than prolate tapers with gaps. When each MBG taper windows only one segment of data, the spectral estimate is nearly identical to an optimized section-averaging scheme. However, if fine-scale frequency resolution is required or the record has many breaks in sampling, then the additional computational difficulties of PRG tapers are worthwhile and will give better results than section averaging or MBG tapers. These techniques can also be extended to cross-spectral and transfer function estimation. The Multi-Gap code package (written in MATLAB) is available at http://earthref.org/ERDA/1494/.

### 2.8 Acknowledgments

We thank two anonymous referees for constructive comments and questions that significantly improved this manuscript. Bob Parker, Glenn Ierley and Duncan Agnew all provided useful advice on computational issues related to solving the multi-taper optimization problems. We also would like to thank Alexey Kuvshinov for providing the external magnetic field model. This work was funded under National Science Foundation Grants EAR 0337712, EAR 0809709, and NASA Earth and Space Science Fellowship.
Chapter 2, in full is a reprint of the material as it appears in Geophysical Journal International 2012, Smith-Boughner, Lindsay; Constable, Catherine. Spectral estimation for geophysical time-series with inconvenient gaps. 190: 1404-1422. The dissertation author was the primary investigator and author of this paper.

2.9 Appendix

2.9.1 Calculating Tapers- Prolates

The optimization problem which defines the prolates (Equation 2.4) has eigenvalues, $\lambda$, which are all very close to one and each other. This provides great energy concentration but poses numerical problems whenever prolate tapers are calculated (Percival and Walden, 1998). To deal with these difficult problems, several methods of computing prolate tapers exist. Bell et al. (1992) suggested computing tapers using an inverse iteration technique and LU decomposition (Press et al., 2007). In Thomson (1982), the prolate multi-tapers are calculated by using a differential equation derived by Slepian (1978) resulting in a tri-diagonal matrix which is simple to solve. Thomson (1982) also proposed calculating the tapers by using numerical integration to solve for the continuous prolate spheroidal wave function in Equation 2.14 where $c = \pi NW$.

$$\lambda_k v_k(x) = \int_{-1}^{1} \frac{\sin(c(x-y))}{\pi(x-y)} v_k(y) dy \quad (2.14)$$

With gaps in the data sampling, considerable care must be taken in choosing the optimal method for computing prolate tapers. The indicator function used in the optimization problem substantially increases the condition number of matrix, increasing the likelihood of accruing substantial round-off error. Because of the large condition number, we decided against using the inverse iteration approach to calculate the PRG tapers.

The numerical integration approach suggested by Thomson (1982) provide PRG tapers with even weighting across the time domain (shown in Figure 2.1a) and good filtering properties in the frequency domain. The integration is implemented using a
Chebyshev quadrature (Clenshaw and Curtis (1960)) scheme and an iteratively restarted Lanczos method (IRLM) (see Lehoucq et al. (1998) for further details) to solve the tapers for 580 sample points. To provide the best possible accuracy with the fewest iterations, we integrate each sample point separately to a tolerance on the order of $10^{-12}$. Several lower tolerances were tested but there was very little detectable improvement in the numerical properties despite the additional computational time required. Using this numerical integration scheme provides prolate tapers with gaps with very good energy concentration. These tapers are also adaptable to a variety of different intermittently sampled time series. The solver generally converges in less than ten iterations. However, because of the integration operation, the computational time increases dramatically for larger problems.

In order to reduce the computational time required and compute the tapers on a personal computer, we downsample the indicator function of length $N$ to length $n_{int} = N/r$ where $r$ is no larger than the length of the smallest gap in the time series of interest. Tapers are computed for the down-sampled indicator function and then interpolated using cubic spline interpolation. Figure 2.16 shows the first-order PRG taper for the indicator function in Figure 2.1c. The tapers are computed from down-sampled indicator functions of 580, 1450 and 2900 sample points for a time-bandwidth product of 12 and interpolated to a length of 14500 samples. For the three different problem sizes, the first order tapers have the same shape but differ slightly in the peak amplitude reached for each continuous segment of data. As the problem size increases, the peak amplitude reached in the smaller sections decreases as the amplitude for the longest two segments increases. This validates our choice of using 580 sample points because it provides the most balanced weighting of each segment of continuous data and can be computed quickly.

### 2.9.2 Calculating Tapers - Minimum bias

In principle, the sinusoidal approximation for the minimum bias tapers can be adapted for gaps by projecting the tapers onto the sampling basis (by multiplying by the indicator function) and orthogonalizing the set using a numerically robust technique. We used singular value decomposition to re-orthogonalize the projected sinusoidal tapers.
The resulting tapers are very jagged in the time domain, particularly as they approach a gap. The jagged time domain behaviour translated to erratic ringing and poor energy concentration in the frequency domain. Because of these properties, the sinusoidal approximations to the minimum bias tapers are poor windowing functions for incompletely sampled data.

As shown in Table 2.1, the optimization problem which defines the minimum bias tapers can also be adapted to compensate for gaps in the data record using an indicator function. Incorporating the indicator function into the optimization problem raises the condition number of the matrix by several orders of magnitude so once again, a numerically stable method for solving the eigenvalue problem is required. To solve the eigenvalue problem, we use a series of Householder transformations to perform a QR decomposition (for further details see Anderson et al. (1999)). This approach requires on the order of $n^3$ computations and because the matrix which defines this eigenvalue problem is full, it requires a large amount of memory. To alleviate some of these computational burdens, the indicator function is down-sampled (using the same method described above for PRG tapers) to only a few thousand sample points before tapers are computed. Then, cubic spline interpolation is used restore the original size while maintaining the frequency domain properties. Ten minimum bias tapers with gaps are shown in Figure 2.1(b). We note that the peak amplitude of the taper is inversely proportional to both the length of the data segment and to the number of tapers which window that region of the data for the examples we examined.

For a set 580 samples long with 14.5% of the data missing, a set of 30 MBG tapers take 0.75 seconds to compute and a set of 24 PRG tapers (NW=12) to a tolerance on the order of $10^{-11}$ requires 1 hours and 45 minutes with a 2.66 GHz Intel Core Duo 2 processor with 4 GB of RAM. For the PRG tapers, the computational time increases by a factor of 2 for every order of magnitude decrease in the error tolerance. The computational time grows like $n_{int}^2$. 
Figure 2.16: Prolate tapers with gaps for $k = 1$ and $NW = 12$, computed for the simulated arrangement of gaps in Figure 2.1c for 580 (solid line), 1450 (dotted line) and 2900 (dashed line) sample points then interpolated using cubic splines.
Bibliography


Chapter 3

Changing Spectrum of Geomagnetic Intensity Variations in a Fragmented 12 My Sediment Record from the Oligocene

Abstract

Time series of relative geomagnetic paleointensity variations derived from marine sediments can be calibrated using absolute data derived from igneous materials. The resulting records may be suitable for spectral analysis of geomagnetic dipole variations. This work re-evaluates the 12 My (22.74–34.77 Ma) sediment record from Deep Sea Drilling Project Leg 73, Site 522, that is a key data set for determining the paleomagnetic power spectrum in the frequency range 1-100 My$^{-1}$. The 12 My record is marred by uneven sampling, with the interval between samples ranging from 1–640 ky, and contains several gaps that are considered too long to interpolate. The relative intensity data are calibrated using 129 globally distributed absolute paleointensity data from the same time interval. The power spectrum of the resulting time series is estimated using direct multi-taper spectral estimation with prolate data tapers adapted to deal with missing sections in the time series. The longest record available for analysis is thereby extended from 5.3 to 12 My. The new paleomagnetic power spectrum confirms the pres-
ence of a broad spectral peak at around 8 My$^{-1}$ for the early Oligocene and uncovers a peak around 2.5 My$^{-1}$ in the late Oligocene. Both peaks may be linked to tiny wiggles in marine magnetic anomalies. The new analysis unambiguously verifies that there is lower overall power in the younger part of the record, where the reversal process appears to dominate the power spectrum of the paleosecular variation. A comparison of the late Oligocene spectrum with that of PADM2M, a model of paleomagnetic axial dipole variations for 0-2 Ma, reveals some broad similarities; both time periods have similar power levels and a reversal rate of 4 My$^{-1}$. During the early Oligocene the reversal rate is about a factor of two lower, the field strength is higher, and the secular variation is stronger, suggesting that a strong magnetic field inhibits reversals but produces more variability in field strength.

### 3.1 Introduction

Changes in the strength of Earth’s internal magnetic field are detectable on time scales ranging from months to hundreds of millions of years (Hulot et al., 2010). Direct observations of the intensity provide information over short time intervals but only since the mid-nineteenth century. Indirect measures of absolute field strength can be obtained from fired archaeomagnetic materials and igneous rocks using the Thellier-Thellier method or one of its many variants (Dunlop and Özdemir, 2007), and in principle extend the record throughout Earth’s history. However, absolute paleointensity (API) measurements tend to come in the form of spot recordings in time, rather than as time series of observations. They are also sparsely distributed geographically. Numerous additional data come in the form of time series of relative paleointensity (RPI) measurements from sediment cores for which changing concentrations of magnetic minerals throughout a sediment core can be accommodated by appropriate normalization. The remaining signal is used as a proxy for the variations in magnetic field strength. RPI require calibration against API data, but can supply records that span millions of years. The temporal resolution of RPI records is highly variable and depends on the chronological information available, rates of sediment accumulation, and the local sedimentary environment (Tauxe, 1993). When the temporal resolution makes it difficult to
recover short term variations (less than a few ky) the geocentric axial dipole hypothesis is often invoked implicitly and RPI records are considered representative of changes in axial dipole moment.

Fluctuations in the strength of the axial dipole part of the geomagnetic field have been studied in detail for the past few hundred (Finlay, 2008) and few thousand (Korte and Constable, 2006) years. Many relative and absolute paleointensity records are available for the last 2 My and these have been used to reconstruct variations in either virtual axial dipole moment (Channell et al., 2009; Valet et al., 2005; Guyodo and Valet, 1996, 1999) or paleomagnetic axial dipole moment (Ziegler et al., 2011) over time intervals ranging from 200 ky to 2 My. However, the availability of both relative and absolute measurements of paleointensity declines rapidly with increasing age and detailed variations in dipole moment before 2.5 Ma are not well documented.

For the past 2 My, PADM2M (Ziegler et al., 2011) provides a reconstruction of variations in paleomagnetic axial dipole moment on time scales longer than 10 ky. This model is based on a comprehensive set of sediment records, and calibrated by absolute data. PADM2M, other VADM models, and individual sediment records show that reversals and excursions are associated with lows in dipole moment (Lund et al. (2006); Laj et al. (2000, 2006)). Continuous records like these readily allow spectral analysis of dipole moment variations (Constable and Johnson, 2005) which may be useful for identifying distinct physical processes that influence the dipole moment. Such studies have sought evidence for external orbital influences on the geomagnetic field (Xuan and Channell, 2008; Thouveny et al., 2008) as well as for spectral signatures associated with the geomagnetic reversal rate, which varies throughout the geologic record. There is also increasing interest in comparing the frequency dependence of geomagnetic power spectra with output from numerical dynamo simulations (e.g. Driscoll and Olson (2009)).

Occurrence times for reversals of the magnetic field and the length of polarity intervals have often been modeled as a Poisson process (Cox, 1968, 1969; McFadden, 1984) without taking account of paleointensity variations. The Poisson process predicts an exponential distribution for the length of polarity intervals, and a simple form for the power spectrum (Constable and Johnson, 2005) that depends on average field strength,
reversal rate, and the time required for the field to change polarity. The distribution of reversal times from the seafloor magnetic anomaly record deviates from the simple Poisson model and this motivated searches for very short polarity intervals in the seafloor magnetic anomaly record (Blakely and Cox, 1972).

In some parts of the seafloor magnetic anomaly record for 0-180 Ma, short wavelength anomalies, known as “tiny-wiggles” or cryptochrons (Cande and Kent, 1995) have now been identified (see Gee and Kent (2007) for a review of magnetic anomalies). Blakely and Cox (1972) first attributed these short wavelength features to short polarity intervals, less than 60ky in duration, while Cande and Labreque (1974) suggested they might also be due to fluctuations in paleointensity. The resolution of the marine magnetic anomaly record is typically limited to features longer than 30 ky, so higher resolution data of the kind available from marine sediment cores is needed to study these variations in more detail. Tiny-wiggles are particularly common in the Eocene-Oligocene period, being widely reported during chron C12-C13, but are also common in chron C5 (9.74-10.95 Ma). Current work still debates their origin. For example, a study by Bowles et al. (2003) using sediment and near-bottom anomaly data attributed the C5 cryptochrons to paleointensity variations, while Roberts and Lewin-Harris (2000) found some of these features corresponding to short polarity intervals using sediment records. Conflicting interpretations have also been discussed for the anomalies from 30.48-33.54 Ma, during chron C12-13. Florindo et al. (2001) suggest that they represent short polarity intervals while Lanci and Lowrie (1997) prefer to relate them to periods of lower regional or global field strength. Several tiny-wiggles have been identified in previous analyses of the DSDP core 522 (Hartl et al., 1993; Tauxe and Hartl, 1997) and associated with major decreases in paleointensity at reversal boundaries. Lanci et al. (2005) collected directional information for a sediment core record from the Pacific Ocean that spanned the entire Oligocene but did not find any polarity chron or excursional directions corresponding to the cryptochrons identified by Cande and Kent (1995).

In this study we re-evaluate a 12 million year record of relative paleointensity variations from site 522 of the Deep Sea Drilling program which spans 22.74-34.77 Ma. Constable et al. (1998) examined the power spectral density of the relative paleointensities in the early Oligocene portion of the record using a 5.3 My interval from 29.4-
34.7 Ma and found enhanced power around frequencies of \(8 \text{ My}^{-1}\) which they associated with tiny-wiggles. Because of intermittent sampling, no high resolution spectral estimate could be computed for the later part of the record (from 22.74–28.77 Ma). Here we provide a brief description of the data, use global API information to calibrate the RPI record and exploit a new method for analyzing time series with large sections missing to re-examine the variability as a function of frequency. A direct multi-taper spectral estimation technique which is adapted for gaps (Smith-Boughner et al., 2011) allows us to examine power spectra of the strength of the magnetic dipole moment throughout the Oligocene.

### 3.2 12 Million years of Paleointensity Data

The time series we analyze comes from work by Hartl et al. (1993) and consists of 2238 relative paleointensity measurements from a sediment core collected at 26°S and 5°W at Site 522 on Leg 73, of the Deep Sea Drilling Program. The paleolatitude at the beginning of the Oligocene was 33°S and changes by about 2° over the 12 My interval considered here. This remains a rare detailed sediment record spanning the time interval from 22.74–34.77 Ma and the longest quasi-continuous record of paleointensity available. The magnetic properties of these sediments have been well studied, initially for magnetostratigraphy (Tauxe et al., 1984; Tucker and Tauxe, 1984), and later for environmental and paleointensity work (Hartl et al., 1995; Tauxe and Hartl, 1997), and magnetite is the primary magnetic carrier. Isothermal remanent magnetization was determined to be the best normalization parameter for relative paleointensity variations (Constable et al., 1998).

The chronology for the DSDP 522 RPI record is based on correlation to 29 magnetic reversals from the timescale of Cande and Kent (1995). The sedimentation rate is primarily controlled by carbonate content and decreases throughout the core, from 10 \(m(\text{My})^{-1}\) in the early Oligocene to 5 \(m(\text{My})^{-1}\) in the later part of the core. Where possible cylindrical samples of 1 cm diameter were taken every 3-4 cm down the working half of the core, resulting in denser temporal sampling in the early part of the record. Each sample averages about 1 ky of geomagnetic behavior. The time series
of relative paleomagnetic measurements have uneven sampling intervals, which range from 1–640 ky between samples and an average sampling interval of 4 ky. Many more gaps occur in the younger portion of the data from 28.77 Ma onward.

The relative paleointensity time series is shown in Figure 3.1. The geomagnetic reversal rate changes throughout the Oligocene, ranging from an average of 1.6 per million years in the older part of the core (34.77–29.40 Ma) to 4 per million years in the younger section (28.77–22.74 Ma). Examining Figure 3.1, we see that the average intensity is higher for the older part of the record, when the reversal rate is lower. Tauxe and Hartl (1997) found a weak correlation between average intensity and length of the polarity interval. In subsequent work Constable et al. (1998) confirmed this and also found a strong positive correlation between the average field strength and its variability, a feature also evident in the time series plot. It is interesting to note that in the older portion of the record, where the reversal rate is lower, there are 16 cryptochrons reported by Cande and Kent (1995) (grey arrows at the top of Figure 3.1). In the younger, late Oligocene part, where the reversal rate is higher and the average intensity is lower, Cande and Kent (1995) found only 6 cryptochrons. Tauxe and Hartl (1997) found a correlation between the intensity data and a cryptochron at 25.3 Ma and report on another possible cryptochron around 25.7 Ma.

### 3.3 Calibration of Relative Paleointensity Data

Absolute and relative paleointensities provide different representations of the field because of the way in which magnetization is acquired in sedimentary and igneous materials. Magnetization is recorded in an igneous material nearly instantaneously as it cools, and the theory describing its relationship to absolute field strength is well developed. In contrast, a sedimentary magnetization is recorded as magnetic materials settle, compact and dewater, aligning with the ambient magnetic field. Suitable normalization allows recovery of relative intensity variations which will typically provide an average of the magnetic field strength over some period of time (hundreds or thousands of years). The theory describing this process is still being developed (see Tauxe and Yamazaki (2007) for a discussion of the issues surrounding both API and RPI measurements).
Figure 3.1: a) The geomagnetic polarity time scale based on Cande and Kent (1995) (with additional chron from Lowrie and Kent (2004)) is shown at the top of the figure with cryptochrons indicated by the grey arrows above. This illustrates the high reversal rate in the recent portion of the record, roughly 4 per My and a lower reversal rate of 1.6 per My during the older part of the record. Shown are the absolute VADM data (scale bar at right) from submarine basaltic glasses (indicated by grey crosses) and those of other igneous materials (grey open circles) from the PINT2010 database used to calibrate the relative paleointensity measurements (plotted in black, scale bar at left.) The grey shading at the top of the y-axis shows the position and size of gaps in the core record longer than 46 ky. There is only one sample point available within the interval 28.76-29.40. b) Plot indicating gaps in the RPI record when I(t), the indicator function is zero, only inter-sample intervals longer than 96 ky are shown.
Absolute paleointensity data this old are relatively sparse, but there are 129 globally distributed absolute paleointensity data available in this 12 My interval, compiled in the PINT2010 database (Biggin et al., 2010). These intensities are converted to virtual axial dipole moment (VADM) after first calculating the paleolatitudes of the data. Paleolatitudes for samples on all plates except the Pacific were calculated using apwp.py, distributed in the PMAGPY software package (available from L. Tauxe at http://magician.ucsd.edu/Software/PmagPy/), which makes calculations based on the plate reconstructions of Besse and Courtillot (2002). A few paleolatitudes for samples on the Pacific plate were taken from their original publication (Tauxe, 2006). The temporal distribution of the data is shown in Figure 3.1; the bulk of the API data available is in the older half of the record, with very few samples between 22-26 Ma and a large cluster of API measurements around 29 Ma, where a large gap in the core measurements occurs. The paleolatitudes of the API data are distributed across a range from 51°S to 54°N.

We calibrated the 522 RPI time series to paleomagnetic axial dipole moment using the maximum likelihood methods developed by Ziegler et al. (2011), and only outlined here. First a very low resolution cubic B-spline model is generated from the VADM data using a regularized least squares fitting procedure. The preliminary model is then used to calibrate the RPI record using the median ratio of the spline model to the RPI record at all RPI data points. The algorithm then generates a regularized maximum likelihood fit to the RPI and VADM data jointly, fine tunes the scale value, and produces a final model fit to the joint data set. We note here, that the maximum likelihood model is not used for further analysis in this study. We use the calibration value found in the above procedure to directly scale the 522 RPI record, and use the calibrated record in further analysis. Also, we do not account of the paleolatitude changes of the 522 drill site over the course of the 12 My, when using the time series. The paleolatitude changes only 2° over the Oligocene, so the effect of paleolatitude changes on the VADM calibration would be a minor decrease in strength for the younger part of the record.

The calibrated mean axial dipole moment for the record is 53 ZAm². The younger and older portions of the calibrated record have distinctly different means (40 ZAm² for 22.5-28.8 Ma and 58 ZAm² for 29.4-34.7 Ma), as is visually evident in Figure 3.1. This
difference was seen previously by Constable et al. (1998) in the uncalibrated record. For context, the raw mean of the 129 VADM values is 44 ZAm$^2$, but this takes no account of the temporal distribution of the data, which is a much sparser sampling of the field than the calibrated 522 record. Notably, the VADM data overlapping the younger portion of the 522 record have a higher mean (47 ZAm$^2$) than the data overlapping the older portion (41 ZAm$^2$). This is opposite to the trend in the RPI record.

The mismatch between the RPI and API trends is attributed to several sources. Some of this mismatch can be attributed to the small number of absolute paleointensity data, in contrast to the dense temporal sampling by the RPI record. The API data come from many sources and locations, but it is notable that the older half of the time interval has a large number of submarine basaltic glass data (plotted as grey crosses in Figure 3.1). On average, submarine basaltic glasses give lower VADMs than other igneous materials (Tauxe, 2006). Temporal sampling and quality issues with all the data sources clearly play a role in this difference. The lower VADM mean in the older half of the igneous record may be related to the different VADM source materials. Additionally, there is large uncertainty in the age of the API data, with age uncertainties ranging from a few hundred thousand to several million years (see PINT2010 database for details) make it difficult to identify any trends in the API data throughout the Oligocene.

### 3.4 Spectral Analysis of Time Series with Gaps

#### 3.4.1 Spectral Estimation Techniques

The simplest direct spectral estimator, the periodogram uses the square of the magnitude of the Fourier coefficients. Multi-taper methods produce a robust spectral estimate by using a set of orthogonal data tapers to window the data in the time domain and generate a suite of independent spectral estimates that are averaged to reduce variance. The data tapers usually result from the solution of an optimization problem. Our method uses the prolate multi-tapers of Thomson (1982). These functions minimize the amount of broadband bias in a spectral estimate by concentrating the energy, \( \lambda \), of the windowing function into a chosen bandwidth, \( 2W \). The resulting tapers are characterized by “\( NW \)” known as the time-bandwidth product, with \( N \) indicating the number of
equally spaced data.

We use a technique described in some detail by Smith-Boughner et al. (2011) to adapt the optimization problem for the prolate tapers to accommodate gaps by requiring the tapers to be zero when a time series sample is not available. We compute prolate tapers for an intermittently sampled time series, hereafter known as the prolate tapers with gaps (PRG), using a numerical integration scheme based on Chebyshev quadrature and an iteratively restarted Lanczos method (see Smith-Boughner et al. (2011) for details). For the 12 My paleointensity record (Figure 3.1) the resulting PRG with a time-bandwidth product of 8 are shown in Figure 3.2.

The frequency domain properties of the prolate tapers with gaps can be determined empirically by examining the average transfer function of the set of $K$ tapers, $H_K(t)$, where $v_k(t)$ denotes the $k^{th}$ taper.

$$H_K(f) = \frac{1}{K} \sum_{k=1}^{K} \left| \frac{1}{N} \sum_{t=0}^{N-1} v_k(t) e^{-2\pi ft} \right|^2$$  \hspace{1cm} (3.1)

The average transfer function of the prolate tapers and tapers with gaps used in this analysis is shown in Figure 3.3 for positive frequencies, the negative frequency response is a mirror image. Figure 3.3(a) compares the average transfer function of the prolate tapers with gaps from Figure 3.2 to the standard prolate tapers. Confidence intervals for the spectra produced from these adapted tapers are computed empirically using the BCA (bias-corrected and accelerated) method from 1000 bootstrap resamplings of the singly tapered spectral estimate (for further details see Davison and Hinkley (1997).)

### 3.4.2 Sampling Intervals of DSDP Site 522

The paleointensity record from DSDP Site 522 is intermittently sampled. Sample spacings vary between 1 and 640 ky in length. Typical spacing between samples was 4-8 ky, and the 11 gaps longer than 0.046 My are shown across the top of Figure 3.1. Large breaks in sampling occur in the gap between core segments or regions where the sediment was too disturbed for sampling. Breaks also occur if the samples fail to produce meaningful paleomagnetic results or when the core material was unavailable because it had been taken for other studies.
Figure 3.2: The prolate tapers with gaps for a time-bandwidth product (NW) of 8, shown are the set of 12 tapers used to analyze the entire 12 My calibrated paleointensity time series, from 22.74-34.77 Ma. The eigenvalues, $\lambda$ indicate the amount of energy contained within the desired averaging bandwidth. The theoretical bandwidth for this set of 12 tapers is 1.3 cycles per million years.
Figure 3.3: Comparison of the average transfer functions, $H_K(f)$ of the prolate tapers with gaps used in analysis. The half-bandwidth, $\frac{w}{2}$ for a set of $K$ tapers is determined by the mid-point of the steepest drop in amplitude. a) Average transfer function of the tapers shown in Figure 3.2 with $NW = 8$, used for analysis of the entire record, spectral estimate is shown in Figure 3.4. b) Transfer function of prolate tapers with $NW = 4$ computed for data availability between 22.7-28.7 Ma, spectral estimate from 6 tapers is shown in Figure 3.5. c) Prolate tapers with $NW = 4$ used for robustness test on data from 29.4-34.7 Ma. The indicator function corresponding to these data tapers is shown in Figure 3.7, the spectrum from 6 tapers is shown in Figure 3.8.
Our technique for analyzing time series with gaps is designed to improve the resolution and reliability of a spectral estimate at low frequencies when there are long gaps in the time series that are difficult, if not impossible, to interpolate. The frequency domain properties of the PRG tapers deteriorate rapidly when there are many short segments of data and few long continuously sampled sections. This sediment record contains gaps which vary widely in length, and to get the best results from this technique we interpolate the shortest gaps in the time series.

We use a two-step scheme to interpolate the time series. First, to even out the sampling rate, we use Akima (1970) spline interpolation to produce an evenly spaced data series with $\Delta t = 4$ ky whenever the interval between samples was 8 ky or less. No Akima spline interpolation was used if the time between samples exceeded 8 ky. This left 182 gaps in the record with an average length of 20 ky and 11 breaks in sampling longer than 46 ky; 99.5% of the intervals between samples are shorter than 46 ky. Because we are concentrating on the low frequency portion of the spectral estimate, this interpolation should have a negligible effect on our results. Next, we interpolate all breaks shorter than 96 ky in the time series using an auto-regressive filtering scheme. Filter coefficients were computed using a modified-covariance technique of Marple (1991) from the data immediately preceding and following a break in sampling. The two filters were used to create forward and backward predictions from the data immediately preceding and following a gap. These predictions were averaged to fill in the missing data. Using an auto-regressive scheme ensures that the frequency content of our interpolated data approximates large scale spectral features of the nearby data without introducing additional complexity. After this interpolation, there are 5 remaining breaks between samples, all lying in the 22.74-29.40 Ma interval, and longer than 96 ky. We used this revised set of “gaps” in computing PRG tapers for our spectral estimation.

### 3.5 Spectrum of the DSDP 522 Paleointensity Record

Now we present the results of our new spectral analysis on the DSDP Site 522 calibrated paleointensity data series. We begin with an analysis of the entire 12 My record, which extends the results of Constable et al. (1998) who derived spectra for
fragmentary pieces ranging up to 5.3 My in length. Then we subdivide the record into an early and late part before and after the large break in sampling between 28.76-29.40 Ma. This further analysis allows a test of whether the spectrum differs in the two time intervals, and of the robustness of our PRG spectral estimates. The younger piece from the late Oligocene spans 22.74-28.76 Ma and has several large gaps, so it requires our PRG taper technique to produce meaningful spectral estimates. The older portion, spanning 29.40-34.77 Ma, has no gaps after interpolation and can be analyzed using standard prolate tapers. The time-bandwidth parameters for both the PRG and PR tapers used for the subsequent analysis were chosen so that the frequency resolution and out-of-band energy rejection of the resulting spectral estimates are similar to those for the entire 12 million years.

3.5.1 Entire 12 My of Paleointensity Data

Prolate tapers with gaps with a time-bandwidth product of 8 were calculated incorporating all gaps in the time series greater than 96 ky. The time-bandwidth parameter, NW, was chosen to balance the trade-off between a large number of tapering functions available, to give sufficient number of tapers to compute reliable confidence intervals (see Davison and Hinkley (1997) for more details) and a small averaging bandwidth, w to produce a high resolution spectral estimate. We chose to use 12 of these functions for the subsequent analysis (shown in Figure 3.2), all with eigenvalues greater than 0.79, indicating the amount of energy contained within our desired bandwidth of 1.4My\textsuperscript{−1}. The transfer functions for this set of tapers are shown in Figure 3.3a where they are compared with the equivalent complete prolate tapers. The PRG tapers have only a slightly larger bandwidth than the standard PR tapers but a much higher level of energy outside the desired bandwidth. This increase by several orders of magnitude, reduces their ability to protect against spectral leakage. The differences between PR and PRG tapers are consistent for all PRG tapers used in the subsequent analysis of the paleointensity record. Given the limited dynamic range of the resulting spectra, we expect the bias from this to be small.

Using these prolate tapers with gaps, we estimate the power spectral density of the entire calibrated paleointensity record (solid black line in Figure 3.4 with 95%
**Figure 3.4:** Spectral estimate of the entire 12 My record produced from a set of prolate tapers with 5 gaps of 0.01 to 0.64 million years in length and a time bandwidth product of 8; spectrum has a frequency resolution of 1.4 Myr$^{-1}$. There is a broad peak at 8 cycles per million years. This signal was also seen in Constable et al. (1998), a similar spectral estimate from Constable & Johnson (2005) is shown (broken line), scaled by 53 ZAm$^2$. 

[Graph showing spectral estimate]
bootstrap confidence intervals shaded in grey). We identify a broad peak at $8 \text{My}^{-1}$ like that seen by Constable et al. (1998) (also examined in Constable and Johnson (2005), using calibration of $53 \text{ZAm}^2$, and plotted with a broken line in Figure 3.4) in their spectral analysis of the 29.40-34.77 Ma interval.

### 3.5.2 Results for Younger Portion, from 22.74-28.76 Ma

Prolate gap tapers with a time-bandwidth product of 3.5 were calculated for the 22.74-28.76 Ma time period with the 4 remaining interruptions in sampling (these gaps are shown in Figure 3.1). The average transfer function of the PRG tapers is shown in Figure 3.3b.

The spectral estimate for the younger data from 6 PRG tapers is shown as the solid black line in Figure 3.5 with 95% bootstrap confidence intervals shaded in grey. We compare this to the (dashed and dotted black line) spectral estimate expected if the magnetic field intensity variations were only those associated with a Poisson reversal process with a rate of $4 \text{My}^{-1}$, and a sharp drop to a plateau of low field strength for 30 ky associated with each reversal (see Constable et al. (1998)). For a more realistic comparison, we use a spectral estimate of the PADM2M model (Ziegler et al., 2011) evaluated at thousand year intervals from 0-2 Ma. This has roughly the same reversal rate as the late Oligocene (spectrum of PADM2M model from 6 PR tapers with NW=4 is shown by the dashed line in Figure 3.5).

The variations in the intensity of the VADM in the late Oligocene are reasonably consistent with the idea that they are dominated by reversal processes, in the sense that the spectrum has about the same variance as the Poisson model in the frequency range 2–45 My$^{-1}$. However, there do seem to be significant differences at long periods and in the details elsewhere. Rather than leveling off at long periods as expected for the Poisson model, there is a broad peak around $2.5 \text{My}^{-1}$, corresponding to a period of around 400 ky, and even more energy at lower frequencies. To test the robustness of the peak at $2.5 \text{My}^{-1}$, we computed PRG tapers and spectra only interpolating gaps less than 46 ky, which resulted in tapers with poor filtering properties but nevertheless detected a very small peak around $3 \text{My}^{-1}$. We also did a test with a larger time-bandwidth product for the 5 gaps longer than 96 ky, which had a larger averaging bandwidth and resolved
Figure 3.5: Spectral estimate of the calibrated paleointensity record from 22.74-28.76 Ma. The estimate was produced by averaging estimates from 6 prolate tapers with gaps (PRG) with a time bandwidth product of 3.5 resulting in a frequency resolution of $1.48\text{ Myr}^{-1}$. The correlation between the spectrum of the data and the model of a Poisson reversal process (dashed and dotted line) indicates that the magnetic field during this period of time is likely dominated by reversal processes. The large scale trend of the spectrum is similar to the paleointensity fluctuations from the last 2 million years from Ziegler et al. (2011) (dashed line)
a slightly broader signal of similar amplitude and shape at $2.4\, My^{-1}$. Two broad arcs of higher spectral amplitudes are visible around 23 and $35\, My^{-1}$, corresponding to periods of roughly 43 and 28 ky respectively. The arc around $23\, My^{-1}$ more or less follows the decreases of the Poisson model of reversals (Constable et al., 1998) and may be related to the average time taken for intensity variations associated with the reversal process. There is also very close agreement between the spectra of the PADM2M model and younger Oligocene data near $23\, My^{-1}$ and the shape of the two spectra between 5-15 $My^{-1}$ and between 20-40 $My^{-1}$.

3.5.3 Results for Older Portion, from 29.40-34.77 Ma

The calibrated paleointensity measurements for 29.40-34.77 Ma are much less fragmented than the younger data; the three small breaks in sampling range from 50-96 ky long and are interpolated using the AR(3) filtering scheme outlined in Section 3.4.2. After interpolation, no breaks in sampling remain and we use 6 standard prolate tapers with a time-bandwidth product of 3 to produce a spectral estimate with a resolution of $1.32\, My^{-1}$ (solid black line in Figure 3.6 with 95% confidence intervals shaded in grey). The spectral density expected from a Poisson process with a rate of $1.6\, My^{-1}$ (broken black line) is also shown, as is the PADM2M spectra (dotted black line in Figure 3.6). Our spectral estimate of the older portion of the data has very different large scale behavior, and more power at frequencies higher than $15\, My^{-1}$.

For the older portion of the record, the variations in magnetic field intensity do not correspond with predictions from the Poisson model. Again there is a very distinct peak around $8\, My^{-1}$, seen in the analysis of Constable et al. (1998), which is more pronounced than that in the 12 My spectrum of Figure 3.4. This corresponds to fluctuations in the intensity of the VADM with a period of roughly 125 ky. Broad arcs of increased spectral amplitude around 17.5 and $38.7\, My^{-1}$ correspond to periods of 57 and 26 ky.

3.5.4 Test of the robustness of our method

To demonstrate the robustness of the results from the PRG tapers, we create an intermittently sampled record. To do this, we remove data from the interpolated time
Figure 3.6: Spectral estimate produced from 6 prolate tapers with a time-bandwidth product of 3.5 for a complete data set, yielding a frequency resolution of $1.32 \text{Myr}^{-1}$. A spectral estimate from the PADM2M model for 0-2Ma is plotted (dotted line) for comparison. Also shown is the theoretical spectral estimate from a Poisson model of a reversal process with a rate of 1.6 reversals per million years (dashed and dotted line).

Figure 3.7: Gaps, similar to the gaps in the younger Oligocene data are imposed on the older Oligocene data (from 29.4 to 34.7 Ma) to test compare the behaviour of PR and PRG tapers. The indicator function, $I(t)$ is 0 when there is a gap in the data at time $t$ and 1 otherwise.
Figure 3.8: Spectrum of data from 29.40-34.77 Ma from using 6 PR tapers on the complete record (blue line with 95% confidence region shaded in light blue) and 6 PRG tapers (red line with 95% confidence region shaded in light red) with gaps imposed to match gaps in 22.74-28.76 Ma (indicator function shown at bottom of Figure 3.7. The two spectra have similar frequency resolution, $1.32\text{Myr}^{-1}$ for the complete data and $1.33\text{Myr}^{-1}$ for the data with gaps. The average transfer function, $H_K(f)$ for these tapers is shown in Figure 3.3c.
series of the older data to match the 4 long gaps in the younger data from 23.52-28.76 Ma. The gap structure imposed and the resulting time series are shown in Figure 3.7. PRG tapers were computed and used to produce a spectral estimate. This spectrum (red line) is compared to that from 6 PR tapers on the interpolated data set (blue line) in Figure 3.8 with 95% confidence intervals shaded for both spectra. These spectral estimates are very similar in frequency resolution and, as expected, we see that the widths of the various features are similar in the two spectra. The PRG tapers (see Figure 3.3 (c)) have a much shallower fall-off in energy outside the desired averaging bandwidth with correspondingly less protection against spectral leakage. This produces a spectral estimate with slightly higher amplitudes than that from the complete prolate tapers but no other large differences. We believe this confirms that the differences we see between the older and younger portions using the PRG tapers for one analysis and the PR tapers for the other are a real feature of the geomagnetic field.

The power in the older portion of the DSDP Site 522 record at $8 \text{My}^{-1}$ is not seen in the younger data just as the signal around $2.5 \text{My}^{-1}$ is not seen in the older data. To examine the possibility that these signals exist in both parts of the record but are not detectable by our method, we perform two tests using the first 1.5 My of the PADM2M model of paleointensity, with one sample every thousand years (this gives a Nyquist frequency four times that of the 522 sediment core record), and introduce a synthetic quasi-periodic signal into the PADM2M model. For each experiment, we removed data from our test time series to synthesize the same availability of data samples as the Site 522 record (see the bottom of Figure 3.7 for the gap arrangement synthesized). Spectral estimates of the complete (using PR tapers) and intermittently (using PRG tapers) sampled model time series with a synthetic signal are compared to determine how accurately we can recover the synthetic quasi-periodic signal with intermittent sampling. To test the robustness of the $8 \text{My}^{-1}$ signal in the older data, we use a synthetic quasi-periodic signal of $32 \text{My}^{-1}$. The synthetic signal is four times higher in frequency than the $8 \text{My}^{-1}$ to compensate for the different sampling rates of the Site 522 (one sample every 4 ky) and the PADM2M model (one sample every 1 ky). No significant differences in the PR and PRG tapers were seen in the spectral estimate around $32 \text{My}^{-1}$. The results of this test lead us to conclude that there is not a significant signal in the younger data.
set at $8\,My^{-1}$. For the second test, we use a signal at $10\,My^{-1}$ as a proxy for the $2.5\,My^{-1}$ signal seen in the younger data (again, the synthetic signal has a frequency four times higher than the $2.5\,My^{-1}$ signal to reflect the different Nyquist sampling rates of the series). Because the $2.5\,My^{-1}$ signal is at a much lower frequency, a higher amplitude is required to detect it in the spectral estimate from the intermittently sampled data. With this higher amplitude signal, the spectral estimates from the PR and PRG tapers detected signals with similar amplitude but the shape of the peak was slightly distorted by the PRG tapers. These synthetic tests verify that the signal present at $8\,My^{-1}$ in the older record is absent in the younger data and it is unlikely that there is a significant signal in the older data at $2.5\,My^{-1}$. These tests also show that the amplitude and shape of the $2.5\,My^{-1}$ signal in the younger data are less robust than those of the $8\,My^{-1}$ in the older record.

### 3.6 Discussion

Analyzing the frequency content of paleomagnetic data from sediments is further complicated by uncertainties in the age of the data. The absolute age of the paleomagnetic sediments is typically estimated by correlating magnetic reversals recorded in the core with those given by the GPTS. For the site 522 data, 29 magnetic reversals in the record were matched to those in the GPTS of Cande and Kent (1995); the absolute age of these reversal boundaries has an estimated average error of 9%. The age of the remaining portion of the sediment core is given by assuming a constant sedimentation rate between the ‘tie-points’ and using linear interpolation to determine the relative age of the samples; this sedimentation rate is thought to reflect the average of the actual sedimentation rate. Guyodo and Channell (2002) modeled the effect that deviations from this assumed average sedimentation rate have on the spectral properties estimated from data auto-correlations and found significant variations in spectral content. Because a direct multi-taper spectral estimator is much more robust, our method should be much less sensitive to errors caused by fluctuations from the average sedimentation rate.

To study the effects of age errors on spectral content, we synthesized a different set of ages for this sediment record and compared spectral estimates of the data using
the original and altered ages. To build a new age scale, we split the magnetic reversal boundaries used for calibration into two pieces, one piece for calibration points younger than 29 Ma with an average chron length of 300 ky and the second piece for the calibration points older than 29 Ma, with an average chron length of 600ky. The age of the calibration points in both sections were shifted by adding a gaussian random variable to the ages; the variance of the gaussian was taken to be 10% of the average chron length of that section. New ages for the 2238 paleomagnetic samples were calculated by linearly interpolating the ages to compensate for the shift in the nearby tie-points. With these altered dates, the data was interpolated using the same scheme of Akima spline and AR(3) interpolation that was applied to the original data. New tapers with gaps were computed for the data with these altered ages and the spectral estimates of the original record and this altered record were compared. The frequencies of the main spectral features were shifted by $1 - 1.4 \text{My}^{-1}$ or less and the amplitudes of some of these features were also slightly different, likely also reflecting the different averaging resolution of the sets of PRG tapers used. At frequencies above 60 M\text{y}^{-1} the slight discrepancies between the spectra from the different time scales increased.

While the possible errors in the relative dates of the samples adds uncertainty to the spectral estimate, a larger source of possible error is the fragmented nature of the data set. To maximize the recovery of the low-frequency features of the spectra, we have interpolated all breaks in the time series of less than 0.1 M\text{y} (with the 99% of these being less than 0.046 M\text{y} in length). This interpolation introduces some bias in the high frequency region of the spectral estimates, mostly at frequencies above 22 M\text{y}^{-1}. Because of this, we refrain from interpreting the fine-scale features of the spectral estimates in this region. This is also the region where possible errors in age calibration would bias the spectral estimate. The precise frequencies of features in the low-frequency region are also sensitive to errors in age calibration and the resolution of the spectral estimate, which are on the order of 1.2 -1.4 M\text{y}^{-1}. However, possible errors due to the frequency resolution are larger than those due to age calibrations. From the tests of robustness and age alteration described above we believe that the large scale features of the spectral estimates produced from the 522 data are relatively reliable despite the uncertainties in the absolute and relative ages of the samples.
Figure 3.9: Top: Close up of the spectral estimates from the entire 12 My record plotted on a log-log scale (black line with 95% confidence region shaded in grey), the older portion (blue line with 95% confidence region shaded in light blue) and the more recent portion (red line with 95% confidence region shaded in light red). Bottom: Full view of the spectral estimates of the entire 12 My record (black line), the older portion (blue line) and the more recent portion (red line).
Figure 3.9 compares the spectral estimates of the entire 12 My Oligocene record (black) with the older (blue) and younger (red) parts of the record. The spectral estimate of the entire 12 My record is dominated by the behavior of the older portion of the record because there is substantially more power in the secular variation there than in the younger data. The broad peak at 8 cycles per million years seen in the early Oligocene is not present in the spectrum of the late Oligocene data, but a signal is observed around 2.5 My$^{-1}$. The older portion (early Oligocene) and more recent portion (late Oligocene) are similar in amplitude around frequencies of 13 and 24 cycles per million years. Starting at frequencies about 26 My$^{-1}$ the spectra diverge, with a lot more power in the spectrum of data from 29.40-34.77 Ma (earlier) than in data from 22.74-28.76 Ma. The pattern of secular variation is very different between the younger and older data, with changes in VADM for the younger data dominated by reversal processes. The high frequency behavior in the spectrum of the entire record is dominated by the early Oligocene portion of the intensity record. The change in the pattern of secular variation during the Oligocene occurs near the middle of the DSDP 522 record but the gaps make this difficult to pinpoint. However, the timing is similar to a change in the statistical properties of magnetic polarity intervals suggested by Lowrie and Kent (2004) at Chron 12.

Our results support the idea that when the magnetic field reverses more often and the average intensity is lower in the late Oligocene (younger data), there is less energy in the secular variation. When the reversal rate is lower in the early Oligocene (older data), there is more energy in the secular variation and higher average intensity. Although some numerical dynamo simulations also exhibit fewer reversals when the average field intensity is higher (Glatzmaier et al., 1999), our findings are not really compatible with theirs. The numerical simulations generally require higher Rayleigh numbers and more energy in the secular variation to increase the number of reversals (Driscoll and Olson (2009)). Low average field strength is correlated to high secular variation. We find the opposite: low field strength corresponds to low power in the secular variation. We do not believe our result is an artifact of our method or the data. Our comparisons of the spectral estimates of the older data from prolate tapers and prolate tapers with gaps in Figure 3.8 suggest that the PRG tapers would tend to overestimate, rather than underestimate, the power in a spectral estimate. It is also worth considering whether the
higher sedimentation rate in the early Oligocene could contribute to the differences in spectral amplitudes for the two time intervals. We can rule this out because the changes in sedimentation rate would only impact the high frequency portion of the spectrum. The time between samples is typically 4–8 ky, and the typical period averaged within each sample is 1–2 ky, so we would not expect any effects at frequencies below 125 My$^{-1}$, the Nyquist frequency for this record.

As mentioned previously, the variations in VADM intensity in the late Oligocene are dominated by reversal processes and show a pattern of secular variation which is very similar to the last 2 Ma, as characterized by the PADM2M model of Ziegler et al. (2011). The average reversal rate for the two time periods are identical, while the field intensities are very different. The average field strength for the past 2 Ma from PADM2M is 59 ZAm$^2$ with standard deviation of 24 ZAm$^2$, actually similar to that of the early Oligocene (with a mean of 58 ZAm$^2$ and standard deviation of 25 ZAm$^2$) while the late Oligocene has an average field intensity of 40 ZAm$^2$ and standard deviation of 18 ZAm$^2$.

However, reversal processes do not seem to have a large effect on the pattern of secular variation in the older part of the record when the average intensity is higher and the reversal rate is much lower. The broad peak in the older Oligocene data around 8 My$^{-1}$ reported in Constable et al. (1998) is not seen in the younger data; just as the peak around 2.5 My$^{-1}$ is not seen in the older data. However, these paleointensity fluctuations have similar amplitudes and widths in the spectra of the two time periods, and we speculate that they may be created by the same kind of source.

One potential candidate for that source might be climate variations resulting in changes in magnetic grain size and periodic contamination of the paleointensity record. For example, Xuan and Channell (2008) suggested that orbital periodicities in sediment record are related to grain size changes due to changing oceanic currents over orbital periodicities. We do not believe the quasi-periodic signals in the Site 522 record are a climate signal because of arguments presented in Constable et al. (1998). They did not find strong coherence at a frequency of 8 My$^{-1}$ between the NRM (natural remanence) and the magnetic susceptibility, which is strongly controlled by carbonate content. Nor is there coherence with other proxies for grain size such as SIRM (saturation remanence) and ARM (anhysteretic remanence), with SIRM being the most relevant here.
since that was used for the RPI normalization. This coherence is also very low at 2.5 $My^{-1}$. Previous analysis of the power spectra of the time series of potential normalizers (magnetic susceptibility, isothermal remanence and anhysteretic remanent magnetization) from Constable et al. (1998) also did not show a significant signal at frequencies of 2.5 and 8 $My^{-1}$, suggesting no substantial change in grain size at these frequencies, we repeated these analyses for the fragmented younger portion of the record and found similar results.

We suppose that these quasi-periodic signals are geomagnetic. Note, however, that despite some broad similarities to orbital eccentricity and obliquity cycles (Laskar et al., 2004), we do not mean to imply that these geomagnetic signals are necessarily related to orbital influences on the geomagnetic field. What is clear from the geomagnetic record is that when the reversal rate is higher, there is less variability in the VADM intensity, a lower average intensity and less frequent paleointensity fluctuations associated with cryptochrons. When the reversal rate is lower, there is more variability in the paleointensity fluctuations at most frequencies and more frequent fluctuations associated with cryptochrons and a higher average field intensity.

Constable et al. (1998) suggested that the signal in the older data around 8$My^{-1}$ could be connected to cryptochrons; we believe this may also be true for the signal around 2.5 $My^{-1}$ in the younger data. There are 14 cryptochrons reported in Cande and Kent (1995) over the 5 million years stretch of older data (2 more occur between 27.8-29.4 Ma, during the break in the sediment core) and 6 cryptochrons in the 6 million years of younger data. Thus these tiny-wiggles occur at a rate of roughly 2.8 $My^{-1}$ during the time period corresponding to the older data where we see a signal at 8 $My^{-1}$ and at a rate of 1 $My^{-1}$ during the younger period where the signal at 2.5 $My^{-1}$ is seen. If we suppose that tiny wiggles and reversals can be lumped together as one geomagnetic process (which may or may not lead to a reversal depending on how low the intensity falls, and whether an excursion occurs instead of a reversal) we find rather similar rates of 4.4 and 5 $My^{-1}$ in the respective parts of the anomaly record. Neither corresponds to the 8 $My^{-1}$ (or 2.5 $My^{-1}$), but the lower rates could in principle be attributed to the fact that not all tiny wiggles or cryptochrons are resolvable in the marine magnetic anomaly record. This would suggest that the occurrence of cryptochrons is
related to core processes leading to changes in paleointensity which appear to occur more frequently when the reversal rate is low.

### 3.7 Conclusions

We have calibrated the relative geomagnetic paleointensity record from Site 522 of the DSDP using a new technique from Ziegler et al. (2011) and finding an average field strength of $53 \text{ ZAm}^2$ for the time interval 22.74–34.77 Ma. We have used a new method for spectral estimation of an intermittently sampled time series to analyze the early Oligocene (older data), late Oligocene (younger) and entire 12 My record of calibrated paleointensities using prolate tapers adapted for gaps.

The new technique for analyzing intermittently sampled time series allowed us to make detailed comparisons of the older portion to the gap-infested younger portion, as well as to the spectrum of the entire record. We see that, in the early Oligocene although the reversal rate is lower and the average field strength is higher, the field intensity is much more variable. The changes in field strength in the late Oligocene appear dominated by intensity drops associated with reversals which occur at the rate of $4 \text{ My}^{-1}$, higher than in the early Oligocene. This rate is similar to that over the past two million years, and the spectrum is similar to that of the PADM2M model covering this interval. This suggests that the reversal rate may be a controlling factor for the spectrum of all long period secular variation, not just the lows which occur near a reversal. Although the reversal rate is higher in the late Oligocene, there is much less variability in magnetic field intensity than during the early Oligocene. This observation supports the idea that a strong average geomagnetic field inhibits reversal even when its strength is highly variable.

There is a small peak in the spectrum around $2.5 \text{ My}^{-1}$ which is not seen in the early Oligocene; just as the signal around $8\text{ My}^{-1}$ seen by Constable et al. (1998) is only seen in the early Oligocene record. It seems likely that these fluctuations in paleointensity are related to the occurrence of cryptochrons, which occur much more frequently in the early Oligocene than late, possibly explaining the shift in the observed frequency.
3.8 Acknowledgements

This work was funded by NSF grant EAR 0809709. We thank Gauthier Hulot, Yohan Guyodo, and an anonymous reviewer for constructive comments.

Chapter 3, in full, is a reprint of the material as it appears in Physics of the Earth and Planetary Interiors 2011. Smith-Boughner, Lindsay; Ziegler, Leah and Constable, Catherine. Changing spectrum of geomagnetic intensity variations in a fragmented 12My sediment record from the Oligocene, Volume 188, Issues 3-4, October 2011, Pages 260-269. The dissertation author was the primary investigator and author of this paper.
Bibliography


Chapter 4

Using adapted multi-tapers to estimate cross-spectra and transfer functions of time series with missing data

4.1 Introduction

In geophysics, many things cannot be measured directly, for example density, temperature, composition or amount of melt in the Earth at depth. However, these properties affect the seismic wave speeds and electrical conductivity, quantities that can be inferred from the response of the Earth to earthquakes and electromagnetic waves.

These response functions, or transfer functions, describe a physical system by its frequency-domain response to input signals. Widely used in engineering and other physical sciences, accurate estimates of the transfer function of a system are extremely important. Along with considering the Earth as a system with an unknown response, the transfer functions of the instruments used to make those measurements must also be modeled in order to remove the effects of the instrument and recover the signal of interest. When the physics involved is well-known, these transfer functions often correspond to the linear differential equation which describes the system and can be used to solve the differential equation.

Often transfer functions are represented as a parametric filter, described by an
analytical expression corresponding to a family of linear, time-invariant filters (e.g. Butterworth, Chebyshev, Gaussian and cosine filters.) This also implies that the input and output characteristics of the system are stationary. The filter coefficients can be obtained by using the Laplace transform of the system to solve for the poles and zeros of the parametric filter (see Lathi (1998) for a full discussion).

For many geophysical systems, the response function cannot be well described by a parametric filter family and Fourier transform-based methods are used. One of the first methods used, the Blackman-Tukey method (Blackman and Tukey, 1958), estimates the covariance between the input and output and Fourier transforms the result. However, there are statistical problems with estimating the covariance of a finite series, the covariance estimate may be biased. More robust, modern estimates are obtained using cross-spectral methods (see Priestley (1981) for a discussion of various methods.) The statistical properties of the resulting response functions depend on the method used to compute the cross-spectra. As with power spectral density estimation, cross-spectra can be estimated using data tapers and/or section averaging by splitting the available data into pieces, then averaging the resulting cross-spectra. Averaging several estimates reduces the variance. Applying a data taper reduces the bias in the estimate. Using multi-taper techniques to estimate cross-spectra can reduce both variance and bias while producing high resolution spectral estimates (Walden, 2000).

Unfortunately, time series often have breaks in sampling. For transfer function estimation, the gaps could be either in the output or input series. Interpolating these series can prove problematic, since any bias introduced by the interpolator can introduce a large amount of bias into the cross-spectral density. Interpolating each data series individually can result in interpolated data sets with in-phase and out-of-phase components that do not match the original data series. As the amount of interpolated data increases, the coherence between the two series can be dramatically altered by an inappropriate interpolation scheme. As a result, estimating a transfer function when there are breaks in sampling is more difficult than estimating power spectral densities.

Calculating a transfer function between two data series with missing samples is a common problem in magnetotellurics and global geomagnetic depth sounding. These breaks occur because of instrument failures, geomagnetic storms or suspected contami-
nation from other magnetic fields. Both methods use a time series of geomagnetic vari-
ations external to the Earth as an input and a time series of the corresponding induced
internal variations as an output to estimate the electromagnetic transfer function of the
Earth. With the transfer function of Earth’s electromagnetic response, inferences can
be made about the electrical conductivity of the Earth at depth (see Kuvshinov (2012)
for a full review). Breaks in the time series reduce the amount of low-frequency infor-
mation available about our system, limiting the depths at which electrical conductivity
can inferred. Better models of electrical conductivity could be obtained by inversion of
robust, high resolution transfer functions with reliable uncertainties.

In Chapter 2, we used the methods of Fodor and Stark (2000) to demonstrate
estimates of power spectral density by adapting the optimization problems which define
the two multi-taper families to compensate for breaks in sampling (forcing the resulting
tapering functions to be zero where there are breaks in sampling.) These adapted win-
dowing functions provide a means to analyze a time series with breaks in sampling as
one long series, rather than using section averaging. This can provide spectral estimates
with higher frequency resolution than section averaging. These adapted multi-tapers can
also be applied to cross-spectra and transfer function estimation to improve the recovery
of long-period spectral characteristics.

To demonstrate this, I first provide background on cross-spectral estimation and
the corresponding confidence intervals. Transfer function estimation in the presence
of noise is discussed. Multi-tapers and the adaptations to compensate for breaks in
sampling are briefly described. Next, two realistic time series are simulated from a
specific transfer function and used for testing purposes.

The synthetic series are constructed as follows. Using a model of the external
and internal dipolar fluctuations of the Earth’s magnetic field from Kuvshinov and Olsen
(2006), I estimate Earth’s response to fluctuations in the externally generated dipole,
the output is an induced internal dipolar signal. Using this known transfer function,
I simulate a pair of input and output time series. The simulated time series are used
for testing. Data are removed from both the input and the output and transfer function
estimates from adapted tapers are compared to estimates from the uninterrupted records.
Confidence limits are investigated for coherence and transfer function estimates with
adapted tapers. Noise is added to both the input and output series to determine how well the adapted tapers estimate multi-variate spectra in the presence of noise. Several different arrangements of missing data are studied to reveal general characteristics and how these compare to estimates from section averaging.

4.2 Cross-spectral estimation with multiple data tapers

There are various ways to examine the relationship between two signals. For two time series, $x(t)$ and $y(t)$, that are assumed to be at least somewhat linearly related, the time-domain relationship can be expressed as:

$$y(t) = g(t) * x(t), \quad (4.1)$$

this is a convolution of the input series $x(t)$ by some function $g(t)$ of length $M$. In the frequency domain this relationship is described as:

$$Y(f) = G(f)X(f) \quad (4.2)$$

where $Y(f)$ and $X(f)$ are the Fourier transforms of the time series computed at a frequency, $f$. Cross-spectral methods can reveal a great deal about the relationship between the $Y(f)$ and $X(f)$ without needing to explicitly calculate $G(f)$. The cross-spectral density between two time series reveals the amplitude of the energy common to both signals and the phase between the series. For continuous time series, $y(t)$ and $x(t)$ the cross-spectra represent the Fourier transform of the true covariance $R_{yx}$. However for discrete, finite length time series, the covariance estimator is often quite biased and would produce a biased estimate of the cross-spectral density.

For discrete, finite length time series, the cross-spectral density between $x(t)$ and $y(t)$ is estimated by:

$$S_{yx} = Y(f)X^*(f) \quad (4.3)$$

where $X^*(f)$ denotes the complex conjugate of $X(f)$.

This is the cross-spectral equivalent of the periodogram and it is subject to the same frequency "ringing" effect. This ringing comes from using a finite length time-
series which effectively windows the data by a box-car function in the time-domain. This corresponds to convolution of the true spectral estimate with a \( \text{sinc} \) function in the frequency domain.

Better estimates can be made by partitioning both time-series, \( y(t) \) and \( x(t) \) into segments, each of length \( L \). A cross-spectral estimate is made between \( y(t) \) and \( x(t) \) from each segment and then averaged. This will remove some of the bias from using a finite number of samples. However, reducing the length of the time-series used to compute each Fourier transform lowers the frequency resolution. Multiplying each time series by a windowing function—also known as a data taper—before Fourier transforming can reduce the bias by windowing the data so that it heads smoothly to zero at the ends of the time segment. Using data tapers and section averaging will reduce the variance of the cross-spectral estimate and the bias. However using a single data taper to reduce ringing discards much of the information near the ends of the record; it is not an efficient estimator.

The multi-taper approach to cross-spectral estimation uses an orthogonal set of data tapers. The minimum bias tapers (MB) (Riedel and Sidorenko, 1995) and the prolate tapers (Thomson, 1982) (both discussed in Chapter 2) can be used for cross-spectral estimation to reduce the variance and bias of the estimate. To use multi-tapers for cross-spectral estimation, both time series are windowed with the same data taper \( v_k(t) \) with weights \( a_k \) from a set of \( K \) orthogonal windows. For prolate tapers, the weights \( a_k = \lambda_k / \sum_{k=1}^{K} \lambda_k \) correspond to the eigenvalues, \( \lambda \) of the \( k^{th} \) taper. Each eigenvalue corresponds to the amount of taper’s total energy concentrated within the frequency band \((-\omega, \omega)\) where \( W = 2\omega \) and \( W \) is set by choosing the time-bandwidth parameter, \( NW \). Using the multi-taper approach, the discrete Fourier transform of series \( x, y \), each of length \( N \) is given by:

\[
Y_k(f) = \sum_{t=0}^{N-1} a_k v_k(t)y(t)e^{-i2\pi ft} \quad (4.4)
\]

\[
X_k(f) = \sum_{t=0}^{N-1} a_k v_k(t)x(t)e^{-i2\pi ft} \quad (4.5)
\]
The cross-spectral density at frequency $f$ is given by the equation:

$$S_{yx} = \sum_{k=0}^{K-1} Y_k(f)X^*_k(f)$$

(4.6)

A cross-spectral density is usually a complex quantity. The in-phase components common to both series are represented by the real part of the cross-spectral density while the out-of-phase components are estimated by the imaginary components of the cross-spectral density. The phase between the series measures the time-delay between the signals at each frequency and is calculated at each frequency:

$$\phi(f) = \tan^{-1}\left(\frac{\Re\{S_{yx}(f)\}}{\Im\{S_{yx}(f)\}}\right)$$

(4.7)

If the phase is linear, the slope of the phase represents the time offset between the two series, with positive phase meaning changes in $y(t)$ occur before similar changes in $x(t)$. If the phase is zero then the two signals are in-phase and any changes in $y$ occur at the same time as similar changes in $x$.

To explicitly characterize the relationship between the input series, $x$ and the output series, $y$ described in equation 4.1, the transfer function, $G_{yx}(f)$ can be obtained using the cross-spectra computed above:

$$G_{yx}(f) = \frac{S_{yx}(f)}{|S_x|^2}$$

(4.8)

The transfer function characterizes the amplitude and phase relationship of the system independently of any input signal. In practice, this method of estimating the transfer function is quite sensitive to noise in $x(t)$. The consequences of this are discussed in section 4.6.

Computing the coherency, another complex quantity, normalizes the cross-spectral density by the energy in both signals.

$$\gamma_{yx}(f) = \frac{S_{yx}(f)}{\sqrt{|S_{yy}(f)S_{xx}(f)|}}$$

(4.9)

It represents the correlation coefficient between the in-phase and out-of-phase compo-
ments of two series as a function of frequency.

For most applications, the mean-squared coherence is used. It ranges from 0—meaning the two signals show no similarities at all at that frequency, to 1, meaning the changes in one time series at this frequency are identical to the changes in the other series but there may be a time-delay between the changes.

\[ |\gamma_{yx}(f)|^2 = \frac{|S_{yx}(f)|^2}{|S_{yy}(f)S_{xx}(f)|} \]  

(4.10)

4.3 Applying Adapted Multi-Tapers to Transfer Function Estimation

In Chapter 2, we extended the technique of Fodor and Stark (2000) and developed methods of evaluating the capabilities and limitations of these adapted multi-tapers for spectral estimation. A set of multi-tapers are adapted to compensate for breaks in sampling using an indicator function. Breaks in sampling are indicated by setting an indicator function \( I(t) \) to be zero when a sample is unavailable at time \( t \), otherwise \( I(t) = 1 \). With this indicator function the optimization problems which define the prolate and minimum bias tapers are adapted to force the resulting set of tapering functions to be zero when there is a break in the sampling. The adaptation for gaps in sampling increases the width of the tapering functions in the frequency domain and reduces the protection against spectral leakage; this reduction is much larger for the prolate tapers with gaps (PRG) than for the minimum bias tapers adapted for gaps in sampling (MBG).

To apply these methods to cross-spectral, coherence and transfer function estimation I restrict the discussion to minimum bias tapers with gaps (MBG). They are faster to calculate than the prolate tapers with gaps (PRG) and provide more flexibility to adjust the frequency resolution of the resulting multivariate estimates without needing to recalculate a set of tapers. However, to accurately estimate the cross-spectra of the signals the same tapers must be used to estimate all of the required quantities. With breaks in sampling, this implies that the two time series studied must have an identical indicator function. This isn’t always the case, so data must be either removed or carefully interpolated to give both time series the same indicator function.
Figure 4.1: Indicator function showing availability of data for an arrangement with 5% missing in three pieces, $I_t = 0$ when no data is available.

Figure 4.2: Sum of the absolute value squared of a set of tapering functions multiplied by the number of sample points (10461 points with data spread over 10953 sampling intervals) and the indicator function shown in Figure 4.1. The blue line is from 100 minimum bias tapers compared to 100 MBG tapers (green line).
We use an initial example indicator function shown in Figure 4.1 for a time series of 10954 samples. The continuous segments of data are of varying lengths and the breaks in sampling are quite short, 5% of the samples are missing. Minimum bias tapers with gaps computed for this indicator function are plotted in the appendix. The time domain weighting provided by a set of 100 MBG tapers is compared to 100 standard minimum bias tapers that have been multiplied by the indicator function in Figure 4.2. The ringing seen in the time domain comes from the higher order tapers which do not always approach a break in sampling smoothly.

The average transfer function of the MBG tapers for the indicator function from Figure 4.1 is shown in Figure 4.3. The average transfer function of the minimum bias tapers with gaps (solid lines) have a similar shape to the standard minimum bias tapers (broken lines) but the energy falls off more slowly. The MBG tapers also have a broader averaging bandwidth, closely resembling the characteristics of a set of minimum bias tapers with 20 more tapers and providing less protection against spectral leakage. These adapted tapers are used to estimate the transfer function between the synthetic time series.

### 4.4 Creating Realistic Synthetic Data

To determine the performance of the adapted prolate and minimum bias tapers for cross-spectral estimation we test the adapted tapers on two time series with a known transfer function. From the chosen transfer function, in this case the 1-D electromagnetic response of the Earth, a realistic input time series, $x(t)$ and output time series $y(t)$ must be simulated to match the relationship described in equation 4.1. Using a time series of external dipolar fluctuations $e_1^0(t)$, representing the magnetic field from the ring current, as an input and the internal dipolar field induced $i_0^0(t)$ from Kuvshinov and Olsen (2006), the Earth’s electromagnetic response, or transfer function is estimated. The two time series are plotted as black lines in Figure 4.4. The transfer function $G_{yx}(f)$ is estimated using equation 4.8 with 100 minimum bias (MB) tapers applied two time series with a four-hour sampling interval.

The transfer function with the input, $x(t)$ as the external dipolar magnetic field
Figure 4.3: Comparison of the average transfer function for a set of MB tapers (broken lines) and MBG (solid lines). The indicator function of the MBG tapers is shown in Figure 4.1. Shown with a sampling frequency, $f_s = 6$ cycles/day.
Figure 4.4: Time series of a model of external (top panel) and internal (bottom panel) dipolar fluctuations of the Earth’s magnetic field in response to changes in the ring current from Kuvshinov and Olsen (2006) shown in black. Synthetic data simulated from Dst and the method of Maus and Weidelt (2004) is shown in blue.
Figure 4.5: Estimated transfer function between input, $x(t) = e^0(t)$ and output, $y(t) = i^0(t)$ from Kuvshinov and Olsen (2006) (shown in black). The transfer function is computed using 100 MB tapers. The transfer function between the synthetic time series using 100 MB tapers is shown in blue. The transfer function corresponding to the 17 C-response values published in Kuvshinov and Olsen (2006) is plotted as red stars connected with a broken line.
and the output, \( y(t) \) as the induced internal dipolar field, is shown in Figure 4.5 as a black line. In this Figure, the red stars show the response function at 17 frequencies, transformed from the C-response values published in Kuvshinov and Olsen (2006). The most notable feature in the response function estimate is a steep drop near 1 cycle/day. This drop-out is related to the magnetic field generated from daily heating of the atmosphere by the sun; the resulting magnetic fields have been modeled and removed from the time series of internal dipole coefficients, \( y(t) \). The mean-squared coherence between the two signals is shown in Figure 4.8 where again this drop out is seen in the amplitude and the phase of the coherence. At frequencies above one cycle per day the coherence remains relatively high, but drops again at 2 cycles per day.

Using this transfer function, a realistic synthetic time series of the input, \( x(t) = e^{0_1(t)} \) and output, \( y(t) = i^{0_1(t)} \) is simulated using a technique devised by Maus and Weidelt (2004). The disturbance storm time index, Dst represents the internal and external contributions to Earth’s magnetic field that are associated with the ring current. Maus and Weidelt (2004) use this to partition Dst into a model of external dipolar fluctuations associated with the ring-current (input) and the induced internal dipolar response (output).

In this case the USGS 1-minute Dst index (Love and Gannon, 2009) is decimated by a factor of 240 to obtain one sample every four hours. The Fourier transform of the external, \( X_D(f) \) and internal dipole coefficients, \( Y_D(f) \) are obtained as follows:

\[
X_D(f) = \frac{1}{1 + G_{yx}(f)} D(f) \quad (4.11)
\]

\[
Y_D(f) = \frac{G_{yx}(f)}{1 + G_{yx}(f)} D(f) \quad (4.12)
\]

where \( D(f) \) is the Fourier transform of the decimated Dst time series.

The inverse Fourier transforms of \( X_D(f) \) and \( Y_D(f) \) yield the time series \( x_d(t) \) and \( y_d(t) \). The transfer function between these two synthetic time series is shown in Figure 4.5 (blue line) and compared to the pilot transfer function estimate from the Kuvshinov and Olsen (2006) model. The transfer function of the synthetic data matches the pilot estimate in amplitude almost everywhere, with a slight mismatch in the imaging of the drop-out at 1 cycle per day. The phase response between the synthetic time series
Figure 4.6: The estimate of the synthetic, $x_d$ (blue line) and original data, $x(t)$ (black line) of the external dipole coefficient time series are both computed with 100 minimum bias tapers. The energy level of the time series differs at 1 cycle per day and beyond 1.6 cycles per day.

Figure 4.7: The estimate of the synthetic, $y_d$ (blue line) and original data, $y(t)$ (black line) of the internal dipole coefficient time series are both computed with 100 minimum bias tapers. The energy level of the time series differs at 1 cycle per day and beyond 1.6 cycles per day.
**Figure 4.8**: Estimate of the mean-squared coherence between the synthetic data mimicking the model of internal and external dipolar fluctuations of the Earth’s magnetic field (blue lines) and compared to the data from Kuvshinov and Olsen (2006) (black lines). The estimate was made using 10954 sample points, spaced 4 hours apart with 100 minimum bias tapers, 95% confidence intervals are shown in grey.
is 5 degrees less than the pilot estimate at 1 cycle and near 1.8 cycles per day. These differences are more apparent when the power spectral density estimates of $x_d$ and $x$ are compared in Figure 4.6 but can also be seen in the comparison of the synthetic internal coefficients, $y_d$ and the model, $y$ shown in Figure 4.7. The differences are largely due to the different filtering applied to the Love and Gannon (2009) Dst series and the Kuvshinov and Olsen (2006) model. Because of the lack of noise in the $x_d(t)$ and $y_d(t)$, the two signals are extremely coherent (shown in Figure 4.8 as blue lines) while the Kuvshinov and Olsen (2006) time series, $x,y$ are as expected, not nearly so coherent.

### 4.5 Confidence Intervals

Constraining the uncertainty when working with multi-variate time series is quite challenging. Most often coherency estimates are used to study the relationship between the two time series. Because this is a complex quantity the variance in the amplitude and the phase of the coherency must be estimated. To accurately estimate the variance the probability distribution of the coherency (see equation 4.9) must be considered when choosing a variance estimator.

To understand the choice of estimator, I will first review the univariate spectral estimates, (discussed in Chapter 2.) Parametric methods of estimating the variance in a spectral estimate are not accurate when the spectra are colored (non-white). The uncertainty in the time-series must be estimated empirically. Each singly tapered power spectral density estimate follows a $\chi^2$ distribution, the average of the singly tapered estimates follow a chi-squared distribution with K degrees of freedom, where K is the number of tapers used to estimate the spectra. Most empirical estimation techniques work best with gaussian distributions. In Chapter 2, a logarithmic transform of the jackknifed estimates of the mean is used to normalize the distribution of singly-tapered spectral estimate. The variance of jackknifed estimates is calculated and used to obtain confidence intervals. This method from Thomson and Chave (1991) provided the most accurate estimates of uncertainty.
The complex coherency is given by:

\[
\hat{\gamma}(f) = \frac{\sum_{k=0}^{K-1} X_k(f) Y_k^*(f)}{\sqrt{\sum_{k=0}^{K-1} |X_k(f)|^2 \sum_{k=0}^{K-1} |Y_k(f)|^2}}
\] (4.13)

where \(X_k(f)\) and \(Y_k(f)\) are the Fourier transforms of the data series windowed by the \(k^{th}\) data taper. The complex coherency follows an F-distribution with parameters \((2, \nu - 2)\), where \(\nu = K\), the number of tapering function used. When more than 50 tapering functions are used this is often approximated by \(\chi^2\). In reality, even with more than 50 tapering functions, for colored spectra, the mean and variance of the coherence are best estimated empirically.

I use the method of Thomson and Chave (1991) for estimating the mean and variance of the coherency using jackknife resampling. Jackknifed estimators are sensitive to outliers, fast to compute, and overly conservative, ideal characteristics for estimating the expected uncertainty in a coherence estimate. The delete-1 jackknife resamples a data set by removing one datum to provide a resampling from which the statistic of interest, \(\theta\) is estimated.

To apply the jackknife to the estimate of coherency:

\[
\hat{\gamma}(j) = \frac{\sum_{k=0,k\neq j}^{K-1} X_k(f) Y_k^*(f)}{\sqrt{\sum_{k=0,k\neq j}^{K-1} |X_k(f)|^2 \sum_{k=0,k\neq j}^{K-1} |Y_k(f)|^2}}
\] (4.14)

However, jackknife resampling (and bootstrap resampling) is most accurate when applied to gaussian distributions (Davison and Hinkley, 1997). Because the coherency follows an F-distribution, jackknifed estimates of the coherency must be transformed to normalize their distribution, and in this case the inverse hyperbolic tangent is applied to the jackknifed estimates. Using this transformation on the jackknifed estimates of the amplitude of the coherency yields the statistic \(M\):

\[
\hat{M}(j) = \sqrt{2K-2} \tanh^{-1}(|\hat{\gamma}(j)|)
\] (4.15)
Under this transformation, $M_{(j)}$ is normally distributed with a mean of:

$$
\mu(M) = \sqrt{2K-2}\tanh^{-1}(\gamma) + \frac{1}{\sqrt{2K-2}},
$$

and a variance of 1.

From these transformed variables the variance is estimated:

$$
\hat{\sigma}_M^2 = \text{var}\{\hat{M}_{all}\} = \frac{N-1}{N} \sum_{j=0}^{K-1} [M(j) - \bar{M}]^2
$$

To use the estimate of $\hat{\sigma}_M^2$ to obtain confidence intervals for the amplitude of the coherence, $|\gamma(f)|$ the estimate of $|\gamma|$ from all $K$ tapers, must be transformed using the inverse hyperbolic tangent, confidence intervals calculated then transformed back. Assuming the transformed variables follow a gaussian distribution, a scale factor, $\beta$ is chosen so that $(1 - \alpha)\%$ of samples drawn from a gaussian distribution will be within $\beta$ standard deviations from $|\gamma|$. For 95% confidence intervals ($\alpha = 0.05$) $\beta = 1.96$.

Confidence intervals of statistic $M$ are given by:

$$
CI_M = \sqrt{2K-2}\tanh^{-1}(|\gamma|) \pm \beta \hat{\sigma}_M^2
$$

To obtain the confidence intervals for the amplitude of the coherency, $CI_\gamma$, the transformation must be reversed.

$$
CI_\gamma = \tanh \frac{CI_M}{\sqrt{2K-2}}
$$

Confidence intervals of the mean-squared coherence- the amplitude of the coherence squared- are obtained by taking the square of the $CI_\gamma$ confidence intervals.

The statistical significance of the coherence is related to the number of degrees of freedom used to compute the quantity of interest. The transformed coherence estimates from jackknife resampling follow a Gaussian distribution with the mean determined by equation 4.16 (however we neglect the $1/\sqrt{(2K-2)}$ term when a large number of tapering functions is used). If the true coherence is zero, then a mean-squared coherence
estimate from $K$ windowing functions would be distributed as:

$$C^2 = \tanh \left( \frac{Z\{1,1\}^2}{2K} \right).$$  \hfill (4.20)

where $Z\{1,1\}$ is a normally distributed random variable with a mean and variance of one. This implies that if 100 tapering functions were used then any coherence above 0.04 (0.1) would be above the 95% (99.95%) significance level and therefore likely to represent significant coherence between the data at that frequency.

Confidence intervals on the phase of the coherency or mean-squared coherence (MSC) are more difficult to obtain. Using the methods described in Thomson and Chave (1991), the uncertainty in the phase response is also estimated by applying a transformation and using jackknife resampling. First, jackknifed estimates of the coherency are divided by their magnitude, forming a new statistic, $e_{(j)}$.

$$e_{(j)} = \frac{\hat{\gamma}_{(j)}}{|\hat{\gamma}_{(j)}|}$$  \hfill (4.21)

Where the jackknifed estimate of the mean is given by:

$$\hat{\phi}_{(\cdot)} = \text{phase}(e_{(\cdot)}),$$  \hfill (4.22)

where $e_{(\cdot)}$ is the mean of the $e_{(j)}$.

The jackknifed variance of the phase, $\phi$ is given by

$$\hat{V}\{\hat{\phi}_{(\cdot)}\} = 2(m - 1)(1 - |\hat{e}_{(j)}|)$$  \hfill (4.23)

Which yields confidence intervals of the phase:

$$CI_{\phi} = \phi \pm \beta \sqrt{\hat{V}\{\hat{\phi}_{(\cdot)}\}}$$  \hfill (4.24)

where again, $\beta = 1.956$ for 95% confidence intervals.

The phase of the transfer function is the phase of the coherency between $x$ and $y$ and confidence intervals are computed in the same manner.

To estimate the sample variance of the amplitude of the response function $G_{xy}$
we must consider the probability distributions involved in equation 4.8 with input, $x$ and output, $y$. The numerator is distributed as $\chi$ with $K$ degrees of freedom where $K$ is the number of tapers used to compute the estimate. The denominator is distributed as $\chi$ with $K$ degrees of freedom, again where $K$ is the number of tapers used to compute the estimate. To empirically test this distribution, a bootstrap resampling of the transfer function estimated from 100 MB tapers was averaged across all frequencies (shown in red on Figure 4.9) for the time series $x, y$ from Kuvshinov and Olsen (2006). This yields a distribution which has more density near the mean and fewer sample points near the tails than the comparable Gaussian distribution. A transformation is needed to normalize the distribution.

Two different transformations were tested on empirical resamplings of the amplitude of the transfer function: the inverse hyperbolic tangent function $(\tanh^{-1}|G_{yx}(f)|)$ and the natural logarithm. Empirical bootstrap resamplings of estimates of the transfer function from 100 MB tapers using both transformations were tested. The inverse hyperbolic tangent function produced a much poorer match to a gaussian distribution.
than the untransformed resamplings. Estimates transformed using the natural logarithm closely match a comparable gaussian distribution, these results are shown in green on Figure 4.9. The natural logarithm, used in Chapter 2 to estimate confidence intervals of spectral estimates combined with jackknife resampling will be used to estimate confidence intervals.

A jackknifed estimate of $|G_{yx}(f)|$ is given by:

$$|\hat{G}(j)| = \frac{\sum_{k=0, k \neq j}^{K-1} X_k(f) Y_k^*(f)}{\sqrt{\sum_{k=0, k \neq j}^{K-1} |X_k(f)|^2}}$$

(4.25)

where $G(j)$ is the jackknifed estimate of the mean of $G_{yx}(f)$ without using the $j^{th}$ taper in the estimation.

Using the natural logarithm, $|\hat{G}(j)|$ is transformed to normalize the distribution.

$$\tilde{\sigma}^2 = var\{\ln|\hat{G}(\omega)|\} = \frac{N-1}{N} \sum_{i=1}^{N} (\ln|\hat{G}(j)| - \ln|\hat{G}(\omega)|)^2$$

(4.26)

Using the results of Thomson and Chave (1991) for spectral estimates, when applied to transfer function estimation, yields confidence intervals:

$$|\hat{G}_{yx}| e^{-t_{N-1}(1-\alpha/2)\tilde{\sigma}} < |G_{yx}| \leq |\hat{G}_{yx}| e^{t_{N-1}(1-\alpha/2)\tilde{\sigma}}$$

(4.27)

### 4.6 Computing Transfer Functions with Noise for Complete Time Series

The statistical properties of the transfer function depend on how the spectra are calculated as well as the noise characteristics of $y$ and $x$, here I reprise the discussion of Olsen (1998). There are three primary cases to consider: noise only in the output, noise only in the input, noise in both. Using the synthetic time series the transfer function between the external dipole fluctuations, $x_d$ and the internal, $y_d$ can be estimated in one of two ways. We could use $x_d$ as the input and $y_d$ as the output, or the inverse. In the
Figure 4.10: Estimate of the transfer function with $x_d$ as the input and $y_d$ the output. The transfer function from the synthetic data is shown in black. 5% noise in $y_d$, the output is shown in red and the estimate with 5% noise in $x_d$, the input is shown in blue. Here the sampling frequency, $f_s = 6$ cycles/day and 100 MB tapers.

In the time domain this is modeled as:

$$y(t) = g(t) * x(t)$$

(4.28)

Which yields the frequency domain relationship:

$$Y(f) = G_{yx}(f)X(f)$$

(4.29)

It is estimated using equation 4.8, by dividing the cross-spectra between $y$ and $x$
by the power spectral density estimate of \(x\). For the synthetic data, the response function with \(x_d\) is the input and \(y_d\) the output is shown in black on Figure 4.10.

If the input series has noise, the time domain relationship is:

\[
y(t) = g(t) \ast (x(t) + n_x(t)), \tag{4.30}
\]

which results in this relationship in the frequency domain:

\[
Y(f) = G_{yx}(f)(X(f) + N_x(f)), \tag{4.31}
\]

where \(N_y(f)\) is the spectrum of the noise. The transfer function estimate is biased:

\[
G_{yx} = \frac{S_{yx}}{|S_x|^2 + |N_x|^2} \tag{4.32}
\]

This means that any noise in the input time series biases the estimate of the response function downward.

Conversely, if noise was in the output the relationship in the time domain becomes:

\[
y(t) = g(t) \ast x(t) + n_y(t) \tag{4.33}
\]

which yields to this frequency domain effect:

\[
Y(f) = G_{yx}(f)X(f) + N_y(f) \tag{4.34}
\]

where \(N_y(f)\) is the Fourier transform of the noise.

\[
G_{yx} = \frac{S_{yx} + S_{n_y x}}{|S_x|^2} \tag{4.35}
\]

With noise in the output time series, the bias in the response function estimate depends on how correlated the input series is with noise in the output series. For most applications this bias is quite small.

I test these effects by adding gaussian noise to the model time series, the variance of the noise is some percentage, \(\eta\) of the total variance in the original signal; these results
**Figure 4.11**: Spectral estimate of $x_d$ with 95% confidence intervals, shown in red is the synthetic time series with dark red confidence intervals. Adding 5% noise to the synthetic time series yields the blue spectral estimate with dark blue confidence intervals. The spectral estimate with 10% added noise is shown in green with dark green confidence intervals are shown in Figure 4.10.

$$\eta_{input} = \frac{\sum_f |N_x|^2}{\sum_f |S_x|^2}$$  \hspace{1cm} (4.36)

When 5% noise is added to the input series, $x_d$ (shown in blue) the resulting transfer function hugely undershoots the amplitude at high frequencies. The 95% confidence intervals of the amplitude of the transfer function do not detect this deviation. When the same amount of relative noise is added to the output series (shown in red), $y_d$ the response function tracks the true amplitude but has much more variability at frequencies above 1 cycle/day. Above 1 cycle/day the spectral estimate of $x_d$ with 5% noise, shown in blue in Figure 4.11 begins to deviate from the model (red line). Further tests with 10% noise in the input showed the transfer function estimate substantially deviating from the model estimate at frequencies higher than 0.5 cycles per day. Noise in the output increases the uncertainty in the estimate but with noise in the input, the transfer function estimate becomes increasingly unreliable.

One method of accurately estimating the transfer function with a noisy input
**Figure 4.12:** Spectral estimate of the $y_d$ with 95% confidence intervals, shown in red is the synthetic time series with light grey confidence intervals. Adding 5% noise to the synthetic time series yields the blue spectral estimate with dark grey confidence intervals. The spectral estimate with 10% added noise is shown in green with black confidence intervals.

The signal is to reverse the relationship, to study the response function between the $x$ and the $y$ in the following way:

\[ x(t) = g(t) \ast y(t) \]  \hspace{1cm} (4.37)

\[ X(f) = G_{xy}(f)Y(f) \]  \hspace{1cm} (4.38)

To obtain $G_{yx}$ from $G_{xy}$:

\[ G_{yx} = \frac{1}{G_{xy}} \]  \hspace{1cm} (4.39)

Using the synthetic data, we estimate the response function with $y_d$ as the input and $x_d$ as the output. Again we add 5% relative noise to the output and estimate the response, shown as the blue line in Figure 4.13, as before the noise in the output increases the variance and bias of the response function estimate, particularly at high frequencies. Transformed jackknifed confidence intervals are shown in grey, at high frequencies the upper confidence interval climbs as high as 25 (not shown on figure).
**Figure 4.13:** Estimate of the transfer function between $y_d$ and $x_d$ with input and output reversed when 5% relative noise has been added to $x_d$-the output (blue line with grey confidence intervals). The transfer function estimate from 100 MB tapers with 5% relative noise has been added to $y_d$-the input is shown in red. The model transfer function is shown in black.
This is because the noise in $x_d$ increases the level of the noise floor, we see this in Figures 4.11 and 4.12 where using $2\eta_x$, adding twice as much noise raises the noise floor by a factor of two. In Figure 4.13, adding 5% relative noise to the input series lowers the response function estimate. This is less dramatic than the previous example with noise in the input because $y_d$ has less power than $x_d$, so total amount of power added as noise is less.

With noise in the input or output time series, the method of computing the transfer function can be adjusted to reduce the influence of noise. In this case, the time domain relationship becomes:

$$y(t) = g(t) \ast (x(t) + n_x(t)) + n_y(t)$$  \hspace{1cm} (4.40)

this yields the following frequency domain relationship:

$$Y(f) = G_{yx}(f)X(f) + N_x(f) + N_y(f)$$  \hspace{1cm} (4.41)

If there is noise in both the input and output then computing the transfer function becomes problematic and the strategy used depends on the ratio of the noise in the input to the noise in the output, $\psi$.

$$\psi = \frac{\eta_{\text{input}}}{\eta_{\text{output}}} = \frac{N_x/S_x}{N_y/S_y}$$  \hspace{1cm} (4.42)

To examine the effect of this, 10% relative noise was added to the $x_d$ synthetic time series and 5% relative noise has been added to $y_d$ synthetic time series. Estimated transfer function is shown as the blue line in Figure 4.14. With twice as much relative noise in the input as output, $\psi = 2$, the transfer function estimate is biased low at frequencies above 0.5 cycles per day, falling off to zero at the highest frequencies. The phase response tracks the synthetic model response quite well at frequencies below 1 cycle per day. Above these frequencies, the noise in both cause the phase to oscillate widely around zero, at high frequencies the 95% confidence limits suggest the phase is poorly estimated at 2.4 c/day and above.

In reality, most time series measurements have some amount of noise, so their
Figure 4.14: Estimate of the transfer function between $y_d$ and $x_d$ when 10% relative noise has been added to $x_d$ and 5% relative noise has been added to $y_d$, $\psi = 2$ (shown in blue with grey confidence intervals). The amplitude of transfer function has been adjusted using equation 4.43, shown in red with confidence intervals. The transfer function of the model is shown in black.
transfer functions would likely be estimated from a noisy input and output series, yielding extremely biased estimates. If however, the ratio of the input noise to the output noise, $\psi$ is known, then this bias can be corrected.

Using this partitioning coefficient, the optimal formula for computing the transfer function (Olsen, 1998) is given by:

$$G_{opt}(f) = \frac{G_{xy}(f)}{2} \left( 1 - \psi + \sqrt{(1 - \psi)^2 + \frac{4\psi}{\gamma(f)^2}} \right)$$ \hspace{1cm} (4.43)

where $\gamma^2$ is the mean-squared coherence between the two series.

The mean-squared coherence between the two series is shown in Figure 4.15, the coherence of the synthetic series is quite high but adding noise to both signals, with $\psi = 2$ (shown in blue) or $\psi = 0.5$ (shown in red) the coherence falls at an identical rate at higher frequencies. Using that estimate of coherence and the above equation, the transfer function estimate from two noisy time series can be adjusted for $\psi = 2$. The adjusted series is plotted in red in Figure 4.14. The adjustment increases the amplitude of the transfer function estimated to approximate the model at frequencies between 1-2 cycles per day. However, because the added noise destroys the coherence between the two series, the transfer function estimate still undershoots the true value by 0.2 at high frequencies.

In practice, the actual amount of relative noise is not usually known a priori and must also be estimated in some way. Despite this limitation, using the noise partitioning coefficient greatly improves the transfer function estimated. However, the reduced coherency between the signals at high frequencies results in unrecoverable bias in the transfer function estimate at high frequencies.

### 4.7 Using MBG tapers to estimate transfer functions with and without noise

Using the indicator function shown in Figure 4.1, MBG (minimum bias) tapers were computed to compensate for the breaks in sampling (tapers shown in the Appendix
Figure 4.15: Comparison of the mean squared coherence between $y_d$ and $x_d$ under various noise conditions. Shown in blue is the coherence when 10% relative noise has been added to $x_d$ and 5% relative noise has been added to $y_d$, $\psi = 2$. For $\psi = 0.5$, 5% relative noise has been added to $x_d$ and 10% relative noise has been added to $y_d$. The MSC of the model is shown in black. The broken line (dashed line) shows the 99.9% (95%) significance levels.
Figure 4.16: Comparison of the response function estimate using 100 tapers for the internal dipole coefficient model from the synthetic data series (black line). With breaks in sampling corresponding to Figure 4.1, the transfer function estimate from 100 MBG tapers is shown in blue. Shown with a sampling frequency, $f_s = 6$ cycles/day.
in Figures 4.26, 4.27, 4.28, 4.29, 4.30.) The largest break in sampling occurs at sample 6200 (Figure 4.1) and coincides with a steep, rapid drop in $x_d$ (see the top panel of Figure 4.4 in late 2003). This nonstationary event could introduce bias into the spectral estimate and transfer function. Transfer functions were computed between the synthetic time series. The synthetic series are extremely correlated. If the tapers with gaps destroy this correlation or bias the estimate of the phase response it will be extremely obvious.

First, a transfer function estimate is computed between the model synthetic series, $x_d$ and $y_d$ and compared to the transfer function from 100 MBG tapers applied to the two data series with samples removed to follow the indicator function shown in Figure 4.1. These results are shown in Figure 4.16 where the black line shows the model response and the blue line, the response from the MBG tapers - which resolves smoother features. The estimates are extremely similar in phase and are similar in amplitude at frequencies below 2 cycles per day. Above two cycles per day, the estimate from the MBG tapers is much flatter and of slightly lower amplitude than the model.

Then, the spectral estimates of $y_d$ is computed using the adapted tapers with gaussian noise added to the time series, this is shown in black on the top of Figure 4.17 along with the estimates from adding 10% relative noise to the input series $x_d$ shown in blue on the bottom panel. The frequency resolution of these estimates from adapted tapers is 0.08 cycles per day, (from the half bandwidth shown in Figure 4.3) while the standard tapers achieve 0.06 cycles per day. This difference is obvious in the spectral estimate, the features resolved by the MBG tapers are broader than those from the standard tapers. The largest difference is seen between 0.8-1.2 cycles per day, this could be from a spurious event in the time series which occurs during the long break in sampling. The larger dynamic range in the power spectral density estimate of $x_d$ shows higher power than the model in both the MB and MBG estimates. This is likely due to the noise added to both time series.

The average empirical transfer function of the MBG tapers, shown with solid lines in Figure 4.3 has a poorer resolution than the MB tapers, so the features recovered are much broader. The slight differences are seen around 1 cycle per day where both amplitude and phase of the response function drop steeply. The MBG tapers also do a reasonable job of capturing the steep drop at 1 cycle per day, a challenge to image
Figure 4.17: Comparison of power spectral density estimates using 100 MBG tapers (blue lines) to 100 MB tapers (red lines). Top: Synthetic model series of $y_d$, shown in black is compared to estimates with 5% added noise. Bottom: Synthetic model series of $x_d$, shown in black is compared to estimates with 10% added noise.
Figure 4.18: Comparison of the response function estimate using 100 tapers for the internal dipole coefficient model from the synthetic time series (black line) and from the synthetic series with 5% noise added to $x_d$ (red line). With breaks in sampling corresponding to Figure 4.1, the estimate from a series with 5% noise in $x_d$ is plotted in blue.

properly. This is also where the contamination occurs in the MB estimate because of the spurious event. With fewer tapering functions, 80 rather than 100, the estimate from MBG tapers exactly matches the model in phase and amplitude.

When noise is added to $x_d$, the transfer function estimated from the MBG tapers (blue line in Figure 4.18 tracks quite closely with the amplitude estimated from complete sampling with the same kind of noise (red line in same Figure). However, there are slight differences near 1 cycle per day, the MBG estimate is closer to the model than the MB estimate because the spurious event occurs during a break in sampling. The confidence intervals of both estimates are similar in size and shape. The largest difference between
the two estimates appears in the phase, the MBG tapers image a similar trend to what is seen in the estimate from noisy data in the phase at high frequencies but because the confidence intervals of both estimates are quite broad, this part of the phase response is unreliable.

Adding 10% noise to $x_d$ and 5% noise to $y_d$, the response function estimates from complete data and with long breaks in sampling (indicator function shown in Figure 4.1) are compared in Figure 4.19 with the corresponding confidence intervals shown. With twice as much noise in the input, $x_d$ as the output, $y_d$, the estimate from the adapted tapers tracks the amplitude and phase of the estimate from a complete data set quite closely, with small deviations between 0.7-2.2 cycles per day where the MB results are slightly higher than the MBG results. The confidence intervals of the estimate from MBG tapers are also slightly larger than those from MB tapers.

When the two estimates are adjusted using equation 4.43 differences between the amplitudes appear. This is due to the different estimates of coherence, shown in Figure 4.20. The coherence of the data with samples missing (shown in blue) is higher than that of the complete data (shown in red), especially around 1.8 cycles per day. This is likely due to bias from the spurious signal in the synthetic data that coincides with the longest break in sampling. Removing this spurious signal improves the coherence between the two signals at some frequencies. With the higher coherence, the adjusted transfer function amplitude from the MBG tapers (shown in magenta in Figure 4.19) matches some of the broad ripples seen in the model between 1.5-2.5 cycles per day. The adjusted MB estimate (in light blue) appears quite different from the model and the MBG estimate, with troughs in some places where the model has peaks. To verify the cause of these differences, the indicator function was reversed in time, so that the missing data segment begins at sample 4754, rather than 6200. In turn the MBG tapers were also reversed in time. This leaves the large spurious signal in the data with samples missing. Transfer function estimates from the complete data and the MBG tapers corresponding to this new indicator function are nearly identical, with slightly more bias in the MBG estimate. Even when the amplitude of the transfer function estimates were adjusted using equation 4.43, the estimates are quite similar. The corrected transfer function from MBG tapers is much smoother because the MBG tapers have less
Figure 4.19: Comparison of the transfer function estimates from 100 tapers with 10% noise added to $x_d$, the input and with 5% noise added to $y_d$, the output (red line). The model is shown in black. Creating breaks in the sampling of both noisy time series (with $\psi = 2$) corresponding to Figure 4.1, the transfer function estimate from 100 MBG tapers is shown in blue. Using equation 4.43, and a value of $\psi = 2$, both noisy estimates have been adjusted. The adjusted transfer function estimate of the complete data is shown in light blue, the adjusted estimate from MBG tapers is shown in magenta.
Figure 4.20: Comparison of the coherence estimate using 100 tapers for the internal dipole coefficient model from a complete data set (black line) and from a complete data set with 10% noise added to $x_d$ and 5% noise added to $y_d$ (red line with grey confidence intervals). With breaks in sampling corresponding to Figure 4.1, the estimate from a complete time series with noise, but no breaks in sampling is shown in blue (with light blue confidence intervals). The broken line (dashed line) shows the 99.9% (95%) significance levels.
Figure 4.21: Indicator function showing availability of data, $I_t = 0$ when no data is available. Out of a record spanning 10954 sample intervals, only 6824 data records are available, 38% of the record is missing.

frequency resolution and are slightly more coherent (blue line in Figure 4.20.)

To further test the capabilities and limitations of using adapted tapers for transfer function estimation, I use a different indicator function, one with nearly equal segment lengths and equal gaps in sampling. This indicator function is shown in Figure 4.21 with 38% of the samples missing. The resulting tapers for this indicator function are shown in Figures 4.31, 4.32, 4.33, 4.34, 4.35.

The minimum bias tapers computed for this indicator function are plotted in the appendix. However, it is more informative to look at the time domain weighting and the frequency domain behavior. The weighted sum of the absolute value squared of 100 minimum bias tapers with gaps for this indicator function (shown in Figure 4.21) is compared to the standard minimum bias tapers, that have been multiplied by the indicator function to show the breaks in sampling. As with the previous set of minimum bias tapers with gaps studied, these tapers also have a ringing effect when viewed as a set but not individually. Any ringing effects in the frequency domain should be apparent in the average transfer function of the set, which in Figure 4.23 does not appear to be the case. The average transfer functions of the tapers adapted for these equal segments has a significantly larger averaging bandwidth than the standard multi-tapers, this bandwidth
Figure 4.22: Sum of the absolute value squared of a set of tapering functions multiplied by the number of sample points (10954) and the indicator function shown in Figure 4.21. The blue line is from 100 minimum bias tapers compared to 100 MBG tapers (green line).

is also larger than the other set of adapted tapers studied previously. As more tapers are added to the set, the difference between averaging bandwidth of the adapted tapers and the averaging bandwidth of a set of standard tapers widens. For large sets of adapted tapers, more than 80, the protection against spectral leakage provided by these tapering functions also diminishes.

Because the remaining data segments are of nearly equal length, section averaging is used. To obtain a desired frequency resolution of roughly 0.06 cycles per day (the half-bandwidth is shown in Figure 4.24), 10 MB tapers would be applied to each of the 5 segments that are 1192 samples long and 7 tapers are applied to the last segment which is 858 samples long. The average transfer function of this set of tapers is shown in Figure 4.24 with a half-bandwidth of 0.04 cycles per day, similar to what would be expected for 60 MBG tapers.

From the average transfer function of the tapers, I would expect that the minimum bias tapers with gaps will provide a very similar estimate to that of section averaging. In fact, if no noise is added to the synthetic data, the MBG estimate is nearly identical to the estimate from section averaging with the same frequency resolution with
Figure 4.23: Comparison of the average transfer function for a set of minimum bias tapers (broken lines) and minimum bias tapers with breaks in sampling (see Figure 4.21). Shown with a sampling frequency, \( f_s = 6 \) cycles/day.
Figure 4.24: Comparison of the average transfer function for a set of minimum bias tapers (broken lines) and minimum bias tapers used in section averaging, with 7 tapers used on the last segment of data and 10 tapers applied to the remainder for a total of 57 tapers. Shown with a sampling frequency, $f_s = 6$ cycles/day.
Figure 4.25: Comparison of the response function estimate from data with 5% noise in the input (see Figure 4.21 for the indicator function).
nearly identical confidence limits. However, when 5% relative noise is added to \( x_d \) and transfer functions are calculated there are some large differences.

Using 60 MBG tapers, shown as the blue lines in Figure 4.25, the amplitude of the transfer function is biased low at frequencies above 1.2 cycles per day, the corresponding phase estimate become increasingly more erratic above 1.2 cycles per day. The section averaging example, with a total of 57 minimum bias tapers used, has higher frequency resolution than the MBG estimate and is closer to the true, noise free model (shown in black) but the phase response is identical to that estimated from the MBG tapers. These results could be improved using the noise correction described in equation 4.43 for cases when the exact amount of relative noise is known.

### 4.8 Discussion and Conclusions

The major goal of this chapter has been to present the general concepts of frequency domain transfer function estimation (including confidence limits) and assess whether the algorithm presented in Chapter 2 for power spectral density estimation on data series with inconvenient gaps can be usefully extended to the bivariate case. The assessment has been conducted using synthetic time series that mimic the kind of data acquired in geomagnetic depth sounding studies where frequency domain transfer functions are used to infer electrical conductivity profiles for Earth’s deep mantle. However, the overall results are generally applicable to estimating transfer functions in any linear system.

The use of simulated paired time series with a specified transfer function has allowed multiple useful tests. Firstly, the synthetic data (produced from the Dst index) were checked for compatibility with the specified transfer function and with real geomagnetic sounding data (Figures 4.5-4.8). We outlined the algorithm for acquiring empirical estimates of confidence intervals using the delete-1 jackknife on suitably normalized statistics (hyperbolic tangent transformation for coherency and natural logarithm for the transfer function). With these tools in hand a suite of experiments were conducted to test the impact of noise in the input, output, and both time series (Figures 4.9-4.14). As has been noted by others (Olsen, 1998) transfer function estimates are
particularly sensitive to noise in the input time series resulting in a bias of the gain to low values. For the output noise case, bias will only occur when the output noise is correlated with the input series. If noise in the output is low, there may be some benefits to switching the input and output series in the transfer function estimation. When both input and output are noisy and the ratio of normalized noise variances is known, Olsen’s (1998) strategy can be used to correct the resulting transfer function using equation 4.43. For the examples used here, the results are adequate at low to moderate frequencies, but increasingly discrepant above 1.5 cpd, where the added noise overwhelms the spectral properties of the model signal.

Two examples were used to test the impact of missing data. In the first, 5% of each time series was removed in 3 sections. One of the segments removed from $x_d$ contained a large non-stationary signal that presents challenges for spectral estimation. Using both MB and MBG tapers, produces spectral estimates that are quite similar to the model time series (Fig 4.16), and only slightly impaired frequency resolution in the MBG transfer function (Fig 4.17).

In the missing data case it was noted again that noise in both the input and output biases the transfer function estimate a great deal, and that if the ratio of the relative noise power in the input, $\eta_x$ to the relative noise power in the output, $\eta_y$ is known this bias can be partially corrected. The correction works well when the two time series are highly coherent. Even with noise in the input, output, or both, the minimum bias tapers with gaps still provide transfer function estimates that are as good as estimates made from the full time series.

The second missing data example studied time series with data segments of nearly equal length and 38% data missing. The MBG tapers produce a transfer function estimate with a phase response that is essentially identical to that computed with section averaging. However, section averaging is less sensitive to noise than the MBG tapers. In such cases, section averaging using a set of MB tapers provides more robust results.

We conclude from these experiments that, for times series containing data segments with widely varying lengths the use of MBG tapers with empirical estimates of the averaging bandwidth combined with accurate estimates of the relative noise between the input and output time series enables better frequency resolution and greater variance
reduction than is obtained with the section averaging approach.

These positive results from MBG tapers are particularly encouraging because of the need to use the longest time series possible for transfer function estimation in geomagnetic depth sounding. Long time series are essential to provide good frequency resolution and the low frequency estimates necessary to see deep into Earth’s mantle. The confidence intervals for the coherence and transfer function estimates quite accurately estimate the variance of the estimates and produce reliable confidence intervals.
4.9 Appendix

**Figure 4.26**: Minimum bias tapers with gaps (K=1-5) for the indicator function shown in Figure 4.1

**Figure 4.27**: Minimum bias tapers with gaps (K=6-10) for the indicator function shown in Figure 4.1
**Figure 4.28:** Minimum bias tapers with gaps (K=11-15) for the indicator function shown in Figure 4.1

![Graph showing minimum bias tapers with gaps (K=11-15)](image1)

**Figure 4.29:** Minimum bias tapers with gaps (K=16-20) for the indicator function shown in Figure 4.1

![Graph showing minimum bias tapers with gaps (K=16-20)](image2)
Figure 4.30: Minimum bias tapers with gaps (K=21-25) for the indicator function shown in Figure 4.1

Figure 4.31: Minimum bias tapers with gaps (K=1-5) for the indicator function shown in Figure 4.21
Figure 4.32: Minimum bias tapers with gaps (K=6-10) for the indicator function shown in Figure 4.21

Figure 4.33: Minimum bias tapers with gaps (K=11-15) for the indicator function shown in Figure 4.21
Figure 4.34: Minimum bias tapers with gaps (K=16-20) for the indicator function shown in Figure 4.21

Figure 4.35: Minimum bias tapers with gaps (K=21-25) for the indicator function shown in Figure 4.21
Bibliography


Chapter 5

Satellite Induction Study using
Adapted Multi-tapers

5.1 Introduction

In this chapter I will apply the methods developed in earlier chapters, using GDS (geomagnetic depth sounding) to recover Earth’s electromagnetic response to external magnetospheric variations. Observations of the vector magnetic field from 2001-2010 taken by the CHAMP low-Earth orbiting satellite are used to estimate the internal and external variations. Much of the theory related to this chapter has been introduced in Sections 1.5, 1.7 and 1.8.

Variations in the magnetic field of the magnetosphere (the purple region in Figure 5.1) induce currents to flow in the mantle. These currents generate magnetic fields. Studying the transfer function between the primary and induced fields yields information about the electromagnetic response of the Earth. With such response estimates, inferences can be made about the electrical conductivity of the Earth.

Globally, the one-dimensional electromagnetic response is still not well constrained, particularly at long periods. The frequency dependence of the skin depth described in section 1.7 implies that only long period magnetic field variations penetrate deep into the Earth. With data from magnetic observatories (Olsen, 1999; Khan et al., 2011), estimates of the long period response characteristics are available. However,
Sources of the Geomagnetic Field and the Modern Data That Enable Their Investigation

Figure 5.1: From Olsen et al. (2010), a sketch of the near-Earth electromagnetic environment.

because magnetic observatories are sparsely distributed over the Earth, with few estimates over oceanic crust, inferences about electrical conductivity from these data sets are biased towards continental conductivities.

To isolate the external field and corresponding internal response all other components of the near-Earth magnetic field must be modeled and removed. The largest component is from the Earth’s core and consists of large-scale structure (spherical harmonic degree, $n \leq 15$). The magnetic field generated from the crust and lithosphere varies extremely slowly and is considered static for these purposes. Because these fields change more slowly than the external fields, it is possible to isolate the more rapidly varying response induced by changing external fields.

The ring-current (Siscoe, 2011) generates a largely dipolar magnetic field which varies on timescales of minutes to years. Along with daily, monthly, seasonal and annual variations, the strength of the ring-current also changes in response to the 27 day solar rotation and the 11-year solar cycle. The symmetric ring current is typically around 20nT but, during magnetic storms can reach 300, even 1000nT. Using these strong magnetic field variations provides a simple geometry to investigate global electrical conduc-
tivity. To simplify the problem further, it is assumed that variations in the ring-current create a time-varying uniform magnetic field that induces a dipolar magnetic field internal to the Earth at the same frequency, in response. In reality, the ring-current is quite asymmetric. Love and Gannon (2009) found that the magnetic field on the dawn side is 18% less than Dst and 20% greater than Dst on the dusk side; the pattern of this asymmetry appears to be consistent across a wide variety of geomagnetic conditions (Siscoe et al., 2012).

Once the magnetic fields from the core and lithosphere are removed, the remaining signal includes contributions of both the internal and external magnetic field variations as well as noise, which includes other unmodeled magnetic field sources such as contributions from the asymmetric portion of the ring-current, the corresponding induced response and from the currents coupling the magnetosphere and the ionosphere.

Using data from magnetic satellites to study the global electromagnetic response allows for global coverage. Unfortunately, these satellites orbit either directly above the ionosphere or within the top 100 km. The largest signal from the ionosphere comes from the Sq or solar-quiet current system, a horizontal current driven by solar radiation on the sunward facing side of the Earth. The magnetic field generated by the Sq current accounts for nearly half of the magnetic field measured at 400-600 km altitude, corresponding to the orbit of CHAMP (Kuvshinov et al., 2007). The magnetic field from this current system has a quadrupolar structure; this violates our geometric assumption about the source field, so daytime measurements of the magnetic field cannot be used.

Time series samples from polar regions and the dayside should not be used to estimate the external and internal dipolar contributions. Measurements from these regions cannot be included in the time series. Large magnetic storms may introduce bias into the estimate. It is an open question as to whether very large geomagnetic storms could introduce a large amount of non-stationarity and overwhelm long period variations and may need to be removed from the time series.

Breaks in the time series reduce the amount of long period information we can recover and complicate the analysis of the entire record, as discussed in Chapters 2 and 4. Typically, long segments of missing data require splitting the time series into smaller pieces and using section averaging to produce a spectral estimate with less frequency
resolution. Short gaps in the time series can be interpolated without introducing bias at long periods. However, because the external and internal components are coupled and respond differently at each frequency, interpolation is quite complicated.

Now that 10 years of vector magnetic field measurements from the CHAMP satellite are available, I can extend estimates of C-response to longer periods. Although there are several 1-D estimates of electromagnetic response from satellites, all have been limited to periods of less than 150 days (Constable and Constable, 2004; Kuvshinov and Olsen, 2006). Conductivity models from these data sets have been limited to depths less than 900 km. If we assume the lower mantle has a conductivity of 10 S/m, then from the skin depth relation, to penetrate deeper than 900 km requires signals with periods of longer than 1 year.

In this chapter, I will describe the time series data used. The data processing steps required to estimate the internal dipolar, $i_1^0(t)$ and uniform external, $e_1^0(t)$ fluctuations from magnetic field measurements is discussed. To separate these contributions, an appropriate model of the main magnetic field must first be removed from the satellite measurements followed by a least squares fit to recover the external and internal field coefficients for each satellite pass. The difference in continuity of the sampling of $e_1^0(t)$ and $i_1^0(t)$ under various data selection criteria (e.g. magnetic local times) are explained and an interpolation method using data predictions from Dst (from Maus and Weidelt (2004) and discussed in Chapter 4) is applied using a pilot estimate of the response. The relative noise in both time series is estimated to improve the response function estimate using the work of Olsen (1998)(see equations 4.43). Several estimates of the electromagnetic and C-responses are discussed with variations in the data selection criteria and how the breaks in sampling are accommodated. These results are compared to previous 1-D C-response estimates.

5.2 Magnetic Satellite Data

Low-Earth orbiting satellite measurements of the magnetic field have greatly improved our understanding of the near-Earth magnetic environment. The launch of OGO-2 in 1965 provided the first precise measurements of magnetic field strength from space.
Figure 5.2: From Olsen et al. (2010), showing the orbital path of a satellite with the orbital inclination shown as a white arrow near the equator.

Table 5.1: High precision magnetic satellites, adapted from Olsen and Stolle (2012); Olsen (1999); Kuvshinov (2008). Note*: Swarm mission consists of three satellites, one with a higher altitude orbit and two lower satellites with 1° separation in longitude.

<table>
<thead>
<tr>
<th>Name</th>
<th>Lifetime</th>
<th>Altitude (km)</th>
<th>Type of Data</th>
<th>Incl. of Orbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magsat</td>
<td>Nov. 1979-May 1980</td>
<td>325-550</td>
<td>Vector &amp; scalar</td>
<td>97°</td>
</tr>
<tr>
<td>Ørsted</td>
<td>Feb. 1999-2012(sc.)</td>
<td>630-880</td>
<td>Vector &amp; scalar</td>
<td>97°</td>
</tr>
<tr>
<td>Swarm*</td>
<td>2013-</td>
<td>530, &lt; 450</td>
<td>Vector &amp; scalar</td>
<td>88, 87°</td>
</tr>
</tbody>
</table>

Vector measurements were available for a short time in 1979-1980 from MAGSAT. With the Ørsted (1999), SAC-C (2000) and CHAMP (2000) satellites, we now have over 10 years of vector measurements of the magnetic field near Earth (see Olsen and Stolle (2012); Kuvshinov (2008) for a summary). The mission durations and orbital parameters of each of these satellite missions are outlined in Table 5.1.

The CHAMP satellite data is used for this study because it is the longest record of satellite vector magnetic field measurements currently available. The satellite recorded the vector magnetic field, sampled once a minute, as it flew in a polar orbit for more than 10 years using a fluxgate magnetometer. The flux-gate vector magnetometer has a sampling rate of 1Hz and an expected error of 1nT. These vector data have an additional error due to the uncertainty in the attitude of the satellite, contributing 5-6 nT of uncertainty (Olsen et al., 2007). In order to accurately determine the position and attitude of the satellite, data from two star cameras are required. With only one star camera, the expected measurement error in the vector field increases by a factor of 2-3 (Holme and
**Figure 5.3:** The magnetic local time of each satellite pass of the available high quality data. Data points with local times between 18h and 6h are shown in red.

Bloxham, 1996). The data sets are provided with several flags indicating the corrections and the overall data quality, only data with a quality flag of 3, the highest quality are used in this study.

The elliptical orbit of CHAMP (cartoon image shown in Figure 5.2) passed through all magnetic local times in 3-4 months with an orbital period of 94 minutes between successive ascending tracks. The orbital altitude of CHAMP also varies from 460 km to 250 km, resulting in a change in the orbital period throughout the 10 years used in this study. This change varies the time between successive ascending tracks slightly as well as the velocity of satellite. With faster orbital velocities, each sample of the magnetic field will represent the average characteristics of a larger spatial area during that sample time.

The magnetic local time of each satellite pass with high quality data is shown in Figure 5.3, measurements taken at magnetic local times between 1800h-0600h are shown in red. There are times when both the ascending and descending orbital paths only pass over the sunward facing half of the Earth.
5.3 Processing Satellite Data

CHAMP flux-gate vector magnetometer data are provided by the Information Systems and Data Center at GFZ, Potsdam in a North-East-Center coordinate system. This system is rotated into spherical geocentric coordinates, $B_r$, $B_\theta$ and $B_\phi$ using the following transformations:

\begin{align*}
B_r &= -B_z \
B_\theta &= -B_x \
B_\phi &= B_y
\end{align*}

(5.1)
(5.2)
(5.3)

5.3.1 Choice of Main Field model

To isolate the external and internal magnetic field contributions from the ring current, a model of the main geomagnetic field contributions must be removed from the data. The CHAOS-4 model (Olsen et al., 2010), which spans 1997-2011 is chosen. This model is based on data from several magnetic satellites and only utilizes data from magnetically quiet days, to minimize the effects of induced internal magnetic fields. CHAOS-4 includes a lithospheric magnetic field model up to spherical harmonic degree 100 and is composed of time-varying spherical harmonics up to degree 20 to model the variations in the core field. The time variations are modeled using $\beta$ splines with 6 month knot spacing. With this knot spacing, estimates of $e^0_i(t)$ and $i^0_i(t)$ will have less power at periods of 6 months and above and will likely produce better estimates of the power at shorter periods. The changes in this model over the 10 years of the CHAMP data used in this study are shown in Figure 5.4. Over this 10 year period, relative changes in the intensity of the magnetic field of up to 3% are seen in the Caribbean and south east and west of Africa.

Various tests were conducted to determine the appropriate order of the spherical harmonic model of the static internal fields. The residuals after fitting to $B_r$ and $B_\theta$ for each satellite pass did not change significantly when the maximum spherical harmonic degree varied between 65-100. The empirical distribution of the least-squares fits to
Figure 5.4: Changes in the magnetic field model of CHAOS-4 between 2010 and 2001. Top panel: Percent relative change in intensity, $F$ at the surface of the Earth. Bottom panel: Relative difference (from the 2001 model) in the radial field at the surface of the Earth (in $\mu T$).
**Figure 5.5**: The empirical smoothed probability distributions of values of $e_1^0(t)$ from 2001 to 2009 when models of maximum spherical harmonic degrees of 50 (solid blue), 65 (broken red), 85 (dotted green) and 100 (dashed green) are used to remove the main field.

**Figure 5.6**: The empirical smoothed probability distributions of values of $i_1^0(t)$ from 2001 to 2009 when models of maximum spherical harmonic degrees of 50 (solid blue), 65 (broken red), 85 (dotted green) and 100 (dashed green).
Figure 5.7: Test of the rotation into magnetic dipole coordinates on a model of DGRF 2010 from $g_1^0, g_1^1, h_1^1$. The model coefficients for the dipole coordinate system are shown in green with geocentric coordinates shown in blue $e_0^0(t)$, shown in Figure 5.5 shifted to a lower mean when degree 85 or 100 models were used. These higher order models resulted in a broadening of the distribution of fits to $e_0^0(t)$, shown in Figure 5.6. From these tests, a maximum degree of 100 was chosen.

After a main field model is removed from the data, they are rotated into magnetic dipole coordinates corresponding to the 2005 DGRF (Definitive Geomorphic Reference Field) using coordinate transformations from Hapgood (1992). This transformation was tested using a simple dipolar magnetic field model; these results are shown in Figure 5.7.

The data are decimated by a factor of two using a low pass filter in equation 5.4 devised by Robert L. Parker. In this equation, $F_{\theta j}$ are the filter coefficients, where $\theta(j) = \pi(0.5N_{fil} + 0.5 - j)$. If a series is being decimated by a factor of $N_{dec}$, then $N_{fil} = 10N_{dec} + 1$. Here $N_{fil}$, the number of coefficients is 5 and $N_{dec}$ is 2.
\[
F_\theta = \frac{\sin(\frac{\phi}{N_{dec}}) \cos(\frac{\phi}{N_{fil}+1})}{\theta}
\]  

(5.4)

### 5.4 Solving for \(e_1^0(t)\) and \(i_1^0(t)\)

Once a model of the core and lithosphere magnetic field is removed then, in geomagnetic dipole coordinates, \(V\) can be written as

\[
V(r, \theta, \phi) = a \{e_1^0(t) \frac{r}{a} + i_1^0(t) \left(\frac{a}{r}\right)^2\} P_1^0(\cos\theta)
\]  

(5.5)

for a point at a distance of \(r\) from the center of the Earth with co-latitude \(\theta\) and longitude \(\phi\), where \(e_1^0\) indicates the amplitude of the external dipole, \(i_1^0\) is the amplitude of the induced internal response and \(a\) is the radius of the Earth.

Using multiple measurements of the field from each half orbit of the satellite, \(e_1^0\) and \(i_1^0\) are modeled as follows:

\[
B_r = \frac{\partial V}{\partial r} = -\cos(\theta)e_1^0 + 2(\frac{a}{r})^3 \cos(\theta)i_1^0 
\]  

(5.6)

\[
B_\theta = -\frac{1}{r} \frac{\partial V}{\partial \theta} = \sin(\theta)e_1^0 + (\frac{a}{r})^3 \sin(\theta)i_1^0 
\]  

(5.7)

\[
B_\phi = 0
\]  

(5.8)

Accurately estimating \(e_1^0\) and \(i_1^0\) requires high quality data measurements. Data of lower quality should not be used, resulting in occasional breaks in the sampling of the field when the available data does not span at least 40 degrees of latitude. This is to ensure adequate geographic coverage of the field.

Several different least squares solution algorithms were tested. The initial least squares estimate with iterative re-weighting increased the norm of the residuals slightly (0.5%) without changing the results. Often, there were single large outliers or short segments of data where the residuals were linearly correlated. Because of the varied nature of the residuals, a two-step method was used. An initial solution was obtained using least squares fitting. Any data points with residuals larger than 3 standard deviations were discarded—at most 7% of the samples were removed. An iterative least squares
algorithm, using a 2-norm, was applied to this reduced data set to obtain a final solution for each satellite pass.

### 5.5 Data Selection Considerations

At long periods, the phase between the external uniform field and the internal dipolar response is very close to zero and the relationship between the external and internal fields is approximately linear. Examining this relationship for each satellite pass should provide a measure how linear this relationship truly is and provide a rough estimate of the strength of the inductive internal response. The dynamics and linearity of this system vary depending upon the magnetic local time of the satellite as the fields involved vary with magnetic local time. These effects are studied using scatter plots of $e_1^0(t)$ vs $i_1^0(t)$ for each magnetic local time. Robust least squares fitting was used to estimate the strength of this relationship as a function of local time and an estimate of the linearity between the two series. Scatter plots of $e_1^0(t)$ versus $i_1^0(t)$ are shown in Figures 5.8, 5.9, 5.10 along with the equation of the best fitting line and the standard deviation, $\sigma$, these quantities are also listed in Table 5.2. There are not sufficient data from all local times to estimate a fit. Few data from 8h-9h and from 14h-15h meet the strict quality criteria.

From the values in Table 5.2, the relationship between magnetic local time and

<table>
<thead>
<tr>
<th>Local Time</th>
<th>Slope</th>
<th>Y-intercept</th>
<th>$\sigma$</th>
<th>Local Time</th>
<th>Slope</th>
<th>Y-intercept</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.170</td>
<td>-4.24</td>
<td>2.38</td>
<td>13</td>
<td>0.259</td>
<td>1.30</td>
<td>5.78</td>
</tr>
<tr>
<td>2</td>
<td>0.171</td>
<td>-4.38</td>
<td>2.45</td>
<td>16</td>
<td>0.254</td>
<td>-6.59</td>
<td>5.85</td>
</tr>
<tr>
<td>3</td>
<td>0.160</td>
<td>-4.15</td>
<td>2.49</td>
<td>17</td>
<td>0.253</td>
<td>-5.23</td>
<td>5.63</td>
</tr>
<tr>
<td>4</td>
<td>0.154</td>
<td>-4.30</td>
<td>2.63</td>
<td>18</td>
<td>0.247</td>
<td>-4.15</td>
<td>3.32</td>
</tr>
<tr>
<td>5</td>
<td>0.161</td>
<td>-4.34</td>
<td>2.59</td>
<td>19</td>
<td>0.210</td>
<td>-3.74</td>
<td>3.09</td>
</tr>
<tr>
<td>6</td>
<td>0.173</td>
<td>-4.47</td>
<td>2.81</td>
<td>20</td>
<td>0.205</td>
<td>-4.08</td>
<td>2.93</td>
</tr>
<tr>
<td>7</td>
<td>0.205</td>
<td>-5.56</td>
<td>3.12</td>
<td>21</td>
<td>0.180</td>
<td>-3.72</td>
<td>2.61</td>
</tr>
<tr>
<td>10</td>
<td>0.287</td>
<td>-0.41</td>
<td>5.70</td>
<td>22</td>
<td>0.172</td>
<td>-3.97</td>
<td>2.62</td>
</tr>
<tr>
<td>11</td>
<td>0.265</td>
<td>3.37</td>
<td>5.78</td>
<td>23</td>
<td>0.163</td>
<td>-3.92</td>
<td>2.55</td>
</tr>
<tr>
<td>12</td>
<td>0.314</td>
<td>3.93</td>
<td>5.85</td>
<td>24</td>
<td>0.159</td>
<td>-3.94</td>
<td>2.53</td>
</tr>
</tbody>
</table>
Figure 5.8: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of 18h to 23h. The equation and standard deviation of robust least squares line fits are shown on each plot.
Figure 5.9: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of 24h to 5h. The equation and standard deviation of robust least squares line fits are shown on each plot.
Figure 5.10: Scatter plots of $e_1^0$ vs $i_1^0$ with local times of, from top left to bottom right: 6h, 7h, 10h, 11h, 16h, 17h. The equation and standard deviation of robust least squares line fits are shown on each plot.
the linearity of $e_0^1(t)$ vs $i_0^1(t)$ is inferred. There is a distinct change in the relationship between $e_0^1(t)$ and $i_0^1(t)$ between night-side and day-side local times. Between 7h-17h, the standard deviation of the line fits increases sharply. The slope of the line relating these quantities changes more slowly, with larger values in the early evening, between 18h-20h. The y-intercept is -4nT on average, during the evening and night hours (from 18h-6h), this value increases in the early morning. This quantity could represent the static contribution of the internal field recorded by the satellite. If we assume this signal is purely inductive, it corresponds to a static external field strength of around -20 nT.

Examining the data this way also provides insight into the influence of various data selection criteria. The estimates of $e_0^1(t)$ and $i_0^1(t)$ plotted above are fit to data within ±45° of latitude are used to estimate $e_0^1(t)$ and $i_0^1(t)$. If a smaller range, from ±35° of latitude is used, the standard deviation of the linear fits increases by 7%. If data within ±55° of colatitude are used, the linear offset increases by 5%.

Estimates of the robust equation of the line relating $i$ and $e$ as a function of local time were computed for 2002 using maximum spherical harmonic degrees of 50, 65, 85 and 100 for the main field model. The strongest linear relationship occurs with 65 degrees of internal field, the strength of night-side (18h-6h) slope of is 4% higher than the linear relationship when the maximum spherical harmonic degree is 100.

The non-linear nature of system does not necessarily imply that a spherical harmonic degree of 65 is the optimal choice for the maximum degree of the model, merely the most linear. As we do not expect this system to be exactly linear, the final data set used a maximum spherical harmonic degree of 100 with the hypothesis that removing as much static, non-uniform or non-dipolar structure as possible would yield a better estimate of the uniform, $e_0^1$ and dipolar $i_0^1$ fields.

### 5.6 Estimation the response function, $G_{ie}$

Assuming that the external field is uniform at satellite altitude, and that it induces an internal $P_1^0(t)$ field structure, ignoring the asymmetries for the moment, then:

$$Q(f) = \frac{P_1^0(f)}{E_1^0(f)}$$

(5.9)
Figure 5.11: Time series of data segments used for pilot estimate of the transfer function. Slightly different data segments were tested. The first test included the large magnetic storm in late October 2003, shown in green. A shorter data segment which omitted the storm is shown in blue. Top panel: $e_0^0(t)$. Bottom panel: $i_1^0(t)$.

This quantity represents the transfer function of a system, the Earth driven by the $e_0^0(t)$ and inducing the output, $i_1^0(t)$. It practice this is estimated as the frequency domain transfer function, $G_{ie}(f)$:

$$G_{ie} = \frac{S_{ie}}{S_{ee}}$$

(5.10)

where $S_{ie}$ is the cross-spectral estimate between $i_1^0(t)$ and $S_{ee}$ is the auto-spectrum of $e_0^0(t)$. The multi-taper method for calculating these quantities is described in Section 4.2.

Because of the large breaks in sampling, a pilot estimate of the transfer function, $G_{ie}$, is required to utilize the method from Maus and Weidelt (2004) to interpolate small breaks in sampling. To ensure the response function is consistent with the data, it must be estimated from a long continuous segment of data. The data segments used to estimate this pilot response function should be as long as possible with few breaks.
in sampling. To obtain a strong inducing signal, the data segments should be chosen from an active part of the 11-year solar cycle. Once $G_{ie}$ has been estimated, then the algorithm described later, in Section 5.7 can be applied.

A little over 1 year of data was chosen beginning in late September 2003. A large magnetic storm occurred in late October 2003, shown in Figure 5.11. The large impulses associated with magnetic storms are a large non-stationary signal which could bias the estimates of the response function. However, these storms also contribute a great deal of power to the ring current and resulting induced currents. To determine if the storm should be used to estimate the pilot response, I compare the results from using a long time series, which includes the storm (shown in green in Figure 5.11) and a shorter time series that begins after the October 2003 storm.

In Figure 5.12, the power spectral density estimates from the longer segments ($e_1^0(t)$ is shown in light green and $i_1^0(t)$ is shown in light blue) that include the storm have more power than the shorter data segments. The low frequency portion of the spectrums of both $e_1^0(t)$ and $i_1^0(t)$ are flatter at frequencies below 2 cycles per day, with a slight peak at 2 cycles per day. This peak is not as pronounced in the spectral estimate or confidence intervals of $e_1^0(t)$ (shown in dark green) from the shorter segment with the storm omitted.

The coherence between $i_1^0(t)$ and $e_1^0(t)$ is significantly lower when the magnetic storm is included in the time series (shown in blue in Figure 5.13.) The phase relationship between $i_1^0(t)$ and $e_1^0(t)$ at frequencies above 3.5 cycles/day does not change if the storm is included or omitted- both estimates of the phase lie within the confidence intervals of the estimate from the shorter time series. However, when the storm is included in the data series, $i_1^0(t)$ lags behind $e_1^0(t)$ by roughly 20 degrees at 1 and at 2 cycles per day.

Response function estimates from time series which include this storm (shown in black ) and omit this storm (shown in blue) are shown in Figure 5.14. Including the October 2003 storm (Pulkkinen et al., 2005) yields a broader and lower plateau in the amplitude of the response function at frequencies below 1 cycle per day and near 2 cycles per day. Both amplitude responses lie within the 95% confidence intervals of the estimate from the shorter data segments at frequencies above 3 cycles per day.
Figure 5.12: Spectral estimate of $e_0(t)$ (dark green) and $i_0(t)$ (dark blue) from the data segments shown in blue in Figure 5.11 shown with 95% jackknifed confidence intervals. The spectral estimates from the longer time series from Figure 5.11 is shown in light green. The broken lines indicate the estimated noise floor for each series.
Figure 5.13: Coherence between $i_0^0(t)$ and $e_1^0(t)$ from the time series (shown in blue in Figure 5.11) using 38 minimum bias tapers is shown in black with 95% confidence intervals. The coherence from 50 MB tapers applied to the longer time segment (from Figure 5.11) containing the October 2003 storm is shown in blue. Both estimates have a frequency resolution of 0.2 cycles per day.
Figure 5.14: Response function estimate between $i^0_1(t)$ and $e^0_1(t)$ from the time series (shown in blue in Figure 5.11) using 38 minimum bias tapers is shown in blue with 95% confidence intervals. The response function estimate from 50 MB tapers applied to the longer time segment (from Figure 5.11) containing the October 2003 storm is shown in black. Both estimates have a frequency resolution of 1 cycle per day.
The flattening in the spectral amplitudes and response functions estimated from the data series that include the magnetic storm suggests that the true long-period characteristics of the response function are being masked by the large magnetic storm. The response function estimated from the shorter time-series is likely to be more accurate and is used as the pilot estimate for equation 5.12.

The pilot estimate of $G_{ie}$ is used to partition Dst into internal and external contributions, as described above with the method of Maus and Weidelt (2004).

5.7 Time Sampling

Breaks in satellite time series occur for a variety of reasons. Accurate estimates of the external and internal fluctuations from each satellite pass require a large number of high-quality vector measurements that span a wide range of latitudes near the equator, away from the auroral zones - where the external magnetic field is not dipolar and violates our source assumptions. The study of $e$ vs $i$ from Section 5.4 suggests that $\pm 45$ of latitude provides the best trade-off between sufficient data coverage and minimizing the influence of the auroral zone.

In Figure 5.15, the times between the midpoint of successive satellite passes (with adequate data) from magnetic local times between 18h-6h are plotted on a logarithmic scale. Small irregularities in the sampling rate are created as the satellite precesses through local times of 6h and 18h. Every 3-4 months, as the satellite precesses through all local times, shown in Figure 5.3, an ascending and descending pass both cover local times between 18h-6h. If the ascending pass of the satellite passed through local times of 6h and the descending pass travelled through 18h, then the two samples are only $0.5\delta_t$ apart (where $\delta_t$ is the orbital period of the satellite). For the next 3-4 months of samples, nightside passes are estimated from the descending path and are offset from the previous sample block by $0.5\delta_t$.

To adjust for the shift in sampling times and short breaks in sampling, the algorithm developed in Section 2.5 is applied. First, Akima (Akima, 1970) spline interpolation is used to shift all data blocks to a regular sampling rate. Short breaks in sampling, of less than 2 sample intervals are also interpolated using this method.
Figure 5.15: The time between satellite passes with local times within 18h-6h on a logarithmic scale. The average sampling rate is shown as a red line at 0.0637 days.
Two threshold values are chosen. If a break in sampling is less than $T_1$ samples long, then the data is interpolated using Akima splines. If the break in sampling is greater than $T_1$ samples long but less than the second threshold value, $T_2$, the missing samples are filled using scaled estimates of internal, $i_1^0$, and external, $e_1^0$ contributions from Dst.

Using a given response function estimate, $G_{ie}(f)$ (Maus and Weidelt, 2004) the Dst time series can be separated into internal and external parts. This method was used in Section 4.4 to generate synthetic data. Here, this technique is used to predict values of $i_D(t)$ and $e_D(t)$ from May 2001 until September 2010. These time series are obtained by computing the inverse Fourier transforms of $E_D(f)$ and $I_D(f)$. These Fourier transforms are estimated using:

$$E_D(f) = \frac{1}{1 + G_{ie}(f)} D(f) \quad (5.11)$$

$$I_D(f) = \frac{G_{ie}(f)}{1 + G_{ie}(f)} D(f) \quad (5.12)$$

where $D(f)$ is the Fourier transform of the Dst time series which has been interpolated (using Akima splines) to match the orbital period of CHAMP. The predictions from Dst are scaled to match the local mean and variance of the $i_1^0(t)$ or $e_1^0(t)$ time series. Here ‘local’ is a window of 100 points before and after the break in sampling. The scaled values fill the gaps in the time series record which are less than $T_2$ samples long.

In Figure 5.16, a section of data with 47 samples missing is shown. The akima spline interpolation of the 47 samples missing from $e_1^0(t)$ (shown as dark green stars) is replaced with the time series of scaled predictions of $e_D$ (shown as magenta diamonds). The resulting time series (black) has a slight jump of 4 nT at the end of the interpolation. For the $i_1^0(t)$, (shown as light green stars) replacing the akima spline interpolation (red diamonds) with the scaled predictions results in a much better match at the end of the interpolation (final time series shown in light grey). Slight jumps due to mismatches with the $i_D$ and the original $i_1^0(t)$ time series do not occur often, this example illustrates one of the largest mismatches in the interpolation.
Figure 5.16: Filling break in sampling 3 days long (47 samples) in October 2004 using scaled predictions. The Akima spline interpolation for $e_1^0(t)$ ($i_1^0(t)$) is shown with dark green (light green) stars. The time series of scaled predictions of $e_D$ ($i_D$) is shown with magenta (red) diamonds. The final time series after interpolation are shown with a black, for $e_1^0(t)$ or dark grey, for $i_1^0(t)$ line.
5.8 Estimating the Noise in the Signals

From the discussion of Olsen (1998) (and Section 4.6), the bias of the noise in both the input and the output can be corrected if the relative noise in the input to the output, $\psi$, is known using equation 5.13.

$$G_{opt}(f) = \frac{G_{xy}(f)}{2}(1 - \psi + \sqrt{(1 - \psi)^2 + \frac{4\psi}{\gamma(f)^2}}) \quad (5.13)$$

To produce an accurate estimate of the response function using equation 5.13, the relative amount of noise in $e_1(t)$ and $i_1(t)$ must be estimated. To determine amount of noise in each signal, an estimate of the ratio of noise power to total power, $\eta = |S_N|/|S_x|$, must be made. Here it is assumed that the noise in both $e_1(t)$ and $i_1(t)$ is Gaussian and white, implying the spectrum of the noise is flat. In Figure 5.17, the spectrum is shown with (red line) and without the noise correction (blue).

The amount of noise in each signal is estimated under the assumption that the power in the variations of both $e_1(t)$ and $i_1(t)$ decreases at higher frequencies and that the fall-off rate is proportional to frequency. When the power spectral density of $e_1(t)$ or $i_1(t)$ is constant across some frequency band it is due to noise in the signal. In Figure 5.12, the power spectral density of $e_1(t)$ flattens out to a power level $2.6nT^2$ day above 6 cycles/day ; for $i_1(t)$, the spectra flattens out at $0.65nT^2$. With these assumptions, the ratio of noise power to total power is 0.2 for $e_1(t)$ and 0.076 for $i_1(t)$. This yields a $\psi$ value of 0.38, indicating that there is far more relative noise in the input, $e_1(t)$ than the output, $i_1(t)$.

5.9 Applying Minimum Bias Tapers with Gaps

Using an indicator function to solve the optimization problem which defines the minimum bias tapers, described in Section 2.2, typically provides a set of tapering functions with a similar frequency domain response to the standard minimum bias tapers but with a wider averaging bandwidth-yielding less frequency resolution and a less out-of-band signal rejection-providing less protection against spectral leakage.

Interpolating all breaks in sampling less than 100 sample intervals, or a little
Figure 5.17: Response function estimate of the November 2003 - October 2004 data (same as the black line in Figure 5.14) from 42 minimum bias tapers is adjusted using equation 5.13 and shown in red. This adjustment lies entirely within the confidence intervals of the original response function.
The empirical average transfer function of minimum bias tapers with gaps (MBG) computed for the indicator function in Figure 5.18 is shown in Figure 5.19. For this indicator function, the MBG tapers have nearly the same frequency resolution as the standard minimum bias tapers but less protection against spectral leakage.

Using 50 of the tapers shown in Figure 5.19, the response function between the internal and external fields was estimated. The relative noise level, $\psi$ was estimated to be 0.38 and used to correct the MBG response. Both results are shown in Figure 5.20. The noise correction generally lowers the amplitude of the response function, however, it also increases the amplitude of the signal at around 1 cycle per day. This noise correction also assumes the relative noise power between the two series is a constant for all frequencies. We also estimate the response using section averaging - which may be less sensitive to noise.
Figure 5.19: Average empirical transfer function of the minimum bias tapers with gaps (solid lines) when a threshold value, $T_2$ of 100 is used. These are compared to the average transfer function from a set of standard minimum bias tapers (broken lines).
Figure 5.20: Response function with noise correction, comparing MBG tapers with and without noise correction (shown in blue without the noise correction with 95% confidence intervals in light blue). Noise correction (red line) is only in amplitude.
5.10 Comparing Response Functions from Section Averaging and MBG Tapers

Response functions for the three long segments of data in this indicator function could also be estimated using section averaging. With the indicator function shown in Figure 5.18 and a threshold value of 100 samples, the response function estimate from 50 minimum bias tapers with gaps is compared to section averaging using minimum bias tapers, as described in Chapter 2. Twenty tapers are applied to the first segment of data, 18 on the second and 3 on the last and shortest segment. The equivalent, optimal average transfer function from section averaging is shown in Figure 5.21. For sets of 20 tapers or less, the minimum bias tapers with gaps can obtain a narrower averaging bandwidth than section averaging.

The interpolation of the shorter breaks in sampling is likely to influence the higher frequency signals. For frequencies above 1 cpd, the variance in the transfer function estimate is quite high. We are interested in the low frequency response and restrict the analysis to frequencies of 1 cpd or lower. The response function from section averaging had higher amplitudes, shown in Figure 5.22. The phase response is almost identical - the only differences occur at very high and low frequencies. Despite the better noise rejection (due to the steeper fall-off) and therefore reduced sensitivity to noise when using section averaging with standard MB tapers, the confidence intervals (not shown) are twice as wide, on average, as the MBG tapers. The reason for this is not clear, a similar result was seen in Chapter 2. The steeper drop off around 0.01 cycles per day of the MBG tapers (shown in Figure 5.21) may play a role.

5.11 Estimates of the C-response

Using the work of Weidelt (1972), the response function can be transformed into C-response for the flat-Earth problem and written as:

\[ C(f) = \frac{a}{2} \frac{1 - 2Q_1(f)}{1 + Q_1(f)} \]  

(5.14)
Figure 5.21: Average empirical transfer function of the standard minimum bias tapers used for section averaging are shown in solid lines. These are compared to the average transfer function from the equivalent set of MBG tapers (broken lines).
Figure 5.22: Response function estimate from MBG tapers for K=50 (shown in black with grey confidence intervals) and from optimized section averaging with K=51 (shown in blue).
Figure 5.23: C-response estimate from 51 MB tapers, the real portion is shown as a sold blue line and the imaginary portion is shown with a broken blue line. The noise correction is applied to this and shown in red. A section averaging approach using 15 MB tapers is shown in green. Previous results from Constable and Constable (2004) are shown as blue diamonds. Kuvshinov and Olsen (2006) shown with green asterisks. The broken black line represents 400 days and is the highest resolution of the MBG estimate.
Figure 5.24: (Top): Mean-squared coherence (MSC) shown between $i_1^0(t)$ and $e_1^0(t)$ plotted in blue for the 51 MBG tapers. The MSC estimated from 50 MB tapers using section averaging is shown in green. (Bottom): Jackknifed standard error estimated from the MBG estimate (blue) and MB section averaging (green).

where $Q(f) = G_{ie}$, the response function. For a large number of tapers, this can be approximated as following a gaussian distribution. Therefore, the variance can be reasonably accurately estimated using jackknife resampling. The C-response for $l = 1$ is close to a direct scaling of $Q(f)$. The variance of C is estimated from a set of $K$ jackknifed estimates of $Q$.

$$\sigma_C^2 = \frac{K-1}{K} \sum_{m=1}^{K} (C_{(m)} - C_{(\cdot)})^2$$  \hspace{1cm} (5.15)$$

where $C_{(m)}$ is the estimate of $C$ with the $m^{th}$ taper omitted from the estimate of $Q$ and $C_{(\cdot)}$ is the mean of the jackknifed estimates of $C$. 

5.12 Comparisons with Other Methods and Previous Results

C-response from two approaches (solid lines for real part and dashed lines for the imaginary part) is shown in Figure 5.23 in blue out to periods of 10 days. Periods below shorter than 10 days are not shown because those periods are more likely to be influenced by ocean induction effects, which are modeled as part of this work. After correcting for a relative noise power, $\psi = 0.38$ (shown in red) the estimate ends up quite close to that from section averaging (shown in dark green). Both the MBG and section averaging approach show a drop in the imaginary portion of the C-response around 150 days. At periods longer than 300 days, the C-response plateaus 2500 km, for the real portion and close to zero for the imaginary part of the response. The mean-squared coherence between the series is shown in the top panel of Figure 5.24; the jackknifed standard error is shown in the bottom panel. The error in the C-response estimate varies as a function of frequency. Some portions of the signal are better resolved than others.

However, at frequencies higher than 0.1 cycles per day this C-response estimate is substantially biased. At these frequencies the heterogeneity of the surface of the Earth, especially the high conductivity of the oceans, represent a substantial 3-D effect. To accurately estimate these periods these surface features must be modeled. Kuvshinov and Olsen (2006) modeled the induced magnetic field contributions from the ocean water and surface sediments and removed the predicted fields from the satellite measurements before estimating the dipole coefficients. The ocean correction largely affects higher frequency (shorter period signals), so we omit estimates at frequencies higher than 1cpd.

The estimate from Figure 5.23 is within one standard deviation of the results from (Kuvshinov and Olsen, 2006) for the imaginary part of the response (shown in green on Figure 5.23). The real portion of the estimate from the 51 MBG tapers, shown in blue, is slightly higher than the Kuvshinov and Olsen (2006) results. However, this estimate is of much higher resolution than the results of Kuvshinov and Olsen (2006). Because of the higher coherence of the section averaging estimate (shown in Figure 5.24), the estimate from section averaging (shown in green in Figure 5.23) is likely to be the most accurate.
Figure 5.25: C-response estimate from 15 MB tapers (applied using section averaging) shown in red (after the noise correction is applied) along with 95% confidence intervals for the uncorrected estimate are shown in grey. These are compared to (Olsen, 1999) (shown as purple stars) with (Constable and Constable, 2004) shown as blue diamonds and (Kuvshinov and Olsen, 2006) shown with green asterisks.
Both Kuvshinov and Olsen (2006) and the method described here relied on spline functions to regularize the time sampling. Kuvshinov and Olsen (2006) used splines with 4-hour knot spacing to combine measurements from the Ørsted, CHAMP and SAC-C satellites. Interestingly enough, the trend with the real component follows quite closely with Constable and Constable (2004), who also used a multi-taper approach.

The section averaging method in Figure 5.25 shows the noise-corrected estimate from 15 MB tapers, with a resolution of 0.0012 cpd. This is shown with 95% confidence intervals shown in grey. Within uncertainty, these results match those of Constable and Constable (2004) and Kuvshinov and Olsen (2006). This estimate provides robust determination of the C-response out to at least 864 days which is substantially longer than 100 day cutoff for previous methods using satellite data sets. Our results are a bit higher than those of Olsen (1999), which went out to roughly 4000 days using European magnetic observatory data. The estimate of Olsen (1999) could be biased to lower values of the C-response because it was only representative of the 1-D response above Europe. At these long periods, it is unlikely that there is a significant bias in our results due to the oceans. Our result is likely to represent a more widespread 1-D view than those of Olsen (1999).

5.13 Discussion and Conclusions

Nearly 10 years of magnetic satellite measurements from CHAMP have been used to estimate the electromagnetic response of the Earth. Strict quality criteria were imposed on the vector magnetic field measurements used in the modeling. From these vector measurements, an estimate of the external dipolar magnetic field variations and the corresponding internal dipolar field were modeled. Only measurements from ±45 of latitude were used in the fitting. Using data from only ±35 of latitude resulted in estimates of $e_0^0(t)$ and $i_1^0(t)$ which were less linearly correlated, as a function of magnetic local time.

The time between successive estimates of $e_0^0(t)$ and $i_1^0(t)$ spanned several orders of magnitude, with the longest gaps spanning more than 20 days. Applying the algorithm from Chapter 2, all breaks in sampling less than 100 samples long (corresponding
to 6 days), were interpolated using predictions from Dst and a pilot estimate of the response function $Q(f)$, yielding only two large breaks in sampling. In this case section averaging produced estimates with higher coherence between $e_0^i(t)$ and $i_0^i(t)$ but the phase response was identical. The response function from standard MB tapers is less sensitive to noise, as discussed in Chapter 4.

Using these methods, the internal field, $i_0^i(t)$ was found to have 0.38 of the relative noise power of the external field, $e_0^i(t)$. This estimate of the relative noise power in the output to the relative noise power in the input, $\psi = 0.38$ was used to correct the response function estimate for bias. However, this correction did not significantly alter the response function estimate.

These results could be improved if a static model of the main internal field were used, rather than the 6 month knot spacing. Using a maximum spherical harmonic degree of 65 for this model should also slightly improve the modeling of $e_0^i(t)$ and $i_0^i(t)$, as suggested by the linear correlation tests performed. This could suggest that lithospheric field model used, estimated using CHAMP data, is not well constrained at very high degrees.

The linear fits to the scatter plots of $e_0^i(t)$ versus $i_0^i(t)$ clearly show the influence of the Sq currents. A slight asymmetry in the field is revealed in late evening, with a static 4 nT field. This is unlikely to be an artifact of the high degree spherical harmonic field model used. The energy contribution of the lithospheric field model from CHAOS-4 from spherical harmonic degrees 65-100 is less than 1 nT at satellite altitude. This evening asymmetry is a feature of the magnetosphere.

Including the October 2003, Hallowe’en storm (Pulkkinen et al., 2005), in the pilot estimate greatly altered the long-period coherence and phase characteristics of the system. Removing strong magnetic storm signals from the time series may provide a more coherent long period signal and, more accurate estimates of the response of the lowermost mantle.

Signal coherence continues to be a problem when studying the long period variations in the magnetosphere, (Khan et al., 2011) restricting the recovery of periods of longer than 100 days. Using the adapted section averaging approach with minimum bias tapers, a maximum frequency resolution of 0.01 cycles/day (slightly less than the half
bandwidth shown in Figure 5.21) is obtainable.

From the C-response calculated in this work, we are able to estimate the electromagnetic response out to periods of 180 days. This estimate could be inverted using a scheme similar to (Medin et al., 2007) to provide further constraints on the electrical conductivity of the deep mantle.
Bibliography


Chapter 6

Conclusions

The goal of this thesis has been to develop power spectral and cross-spectral tools for analyzing time series with long breaks in sampling. The methods of Fodor and Stark (2000) for adapting multi-tapers to compensate for gaps in sampling using an indicator function, were extended for cross-spectral estimation. As part of this work a software package and algorithm were developed and applied to various geophysical problems. The minimum bias tapers, adapted to compensate for breaks in sampling, provide nearly the same tapers and spectral properties as an optimized section averaging scheme.

These tools were applied to study the long-term behavior of the geomagnetic field during the Oligocene from a sedimentary record with long interruptions in sampling. This relative paleointensity time series was calibrated using the methods of Ziegler et al. (2011). The new analysis reveals a somewhat abrupt change in field behaviour around 29 Ma. The earlier portion of the time series, studied by Constable et al. (1998) included a quasi-periodic signal at 8 cycles per million years and corresponded to a period with strong secular variation but relatively few magnetic reversals. The younger data, from the late Oligocene, contained a quasi-periodic signal at 2.5 cycles per million years and had less power in secular variation. These results appear to imply that a strong magnetic field inhibits reversals while periods with more magnetic field reversals have less powerful secular variation. This could indicate that magnetic field reversals and secular variations are caused by similar non-linear processes.

The minimum bias tapers, adapted for breaks in sampling are applied to trans-
fer function and coherence estimation. A method for estimating confidence intervals for coherence spectra (from Thomson and Chave (1991)) was described. A method of estimating the confidence of a transfer function estimate was also developed. The behaviour of the MBG tapers in the presence of noise was tested using a synthetic time series of the external dipolar fluctuations associated with variations in the ring-current and the resulting induced signals. The bias in the response function estimate due to noise and uncertainty in the fitting of the external and internal coefficients was studied using the work of Olsen (1998). This requires an estimate of the ratio of the relative noise in the output and the relative noise in the input, which was found to be 0.38 - indicating far more noise in the estimate of \( e_1^0(t) \). With this ratio and an estimate of the coherence, much of the bias from the noise can be corrected but not all. The MBG tapers provide reliable estimates of the coherence and transfer function from intermittently sampled time series. However, when the continuous data segments of the time series are of roughly equal length, section averaging provides transfer function estimates which are less sensitive to noise than the MBG tapers.

The techniques developed for transfer function estimation were used for geomagnetic depth sounding with satellite data. From vector measurements of the magnetic field at mid-latitudes, estimates of the external dipolar fluctuations and the internal dipolar response were modeled from high-quality data sets. Care must be taken when deciding which time variations of the internal field model to include. Using CHAOS-4 with 6 month knot spacing may have removed much of the signal of interest at long periods. The C-response estimated here could be slightly improved if the maximum degree of the spherical harmonic model was 65, rather than 100. Longer periods would be available if the internal field model used had more widely-spaced spline knots or a static average field was removed.

Despite this, the C-responses estimated from the interpolated data and MBG tapers included longer periods than those estimated from previous satellite studies. The empirical confidence intervals propagated through to the C-response yield more realistic estimates of the associated uncertainties. Ultimately, the best estimate of the C-response at long periods (with frequency resolution 0.0012 cpd) came from 15 minimum bias tapers which were applied using section averaging. The MBG tapers had narrower con-
fidence intervals but resulted in lower coherency between $i_0(t)$ and $e_0(t)$, particularly at higher frequencies (shorter periods). These techniques allowed us to extend the estimate of the C-response out to 180 days. At this period, the C-response was higher than the results of Olsen (1999), estimated from European observatory data.

6.1 Future Work

With faster eigenvalue solvers, the prolate tapers with gaps could be more easily applied to longer time series with interruptions in sampling; the cost of recalculating the tapers with gaps with a different frequency resolution would be less. The unique ability of the prolate tapers to minimize broadband bias could prove useful for analyzing multivariate problems to improve long period recovery. Adaptive weighting schemes for tapers have not been implemented here, so only a worst case scenario has been studied. Utilizing these schemes with the tapers adapted for gaps should provide spectral estimates with optimal resolution and uncertainty. As more paleomagnetic data becomes available, applying these methods increases the amount of frequency-domain information available and adds to our knowledge of the variations in the Earth’s magnetic field.

Inverting the C-response derived from the spectral techniques developed in this thesis should extend the depth to which the conductivity of the mantle could be estimated. The empirical estimates of the associated uncertainties should provide more realistic error estimates for these inversions for electrical conductivity. From these models of electrical conductivity, many mantle properties can be better constrained. The electrical conductivity of olivine varies with water content; models of electrical conductivity from C-response estimates can constrain these quantities (Medin et al., 2007). In addition to the water content, the thermal gradient and chemical composition of a region changes the electrical conductivity; Khan and Shankland (2012) used laboratory profiles and regional C-response estimates to study regional properties.

The tools developed in this thesis have broad applications in the physical sciences. For unique time series with long breaks in sampling, these methods of power spectral and transfer function estimation with empirical confidence intervals can extend the amount of information to be gleaned about a physical system by studying its fre-
quency domain behavior.
Bibliography


