SPECTRAL ANALYSIS OF TIME SERIES WITH LATENT AND IRREGULAR TIMES

by

AARON SPRINGFORD

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
December 2017

Copyright © Aaron Springford, 2017
Abstract

Many standard methods of time series analysis assume that observation times are both known and regularly-spaced. Regular sampling and known observation times are cornerstones of methods such as autoregressive/moving-average models and spectral estimation using the Fast Fourier Transform. When the measurement process is controlled by the experimenter, these assumptions can be largely met by design. However, there are cases in which the measurement process is not under complete experimental control, and the observation times are either irregular or unknown. For example, many data sets in astronomy have irregular sampling due to the effects of orbital geometry and interfering processes such as celestial bodies or atmospheric processes. In paleoclimate studies, time series data may consist of core samples of known depth, but unknown age. Extending common time series analysis methods to these types of data is a challenge.

This thesis makes three key contributions. The first is a new Bayesian method for inferring the chronology – or age versus depth relationship – of a core taken from a sedimentary record. The second is an approximate multitaper statistic (mtLS) for irregularly sampled time series. The third is a Bayesian model for the spectrum inspired by the previous work of Mann and Lees (1996) and Thomson et al. (2001), who separate the spectrum into noise and signal components. Together, the three
contributions are used for spectral inference of time series obtained from one peat core and three lake cores. The approach quantifies and includes uncertainty from both the chronology and the time series process. In addition, the application of the mtLS statistic as an estimator for the spectrum of an irregularly sampled time series in astronomy is presented.
Co-authorship

This work described herein is of my own conception and execution. Portions of Chapter 2 have been published with Dr. David Thomson as a co-author in the Joint Statistical Meetings Proceedings (Springford and Thomson, 2015). In this case, Dr. Thomson supported and was involved in discussions about the work but did not otherwise contribute. His involvement in the remainder of the thesis was similar.
Acknowledgments

Thank you to my supervisors Dave Thomson and Glen Takahara, and to all of the students at Queen’s that I’ve had the pleasure of interacting with during my time here. In particular, thank you Dave Riegert, Karim Rahim, Charlotte Haley, Carly Rozins, Wesley Burr, Josh Pohlkamp-Hartt, Frank Marshall, and Claire Boteler for giving me an education that was unique in many ways. Thank you to my family for getting me to this point. Thank you Gwen – you mean more to me than anything.

Thank you to the Paleoecological Environmental Assessment and Research Laboratory (PEARL) at Queen’s University for providing the prairie lakes data, and Tim Bedding for providing the Kepler flux data. This work was funded in part by a scholarship from the Natural Sciences and Engineering Research Council of Canada.
Contents

Abstract i
Co-authorship iii
Acknowledgments iv
Contents v
List of Tables vii
List of Figures viii
Lists of abbreviations and symbols xxii

Chapter 1: General Introduction 1
  1.1 Spectral analysis of time series . . . . . . . . . . . . . . . . . . . . . . 1
  1.2 Bayesian data analysis . . . . . . . . . . . . . . . . . . . . . . . . . 4
  1.3 Paleoclimate and sedimentary records . . . . . . . . . . . . . . . . . . 5
  1.4 Organization of thesis . . . . . . . . . . . . . . . . . . . . . . . . . . 7

Chapter 2: A Bayesian Hierarchical Chronology Model for Time Series Analysis of Paleoenvironmental Data 11
  2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.1.1 Modelling chronologies . . . . . . . . . . . . . . . . . . . . . . . 14
  2.2 Motivating datasets and methods . . . . . . . . . . . . . . . . . . . . 17
    2.2.1 Bayesian hierarchical chronology model . . . . . . . . . . . . . 19
    2.2.2 Sampling chronology model posterior parameter distributions 25
  2.3 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
  2.4 Comparison to other chronology methods . . . . . . . . . . . . . . . . 30
    2.4.1 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
  2.5 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
  2.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55
List of Tables

1 List of symbols used throughout this thesis. . . . . . . . . . . . xxii
2 List of abbreviations used throughout this thesis. . . . . . . . . . xxvi
List of Figures

1.1 Photos of two sediment cores extracted from Owkeno Lake in British Columbia, August 2004. The left photo shows the frame used to level the core against the bottom. Various methods are used to drive the coring mechanism into the sediment and retain the sediment for retrieval. The core is extruded and sectioned into samples as a function of depth once retrieved. The author was involved in collection of these core samples. ................................................................. 6

1.2 The distribution of radiocarbon ages as a function of true calendar ages for the Holocene. Each panel shows the mean radiocarbon age ± two standard deviations for a given calendar age, as estimated by Reimer et al. (2013) using the methodology of Niu et al. (2013). The relationship between radiocarbon age and calendar age is not monotonic, nor does it follow a 1:1 relationship on average. A given radiocarbon age can imply a complicated and possibly multimodal distribution of calendar ages. ................................................................. 8
2.1 Salinity index data for the three lakes described in Michels et al. (2007). Note that the cores are not all sampled to the same sediment depth. The salinity index is derived from counts of diatom fossils within each layer of the core. The salinity index is available at a resolution of 2 cm for Chauvin and Oro lakes, and 1 cm for Humboldt Lake. Open circles represent the depths of radiocarbon dating samples used to infer chronologies.

2.2 Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}_{w}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.

2.3 One thousand posterior time series of equal probability (light grey lines, over-plotted) and the mean posterior time series (white line). The top panel are the time series for the Greyscale measurements, the middle panel Humification measurements, and the bottom panel Ash Content.
2.4 Chauvin Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.

2.5 Humboldt Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.

2.6 Oro Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.

2.7 The varve-derived (assumed true) age versus depth relationships for the three cores used to compare chronology methods.
2.8 Elk Lake, 41 simulated dated intervals: Posterior distributions for the mean age increment \( \mu \), the age increment standard error \( \tau^{-1}_w \), and the age increments \( w \) (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated \(^{14}\text{C}\) dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).

2.9 Elk Lake, 21 simulated dated intervals: Posterior distributions for the mean age increment \( \mu \), the age increment standard error \( \tau^{-1}_w \), and the age increments \( w \) (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated \(^{14}\text{C}\) dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
2.10 Elk Lake, six simulated dated intervals: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS). . . . . 41

2.11 Holzmaar, 41 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS). . . . . 42
2.12 Holzmaar, 21 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), $2/3$ credible region coverage, and the continuous ranked probability score (CRPS).

2.13 Holzmaar, six simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), $2/3$ credible region coverage, and the continuous ranked probability score (CRPS).
2.14 Iceberg Lake, 41 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).

2.15 Iceberg Lake, 21 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
2.16 Iceberg Lake, six simulated dated intervals. Posterior distributions for
the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and
the age increments $w$ (left to right, top panels). Posterior age versus
depth trajectories based on 1000 posterior samples (bottom panel).
Also plotted are violin plots (in orange) of the likelihood from cali-
brated $^{14}$C dating, and the true age versus depth relationship in blue,
obtained from varve information. The reported measures of perfor-
manence are the root mean square error of prediction (RMSEP), 2/3
credible region coverage, and the continuous ranked probability score
(CRPS).

2.17 Posterior age versus depth trajectories based on 1000 posterior samples
(black), plotted as residuals from a cubic polynomial fit to the mean
likelihoods of the calibrated $^{14}$C dates. Also plotted are the violin
plots (orange) of the likelihoods from calibrated $^{14}$C dating, a dashed
zero-line (blue) representing the cubic polynomial fit, and a piecewise
linear function connecting the means of the calibrated likelihoods. The
Bayesian hierarchical trajectories can be seen as a compromise between
the cubic polynomial fit and the piecewise linear function.

3.1 A simple age model to demonstrate the effect of interpolating an irreg-
ular series to a regular mesh before estimating the spectrum.

3.2 A comparison of multitaper spectrum estimates for a white noise pro-
cess sampled at a time interval of five (top panel), and the same real-
ization sampled irregularly and interpolated to the same time interval
of five (bottom panel), showing the effect of interpolation.
3.3 Application of the Mann and Lees (1996) method to the interpolated white noise series spectrum seen in Figure 3.2. The light gray is the locally smoothed spectrum estimate used to fit the autoregressive spectrum in solid black. The dotted lines are 90, 95, and 99th percentiles of the background spectrum. Estimated spectrum values exceeding these percentiles would be interpreted as significant signals. 66

3.4 The multitaper spectrum estimate for a white noise process sampled irregularly. The spectrum estimate is comparable to the estimate that would have been obtained under regular sampling (Figure 3.2, top). 73

3.5 Three mtLS pseudowindows for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year. The parameters of the underlying DPSS were $N = 649$, $W = 0.0046225$, and $k = 5$. 75

3.6 Three Lomb-Scargle (non-tapered) pseudowindows for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year. 76
3.7 A comparison of the pseudowindows for the mtLS (in blue) and non-tapered Lomb-Scargle (in gray) for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year. The parameters of the underlying DPSS were $N = 649$, $W = 0.0046225$, and $k = 5$.  \\

3.8 mtLS spectral statistics for the Tibetan peat core data: Greyscale (top panel), Humification (middle panel), and Ash Content (bottom panel). The frequencies are jittered, and are in units of cycles per year. Each point represents chronology drawn from the posterior distribution of chronologies (Figure 2.2).  \\

3.9 Relative flux data for the red giant star Kepler-91.  \\

3.10 Distribution of first time differences $\Delta t$ for the red giant flux data in Figure 3.9. Most data are sampled nearly regularly, but there are also some longer gaps.  \\

3.11 Four spectrum estimates using the Kepler-91 flux data from Figure 3.9. The top panel is the usual Lomb-Scargle periodogram. The second panel is the mtLS spectrum estimate with NW=4 and K=7 tapers. The third panel is the mtLS spectrum estimate with NW=10 and K=19 tapers. The bottom panel is the mtLS spectrum estimate with NW=20 and K=39 tapers. The progression shows the tradeoff between frequency resolution and estimator variance.
3.12 The mtLS spectrum estimate with NW=10 and K=19 tapers. Panel (a) is the spectrum estimate across all frequencies. Panel (b) is the low-frequency portion of the spectrum; lines show the fundamental frequency and harmonics of the previously identified transiting planet. Panel (c) is the band from 8-12 cycles per day, showing many apparent signals. Panel (d) is the highest-frequency portion of the spectrum. This portion of the spectrum should be interpreted with caution due to the possibility of leakage or aliasing from distant parts of the spectrum.

3.13 Frequency response of the mtLS for the Kepler-91 time series. Frequencies above about 10 cycles per day show pseudo-aliasing to higher frequencies.

4.1 The prior hierarchy for linear model coefficients from Ishwaran and Rao (2005), equation 4 therein.

4.2 Posterior distribution of non-centrality parameters for the Greyscale series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
4.3 Posterior distribution of non-centrality parameters for the Humification series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.

4.4 Posterior distribution of non-centrality parameters for the Ash Content series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.

4.5 Posterior distribution of autoregressive noise spectrum parameters from the Tibetan peat core. The variance parameters for the three series are: (a) for Greyscale; (b) for Humification; (c) for Ash Content. Similarly, the $\phi_1$ parameters for the three series are: (d) for Greyscale; (e) for Humification; (f) for Ash Content.
4.6 Frequency-wise mtLS spectral statistics for the series from the Tibetan peat core. The spectral statistics have been calculated at high frequency resolution in order to examine characteristics within-band. The heavy line is the mean spectral statistic across 1000 posterior chronology samples, and finer lines are the 2.5 and 97.5 percentiles. Gray lines delineate the $2W$ intervals.

4.7 Posterior distribution of non-centrality parameters for the salinity index series from Chauvin Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.

4.8 Posterior distribution of non-centrality parameters for the salinity index series from Humboldt Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
4.9 Posterior distribution of non-centrality parameters for the salinity index series from Oro Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.  

4.10 Posterior distribution of autoregressive noise spectrum parameters from the prairie lakes data. The variance parameters for the three series are: (a) for Chauvin; (b) for Humboldt; (c) for Oro Lake. Similarly, the $\phi_1$ parameters for the three series are: (d) for Chauvin; (e) for Humboldt; (f) for Oro Lake.  

4.11 Frequency-wise mtLS spectral statistics for the series from the prairie lakes data. The spectral statistics have been calculated at high frequency resolution in order to examine characteristics within-band. The heavy line is the mean spectral statistic across 1000 posterior chronology samples, and finer lines are the 2.5 and 97.5 percentiles. Gray lines delineate the 2$W$ intervals.
Lists of abbreviations and symbols

Table 1: List of symbols used throughout this thesis.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>Data, a realization of a time series process</td>
</tr>
<tr>
<td>$t$</td>
<td>Times at which a time series process is observed</td>
</tr>
<tr>
<td>$\Delta t$</td>
<td>Time increments – the first differences between $t$</td>
</tr>
<tr>
<td>$N$</td>
<td>The number of observations in a time series</td>
</tr>
<tr>
<td>$d$</td>
<td>The section depths in a record</td>
</tr>
<tr>
<td>$M$</td>
<td>The number of ageing estimates in a record</td>
</tr>
<tr>
<td>$a$</td>
<td>The $^{14}C$ ages (non-calibrated)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>The $^{14}C$ age standard errors</td>
</tr>
<tr>
<td>$f(a^c; a, \sigma)$</td>
<td>The calibrated age likelihoods</td>
</tr>
<tr>
<td>$a^c$</td>
<td>The calibrated ages</td>
</tr>
<tr>
<td>$v$</td>
<td>Depth increments – the first differences between $d$</td>
</tr>
<tr>
<td>$u$</td>
<td>Time increments</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>The set of time indices corresponding to the calibrated ages $a^c$</td>
</tr>
<tr>
<td>$\alpha, \lambda$</td>
<td>Accumulation process parameters (Gamma)</td>
</tr>
<tr>
<td>$w$</td>
<td>Standardized time increments (time increment per unit depth)</td>
</tr>
</tbody>
</table>
\( \beta \) Accumulation process rate parameter for standardized time increments

\( \alpha_1, \beta_1 \) Parameters of hyperprior for \( \beta \) (equation 2.2)

\( \alpha_2, \beta_2, \gamma_2 \) Parameters of hyperprior for \( \alpha \) (equation 2.4)

\( \mu, \tau^2_w \) Accumulation process parameters (Normal)

\( \mu_1, \tau^2_1 \) Parameters of hyperprior for \( \mu \) (equation 2.6)

\( \sigma_L, \sigma_U \) Parameters of hyperprior on \( \tau^2_w \)

**Chapter 3**

\( X \) Sequence of random variables generating a time series

\( t \) Observation time

\( D_t \) Set of observation times

\( \mu_t \) Mean of the random variable \( X_t \)

\( c \) Time interval for regular sampling

\( s \) Time lag

\( N \) Length of a time series

\( R() \) The autocovariance function

\( S() \) The spectral density function

\( f \) Frequency

\( \hat{S}_P() \) The periodogram

\( \mathcal{F}() \) Fejér’s kernel

\( \hat{S}_D() \) Direct estimator of the spectrum

\( h_t \) Data taper

\( v_t^{(k)}(N,W) \) The DPSS taper of order \( k \) with parameters \( N \) and \( W \)

\( W \) The bandwidth parameter of the DPSS tapers
The number of DPSS tapers used in a multitaper spectrum estimate

The $k$th eigenspectrum of the multitaper estimator

The multitaper spectrum estimator

The Lomb-Scargle periodogram

The Lomb-Scargle periodogram time offset

The angular frequency

Components of the decomposition of the Lomb-Scargle periodogram for fast computing (equations 3.11 and 3.12)

Components of the Lomb-Scargle periodogram (equation 3.15)

nfft library time series and Fourier coefficients

nfft library matrix notation for Fourier transform

The output of the adjoint (conjugate transpose) nfft

The average sampling interval for a set of irregular times

Interpolated DPSS weights

First time differences

Spectral density function

Variance parameter for noise distribution

The $j$th autoregressive parameter for an AR($p$) process

The autoregressive parameter for an AR(1) process

The characteristic decay time for an AR(1) process

The non-centrality parameters for noncentral chi-squared distributions
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\varrho)</td>
<td>Mixture-distribution parameters for (non)central chi-squared distributions</td>
</tr>
<tr>
<td>(\zeta)</td>
<td>Vector of hyperparameters for non-centrality parameters</td>
</tr>
<tr>
<td>(\beta^{SS})</td>
<td>Linear model coefficients in the spike-and-slab prior model of Ishwaran and Rao (2005)</td>
</tr>
<tr>
<td>(\delta_{v_0}, \delta_1)</td>
<td>Discrete probability measures in the spike-and-slab prior model of Ishwaran and Rao (2005)</td>
</tr>
<tr>
<td>(\mathcal{I}, \tau^2, w^{SS}, v_0)</td>
<td>Parameters of the spike-and-slab prior for linear model coefficients from Ishwaran and Rao (2005)</td>
</tr>
<tr>
<td>(a_1, a_2)</td>
<td>Parameters of the prior on (\tau^2) for the spike-and-slab prior model of Ishwaran and Rao (2005)</td>
</tr>
<tr>
<td>(\alpha_\tau, \beta_\tau)</td>
<td>Parameters of the prior on (\tau^2) for the current spike-and-slab prior</td>
</tr>
<tr>
<td>(p_\varrho)</td>
<td>Parameter of the prior on the (\varrho) parameters</td>
</tr>
<tr>
<td>(\vartheta)</td>
<td>Vector of spectral model parameters</td>
</tr>
<tr>
<td>(x)</td>
<td>A time series</td>
</tr>
<tr>
<td>(\mathcal{T})</td>
<td>Set of time parameters</td>
</tr>
<tr>
<td>(S(x, \mathcal{T}))</td>
<td>Spectral statistics as a function of the time series and the times</td>
</tr>
<tr>
<td>(\sigma_{p,L}, \sigma_{p,U})</td>
<td>Parameters for the prior on (\sigma_p^2)</td>
</tr>
</tbody>
</table>
Table 2: List of abbreviations used throughout this thesis.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR and ARMA</td>
<td>Autoregressive and Autoregressive Moving Average</td>
</tr>
<tr>
<td>BP</td>
<td>Before present (1950 <em>anno Domini</em>)</td>
</tr>
<tr>
<td>CRPS</td>
<td>Continuous ranked probability score</td>
</tr>
<tr>
<td>DPSS</td>
<td>Discrete prolate spheroidal sequences</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov Chain Monte Carlo</td>
</tr>
<tr>
<td>mtLS</td>
<td>Multitaper Lomb-Scargle</td>
</tr>
<tr>
<td>NDFT</td>
<td>Non-equispaced discrete Fourier Transform</td>
</tr>
<tr>
<td>nfft</td>
<td>Non-equispaced Fast Fourier Transform library</td>
</tr>
<tr>
<td>RMSEP</td>
<td>Root mean squared error of prediction</td>
</tr>
<tr>
<td>T&amp;T</td>
<td>Trachsel and Telford (2017)</td>
</tr>
</tbody>
</table>
Chapter 1

General Introduction

Time series data are collected sequentially as a function of time. The majority of time series data are collected at known times, because you usually have to decide to collect a data point. This control over the data collection process means that observation times tend to be known and regularly-spaced, meaning that the first differences of the observation times are all the same. As a result, most methods of time series analysis are for data collected at known and regularly-spaced times. However, there are some cases in which the observation times cannot be controlled. In these cases, the times will be irregular and possibly unknown, making the usual methods of analysis inappropriate. In these situations, time series have sometimes been referred to as temporal series, to emphasize that the times are not known precisely.

1.1 Spectral analysis of time series

What differentiates time series analysis from many other fields of statistics is: (1) an interest in the dependence between random variables that generate the time series data; and (2) a lack of replication. Colloquially, (1) knowing the weather yesterday tells us something about what the weather will be today, and (2) we can only
observe today’s weather once. For Gaussian processes, the dependence between the random variables depends only on the covariance between them. Wanting to know about the covariance between random variables without replication requires that we make a simplifying assumption, the most popular being that of stationarity. Precise definitions are presented in Chapter 3, but in lay terms stationarity assumes that only time differences matter rather than the times themselves. This assumption allows covariance to be written as a function of time lag, and structures the covariance between the random variables that generate a time series to depend only on the difference between their observation times. This covariance function is often called the autocovariance. In the case of regularly sampled time series, the stationarity assumption creates replicates that can be used to estimate the autocovariance function at a set of discrete lags. There exists an alternative representation called the spectrum which is the Fourier transform of the autocovariance. Thus, statistical inference can be performed by estimating either the autocovariance or the spectrum, and there is a rich literature in spectrum estimation for regularly sampled time series (see Percival and Walden, 1993, for a comprehensive overview).

In their review of spectral analysis for irregularly sampled time series, Babu and Stoica list several categories of irregularity:

1. A time series that is sampled regularly but some samples are missing.

2. A time series that is sampled randomly but following a specified distribution.

3. A time series that is sampled by a sampling device that introduces jitter but that would otherwise be regularly sampled.

4. A time series that is sampled randomly within intervals of fixed time length.
1.1. SPECTRAL ANALYSIS OF TIME SERIES

5. A time series that is sampled arbitrarily.

For the purposes of this thesis, I am interested mostly in time series that would be considered to be sampled arbitrarily because this is the least restrictive assumption and the other categories of irregularity are not likely to be suitable for the particular application.

Babu and Stoica (2010) define three classes of spectral analysis methods for irregularly sampled time series. The first is based on least squares, and includes the classical periodogram (Schuster, 1898), the Lomb-Scargle periodogram (Scargle, 1982), and the Iterative Adaptive Approach of Stoica et al. (2009). This would also include the work of authors focused on the case where sampling times are realizations of a stationary random process that is independent of the process that gave rise to the time series (Brillinger, 1984; Masry, 1978; Brillinger, 1973; Thomson and Robinson, 1996). The second class of methods is based on interpolation of the autocovariance sequence using various interpolation techniques. The interpolated autocovariance can then be transformed to give a spectrum estimate. The third class of methods is based on slotted resampling, a heuristic approach which involves: (1) defining a set of regularly-spaced points in the time domain; (2) defining a slot around each point; and (3) assigning the nearest sampled value to each regularly-spaced point, provided there is a sampled value within the slot, otherwise the point is treated as missing. The result of the slotted resampling is a regularly spaced time series, possibly with missing values, which can be analyzed using methods for regularly sampled series.

In this thesis we require a statistic which is fast and reliable to compute because we are in a situation of not only irregular but unknown times, and will be computing distributions of results using Monte Carlo methods. We also wish to avoid making
strong assumptions about the spectrum or autocovariance of the process. The spectral statistic in Chapter 3 is in the class of least-squares estimators and satisfies both of these requirements.

1.2 Bayesian data analysis

Bayesian data analysis is the application of Bayes’ Rule to infer a distribution of statistical model parameters given observed data \( x \). Consider a sampling model for \( x \), written \( p(x|\theta) \), indexed by a parameter vector \( \theta \). In Bayesian data analysis, the objective is to evaluate the distribution \( p(\theta|x) \), known as the posterior distribution for \( \theta \). This is done using Bayes’ Rule:

\[
p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}.
\]

Specification of a prior for model parameters, \( p(\theta) \), is required. This prior does not depend on the data \( x \), so as the name suggests, it should represent a priori knowledge of the parameter \( \theta \). The final term, \( p(x) \), is a constant for fixed \( x \), so Bayes’ Rule can be written \( p(\theta|x) \propto p(x|\theta)p(\theta) \).

In some simple cases, the posterior distribution can be evaluated directly. In the general case, the posterior distribution can be approximated using numerical integration. A very common approach is to obtain samples from the posterior distribution by construction of a Markov Chain having stationary distribution that is the same as the posterior. Gelman et al. (2013) and Carlin and Louis (2008) provide comprehensive guides to Bayesian data analysis.
1.3 Paleoclimate and sedimentary records

This thesis is concerned primarily with the analysis of time series collected in paleoclimate research. These time series are comprised of samples obtained from cores extracted from a sedimentary record. For example, sediment collects at the bottom of lakes over time, with the youngest sediment at the sediment/water interface and older sediment buried below. In studies of paleolimnology, cores are extracted from the lakebed and sampled as a function of depth (Figure 1.1). Rather than a time series, these data are a depth series, and a necessary first step is to map depths to ages of samples.

Climate change is a topic of contemporary interest and importance (IPCC et al., 2014). However, anthropogenic effects on climate exist together with baseline climate variation. A historical record of climate changes over the Holocene (the period since the last major ice age, about 11,700 years ago (Walker et al., 2009)) thus puts current climate and recent climate changes in context.

A major source of information about climate over the Holocene are environmental proxies residing in sedimentary records such as lake cores (Last et al., 2001). Quantifying variation in climate proxies as a function of depth can be achieved using core sampling. What remains is to link depth in the core to age of the sample. Many ice cores, some lake cores, tree rings, and speleothems show regular layers called *varves* that can be counted to determine the age of samples (recognizing that some layers could be missed). For records without varves, another method is required. In these records, samples of known depth are taken and dated using dating methods that are subject to error. In particular, radiometric dating using $^{14}$C or $^{210}$Pb is common.

Radiocarbon dating can be used to age biological samples up to 50,000 years old...
1.3. PALEOCLIMATE AND SEDIMENTARY RECORDS

Figure 1.1: Photos of two sediment cores extracted from Owikeno Lake in British Columbia, August 2004. The left photo shows the frame used to level the core against the bottom. Various methods are used to drive the coring mechanism into the sediment and retain the sediment for retrieval. The core is extruded and sectioned into samples as a function of depth once retrieved. The author was involved in collection of these core samples.
The radioactive carbon isotope $^{14}\text{C}$ is produced by cosmic rays in Earth’s atmosphere. The $^{14}\text{C}$ is taken up by plants and decays to nitrogen over time. The amount of $^{14}\text{C}$ relative to other isotopes of carbon can thus be used as a clock if the initial abundances are known, given a known rate of decay of $^{14}\text{C}$. Unfortunately, the amount of $^{14}\text{C}$ in the atmosphere has not been constant over time, and the age determined from the relative abundance of $^{14}\text{C}$ must be calibrated by referencing to samples whose age can be estimated independently. The calibration results in a complicated, sometimes multi-modal, and non-parametric calendar age distribution for a given sample (Figure 1.2). Recent human activities have meant that radiocarbon dating can’t be used to determine ages of samples deposited prior to 1950; it is conventional to report dates as years before present (BP), where present is the year 1950 AD. For these more recent samples, $^{210}\text{Pb}$ dating can be used instead (Appleby, 2008).

For core data taken from a sedimentary record, the mapping of depths to ages is imperfect and the times are not known precisely. Despite this uncertainty, we would like to analyze the time series of samples from the sediment core to infer past climate conditions. A method of time series analysis that incorporates timing uncertainty is required.

1.4 Organization of thesis

This thesis is organized into three chapters, representing three key contributions. The chapters are intended to stand on their own and contain relevant reference to the literature. Together, these three contributions achieve the overall objective of time series analysis and inference for paleoclimate records with unknown times.
Figure 1.2: The distribution of radiocarbon ages as a function of true calendar ages for the Holocene. Each panel shows the mean radiocarbon age ± two standard deviations for a given calendar age, as estimated by Reimer et al. (2013) using the methodology of Niu et al. (2013). The relationship between radiocarbon age and calendar age is not monotonic, nor does it follow a 1:1 relationship on average. A given radiocarbon age can imply a complicated and possibly multimodal distribution of calendar ages.
The first chapter (Chapter 2) is a new Bayesian hierarchical model for the chronology of a sedimentary record. The chronology of a sedimentary record is the relationship between depth and age. This chapter builds on previous chronology modelling approaches, and results in a posterior distribution for the chronology that is used in subsequent analysis of climate proxy series collected from the same core. The key difference between the hierarchical model of Chapter 2 and previous approaches is the inclusion of a group-level prior on accumulation time between proxy measurements. This prior structure results in a compromise between the ages obtained from radiocarbon dates and the overall chronology of the core, limiting unrealistic accumulation rates that plague some previous approaches. Moreover, the method does not require specification of informative priors, making it broadly applicable. Some of the work in Chapter 2 was previously published (Springford, 2013; Springford and Thomson, 2015).

The second chapter (Chapter 3) is a new method for fast computation of a spectral statistic for irregularly sampled time series. Viewed as an estimator of the spectrum, the method is a compromise between an optimal estimator that is difficult and expensive to compute, and the most common estimator that is less expensive to compute but far from optimal. The estimator is identical to the optimal solution in the case of regular sampling. The statistic can be used as an estimator of the spectrum for time series that have irregular but known times, and I provide an example application to a time series from astronomy. Thanks to a method for rapid computation, the statistic can be used in numerical sampling schemes such as Markov Chain Monte Carlo for Bayesian inference.

The third chapter (Chapter 4) is a Bayesian model for the spectrum of a time
series process that includes timing uncertainty (using the posterior chronology from Chapter 2) as well as time series process uncertainty. This third chapter makes use of a product likelihood based on the spectral statistic from Chapter 3. A parametric model for the spectrum that combines concepts from previous spectral analysis of climate time series is cast in a Bayesian context. A prior structure that is used in modern analyses of high-dimensional regression models is employed to identify signals in the spectrum. The approach is applicable to Bayesian spectral modelling in cases of known or unknown as well as regular or irregular times.
Chapter 2

A Bayesian Hierarchical Chronology Model for Time Series Analysis of Paleoenviromental Data

In this chapter, I present an approach for modelling the chronology of a time series record that is common in paleoenvironmental studies consisting of cores. In these studies, a core is extracted from a sedimentary record, so that measurements are taken as a function of depth. A necessary first step in the analysis of these data is to translate depths to ages. This is what I refer to as the chronology. Modern approaches model the chronology by modelling the accumulation of the core over time. The major contribution presented here is to include not only a group-level prior distribution for the time-to-accumulation of each core section, but also a specified hyperprior distribution on the parameters of the prior that is at most weakly informative. The model hierarchy results in a compromise between the average time-to-accumulation of the core and the local age likelihoods, which due to the process of selecting samples can have larger errors than those implied by ageing labs. The hierarchy has the added benefit of absolving the analyst from having to make informative assumptions about the distribution of times-to-accumulation when no prior
knowledge exists. The chronology model parameters are sampled using a Markov Chain Monte Carlo (MCMC) sampler. Samples from the posterior distribution of chronologies (i.e. times) will then be used to compute posterior distributions for time series parameters in subsequent chapters.

2.1 Introduction

Time series data consist of observations collected sequentially in time. The basic structure of the data is

\[
\begin{array}{ccccccc}
  x_1 & x_2 & x_3 & \ldots & x_{N-1} & x_N \\
  t_1 & t_2 & t_3 & \ldots & t_{N-1} & t_N \\
  \Delta t_1 & \Delta t_2 & \ldots & \Delta t_{N-1}
\end{array}
\]

where \( t_i \leq t_j \) for \( i < j \). In most cases, the \( \{t_i, i = 1, 2, \ldots, N\} \) are known, and in many cases the time increments \( \Delta t_i = t_{i+1} - t_i \) are equal for all \( i = 1, 2, \ldots, N - 1 \). However, it is sometimes the case that the time increments are not all the same, especially in observational data where the observer doesn’t control the observation times. Common examples include time series that are evenly sampled but that contain missing values, or event data such as nerve impulses (Brillinger, 1984). In other cases, we may not observe the times directly and are forced to treat them as latent. Common examples are paleoenvironmental studies based on extracted cores, which are composed of material that has accrued over time. For example, paleolimnology studies often collect cores from lakebed sediments. The cores are sectioned into depth slices, and the material from the slices are analyzed for proxies of past environmental conditions (Last et al., 2001). Although the slices are often (but not necessarily) uniform in depth, the corresponding time increments are unknown and vary in length. Other
common examples of core data include ocean sediments, peat, rock and terrestrial soil, trees, and ice.

Paleoenvironmental research is focused on describing phenomena that vary with time and often have periodic dynamics. Thus, time series modelling and spectral analysis are the statistical tools necessary for description and inference. Techniques for time series modelling and spectrum estimation are well-established for series that are evenly sampled at known intervals. Common models for time series include autoregressive moving average (ARMA) models, state-space models such as Kalman filtering (Brockwell and Davis, 2009), and additive models that attempt to separate trend, seasonal, and noise components (Cleveland and Tiao, 1976). When producing spectral estimates, averaging estimates from different segments in time or from multiple orthogonal tapers reduces variance, and the use of a fast Fourier transform reduces computation time (Percival and Walden, 1993).

Core data face particular challenges when it comes to obtaining time series model parameter estimates and spectral estimates. First, the time series data tend to be limited in length due to the logistics of the sampling process, the availability of records in nature, and high sampling costs associated with sectioning and analyzing the core for proxies of interest, leading to estimates with large variance compared to longer series of the same process. Second, the data are irregularly sampled in time. The irregular sampling limits the types of models that can be used, and can also increase computational burden. For example, spectral estimates are more expensive to compute because the fast Fourier transform relies on uniform sampling. Third, the data are sampled at unknown times; the relationship between depth in the core and time is uncertain. Errors in timing result in errors in estimates of model parameters and
difficulties estimating the spectrum. In particular, detection of periodic components that manifest as peaks in the spectrum, and the estimation of the frequencies of peaks can suffer when times are unknown. A quantitative way of incorporating this uncertainty into the analysis is required (Blaauw and Heegaard, 2012). The balance of this chapter is focused on addressing the unknown relationship between depth in the core and age of a sample using a Bayesian hierarchical chronology model.

2.1.1 Modelling chronologies

Chronology modelling involves estimating times (ages) using at a minimum some sort of age proxy. Times are obviously necessary for the analysis of time series. However, chronology models can vary in their interpretation of the available data, and as a result provide different estimates of the chronology. I propose a short list of characteristics that a chronology model for core data should possess:

1. The ages are monotonically increasing with depth in the core.

2. The ages are robust to anomalies in dating estimates.

3. The model makes full use of the information contained in the dating estimates (i.e. the likelihood of the data).

4. Quantification of uncertainty in the ages can be propagated to spectrum estimates or estimates of time series model parameters.

Feature (1) is fundamentally a result of the deposition process assumed to have taken place in the core – older layers of the core were deposited before newer ones, and thus must be deeper in the core. Although this assumption might be violated in
practice, such violations can be assumed to be evident to trained experts when the core is examined.

Feature (2) is desirable because several factors can influence the dating estimates that are reported by labs. For example, if the core consists of sediments and the dating estimate is inferred from a $^{14}$C isotope ratio, then the material chosen for ageing (e.g. a piece of organic material such as a chip of wood) might correspond to a different age than the bulk material at the same depth. This type of error is exacerbated when the variance of the dating estimate as reported by the lab is underestimated, because we are falsely confident in the apparent age, which is different than the true age of the bulk material.

Feature (3) is desirable because of the somewhat unconventional distributions that can arise when radio-isotope dates, in particular $^{14}$C dates, are calibrated. $^{14}$C dates must be calibrated because the ambient concentration of $^{14}$C is not constant over time (Damon et al., 1978; Stuiver et al., 1998; Reimer et al., 2013). In most cases, a particular current-day $^{14}$C isotope ratio maps to multiple possible deposition dates. As a result, the calibrated age distribution for a $^{14}$C sample is often multimodal, and the expected value of the calibrated distribution might have a very small or zero likelihood. In such cases, a point estimate for the calibrated age may not be ideal. Rather, the information contained in the full likelihood function is more representative of the distribution of possible dates and is thus more appropriate.

Feature (4) is desirable because the chronology model does not exist in isolation but rather in the larger context of a time series analysis. As mentioned above, if measures of uncertainties in the timing of events are not propagated through to the time series analysis, then the total amount of uncertainty in time series estimates
will be too small, and we will be overconfident in our conclusions. This could, for example, result in incorrect interpretations of the time series analysis results, such as false positive detection of periodic components.

Previous chronology models can be categorized into three general approaches. The first approach is to assume that time increments are piecewise linear and focus on interpolating between available dating estimates. In the simplest case, this amounts to “connecting the dots” between dating estimates, where the “dots” are often the mean or median of the calibrated age likelihoods. The second approach is to fit a function through the dating estimates. The choice of function family is often arbitrary. For example, in the analysis of the motivating dataset (Section 2.2) taken from Yu et al. (2006), a two-piece linear function was chosen by the authors. Polynomial functions or spline functions are also possible choices. These approaches were compared in a paper by Telford et al. (2004), whose results suggested that when there are a large number of calibrated dates available, cubic smoothing splines were most appropriate. When the number of calibrated dates was small, however, none of the function fits was able to properly characterize the true age versus depth relationship, and estimates were overconfident.

Another related approach is wiggle match dating (Pearson, 1986; Blaauw et al., 2003), where reversals in uncalibrated $^{14}$C dates are used to tune the chronology of the core. This approach can be successful for certain regions of some cores that display the required wiggles.

The third approach is to model the accumulation of the core using what amounts to a state-space model in which the observations are the dating likelihoods and the hidden process is the deposition of the core (e.g. Bronk Ramsey, 2008; Haslett and
Parnell, 2008). Additional data or information thought to influence the deposition process can also be included if available. Previous approaches of this type have met criticism for allowing unrealistic accumulation rates (Blaauw and Heegaard, 2012), but a recent appraisal of the performance of this class of models (Trachsel and Telford, 2017) was more favourable than the function fitting approach examined in Telford et al. (2004).

In this chapter, I present a chronology modelling approach that is based on accumulation of the core, but that avoids unrealistic accumulation rates by assuming that the time increments share a common distribution. This also has the benefit of being an “automatic” procedure in the sense that informative priors on accumulation rate are not required.

2.2 Motivating datasets and methods

In this section I present data obtained from a peat core and a set of three lake sediment cores. I then define my chronology modelling approach and obtain samples from the posterior distribution of model parameters given each of the motivating datasets. Further analysis using the resulting chronologies is presented in subsequent chapters.

Tibetan peat core

Yu et al. (2006) describe a study of past climate on the Tibetan Plateau, with particular attention paid to the strength of summer and winter monsoons in the region. The data come from a peat core that was sectioned into 649 one-centimetre intervals. The core covers much of the Holocene, roughly 11,000 years. Proxy records from each section were obtained, as well as $^{14}$C isotope ratios from a subset of eleven sections.
I converted these raw $^{14}$C age estimates into likelihoods using the calibration software available in the Bchron R package (Parnell, 2016). Calibrated ages are reported as years prior to 1950. The proxy records include peat greyscale and humification, which are thought to be proxies for the strength of summer monsoons, and ash content which is thought to be a proxy for the strength of winter monsoons (Yu et al., 2006).

**Canadian prairie lake cores**

The Tibetan peat core data is in many ways amenable to chronology modelling: (1) There are a reasonable number of $^{14}$C samples; (2) The $^{14}$C samples are spaced throughout the core; (3) Deeper samples have older $^{14}$C ages, so that even if the chronology modelling approach was simply to interpolate calibrated ages there would be no reversals. However, not every core will have these properties. In a given core, it is not possible to determine in advance which layers will be dated using $^{14}$C because dating relies on the availability of organic material within the core. Moreover, interfering processes on the landscape will cause variation in the position of organic material within the core, introducing additional variation in the inferred ages. It is also possible that reported lab uncertainties are underestimated or generally in error (Scott et al., 1990; Boaretto et al., 2002; Scott et al., 2010).

Michels et al. (2007) present evidence for shifts in drought conditions in the Canadian prairies based on records of diatom assemblages collected from three sediment cores taken from three different lakes. Diatoms are algae whose fossils are used to infer an index of lake salinity. They are useful for this purpose because the relative abundances of different species of diatom change as a result of changes in lake salinity and
climate. The three lakes studied were Chauvin Lake in eastern Alberta, Humboldt Lake in central Saskatchewan, and Oro Lake in southern Saskatchewan. Salinity index data are available at a resolution of 2 cm for Chauvin and Oro Lake, and a resolution of 1 cm for Humboldt Lake. The depths of radiocarbon dating samples are irregular to varying degrees (Figure 2.1), and the calibrated ages sometimes exhibit reversals – deeper samples having younger calibrated ages compared to shallower samples.

2.2.1 Bayesian hierarchical chronology model

Here, I will describe a Bayesian hierarchical model for the chronology of a core extracted from a sedimentary record such as a lakebed or a peat bog. The necessary background for the Bayesian approach can be obtained from either Gelman et al. (2013) or Carlin and Louis (2008).

Generally, a chronology model relates a time surrogate to actual times. Core data such as those in the motivating datasets are comprised of depth sections, a subset of which are dated using radio-isotopes or other methods. The data consist of the section depths \( \{d_i : i = 0, 1, \ldots, N\} \), assumed known, and the ages \( \{a_j : j = 1, 2, \ldots, M\} \) with corresponding ageing errors \( \{\sigma_j : j = 1, 2, \ldots, M\} \), from which the likelihoods \( \{f(a^c_j; a_j, \sigma_j) : j = 1, 2, \ldots, M\} \) for the calibrated ages \( \{a^c_j : j = 1, 2, \ldots, M\} \) are obtained. From the section depths, the depth increments are the first differences \( \{v_k : k = 1, 2, \ldots, N\} \), \( v_k = d_k - d_{k-1} \), each with a corresponding (latent) time increment \( u_k \). The focus of the chronology model is to estimate the latent time increments \( \mathbf{u} \) to obtain times as the cumulative sums \( t_k = t_0 + \sum_{l=1}^{k} u_l, k = 1, 2, \ldots, N - 1 \). Denote \( t_0 \) as the baseline time – that is, the time at the very top of the core. The prior for the baseline time can be taken to be diffuse, for example a Normal distribution.
2.2. MOTIVATING DATASETS AND METHODS

Figure 2.1: Salinity index data for the three lakes described in Michels et al. (2007). Note that the cores are not all sampled to the same sediment depth. The salinity index is derived from counts of diatom fossils within each layer of the core. The salinity index is available at a resolution of 2 cm for Chauvin and Oro lakes, and 1 cm for Humboldt Lake. Open circles represent the depths of radiocarbon dating samples used to infer chronologies.
with large variance. To relate the ageing data to the latent times, let $\varphi_j$ be the index $k$ of $t$ corresponding to the age $a_j^c$. Modelling the time increments and forcing the condition $u_k \geq 0$ for all $k$ restricts the time-depth relationship to be monotonically increasing.

I developed the prior distribution on the time increments using the following logic: Imagine the accumulation of a depth increment $v_k$ to be the accumulation of $\alpha_k$ smaller increments, each of which has a waiting time that follows an Exponential distribution with rate parameter $\lambda_k$ (a Poisson process). Then the waiting time $u_k$ for the entire increment $v_k$ will follow a $\text{Gamma}(\alpha_k, \lambda_k)$ distribution. If we standardize the time increments $w_k = u_k / v_k$, then $w_k \sim \text{Gamma}(\alpha_k, \lambda_k / v_k) = \text{Gamma}(\alpha_k, \beta_k)$.

Based on this development, I assume that the standardized time increments follow the prior

$$w_k \sim \text{Gamma}(\alpha, \beta) \text{ for all } k \quad (2.1)$$

Note that this prior is common to all of the standardized time increments (only one $(\alpha, \beta)$ rather than $k$ separate prior distributions).

The radiocarbon age likelihoods are the probability distributions for the radiocarbon ages given calendar ages representing the true ages of the samples and ageing errors $\sigma$. Likelihood functions $f(a_j^c; a_j, \sigma_j)$ for the calendar ages $a_j^c$ were obtained from Parnell (2016) using the calibration estimated by (Reimer et al., 2013) and the methodology of (Niu et al., 2013), where $a_j^c = t_{\varphi_j} = t_0 + \sum_{l=1}^{\varphi_j} u_l = t_0 + \sum_{l=1}^{\varphi_j} w_l v_l$. These likelihoods tend to be highly variable in form from sample to sample; calibrated likelihoods associated with younger ages tend to be better estimated, but are often multimodal.
The hyperpriors for $\alpha$ and $\beta$ were chosen to be conjugate to the Gamma distribution in equation 2.1. A conjugate prior is one that results in a posterior from the same family of distributions as the prior, which can be advantageous because the posterior is available in closed form. The advantage of having a hyperprior that is conjugate to the prior is that the conditional distribution of hyperprior parameters is available in closed form, making implementation of a Gibbs sampler (Geman and Geman, 1984) more straightforward.

The gamma distribution as used here is defined

$$\text{Gamma}(w_k|\alpha, \beta) = \beta^\alpha \frac{1}{\Gamma(\alpha)} w_k^{\alpha-1} e^{-\beta w_k}.$$  

The hyperprior for $\beta$ was taken to be a Gamma distribution with hyperparameters $\alpha_1 > 0$ and $\beta_1 > 0$. Conditional on other model parameters, this results in a Gamma distribution for $\beta$ with updated hyperparameters

$$\alpha'_1 = \alpha(N - 1) + \alpha_1$$
$$\beta'_1 = \frac{\beta_1}{1 + \beta_1 \sum_{k=1}^{N-1} w_k}.$$  

The hyperprior for $\alpha$ was taken to be

$$p(\alpha|\alpha_2, \beta_2, \gamma_2) \propto \begin{cases} \frac{\alpha^{\alpha-1} \beta^{\gamma_2}}{\Gamma(\alpha) \beta_2^{\gamma_2}} & \alpha > 0 \\ 0 & \text{otherwise} \end{cases}$$

with hyperparameters $\alpha_2, \beta_2, \gamma_2 > 0$. Conditional on other model parameters, this
results in a distribution as in equation 2.3 for $\alpha$ with updated hyperparameters

$$
\alpha'_2 = \alpha_2 \prod_{k=1}^{N-1} w_k
$$

$$
\beta'_2 = \beta_2 + (N - 1)
$$

(2.4)

$$
\gamma'_2 = \gamma_2 + (N - 1).
$$

Notice that the conditional distribution for $\alpha$ requires integration of the kernel in equation 2.3, which has no closed form and must be evaluated numerically.

Numerical issues arise if $\alpha$ becomes large. For example, $\alpha = 100 \implies \Gamma(\alpha) = 9.3 \times 10^{155}$. When posterior weight is given to values of $\alpha > 100$, which is the case for all data sets examined in this thesis, this computing limitation makes the use of a Gamma prior on $\mathbf{w}$ difficult or impossible. However, if $\alpha$ is large the Gamma prior is approximately Normal. For fixed $\beta$, large values of $\alpha$ correspond to large numbers of increment layers, and each increment layer’s waiting time is assumed to follow an Exponential distribution. Denote the waiting time for increment layer $i$ as $\delta_i \sim \text{Exp}(\beta)$. If there are $\alpha$ such layers, then $\sum_{i=1}^{\alpha} \delta_i \sim \text{Gamma}(\alpha, \beta)$. Applying the Central Limit Theorem,

$$
\frac{\sum_{i=1}^{\alpha} \delta_i - \alpha/\beta}{\sqrt{\alpha/\beta}} \xrightarrow{D} N(0,1)
$$

as $\alpha \to \infty$, where $\xrightarrow{D}$ means converges in distribution, and $\sum_{i=1}^{\alpha} \delta_i$ has approximately a $N(\alpha/\beta, \alpha/\beta^2)$ distribution. Thus, when posterior weight is given to values of $\alpha > 100$ (and computational issues arise), I assume a Normal prior for the standardized latent time increments

$$
w_k \sim N(\mu, 1/\tau^2_w) \text{ for all } k
$$

(2.5)
where $\mu$ is the mean and $\tau_w^2$ is the information (reciprocal of variance) of the distribution. Equation 2.5 replaces the Gamma prior (equation 2.1).

Hyperpriors for the Normal prior (equation 2.5) are on $\mu$ and $\tau_w^2$. The hyperprior for $\mu$ was taken to be

$$p(\mu|\mu_1, \tau_1^2) = \frac{\tau_1}{\sqrt{2\pi}} \exp\left(-\frac{\tau_1^2}{2}(\mu - \mu_1)^2\right) \quad (2.6)$$

with hyperparameters $\mu_1$ and $\tau_1^2$. The conditional distribution for $\mu$ in the posterior is as in 2.6, but with updated hyperparameters

$$\mu'_1 = \frac{\mu_1 \tau_1^2 + \tau_w^2 \sum_{k=1}^{N-1} w_k}{\tau_1^2 + (N-1)\tau_w^2} \quad (2.7)$$

$$\tau_1' = \tau_1 + (N-1)\tau_w^2$$

For $\tau_w^2$, the conjugate hyperprior is a Gamma distribution, however this hyperprior choice is not recommended in general because it is difficult to make weakly informative with respect to the posterior (Gelman, 2006). I chose instead a weakly informative uniform prior on $\tau_w^{-1}$

$$p(\tau_w^2|\sigma_L, \sigma_U) = \begin{cases} \frac{1}{\sigma_U - \sigma_L} & \sigma_L \leq \tau_w^{-1} \leq \sigma_U \\ 0 & \text{otherwise} \end{cases}$$

with hyperparameters $\sigma_L$ and $\sigma_U$, $0 \leq \sigma_L < \sigma_U$. The conditional distribution for $\tau_w^2$
in the posterior is

\[ p(\tau_w^2 | \sigma_L, \sigma_U, \mu, w_1, \ldots, w_k) \propto \begin{cases} 
\tau_w^{N-1} \exp \left( -\frac{\tau_w^2 \sum_{k=1}^{N-1} (w_k - \mu)^2}{2} \right) & \sigma_L \leq \tau_w^{-1} \leq \sigma_U \\
0 & \text{otherwise} \end{cases} \]

(2.8)

which can be sampled numerically.

### 2.2.2 Sampling chronology model posterior parameter distributions

The conditional definition of the chronology model in the previous section makes computation of the posterior distribution for model parameters using a Gibbs sampler (Geman and Geman, 1984) straightforward. I used a Gibbs sampler with parameter-specific update methods depending on the availability of an analytic conditional posterior (Gelman et al., 2003). This is sometimes referred to as a *Metropolis-within-Gibbs* sampling approach (Carlin and Louis, 2008). For those parameters with analytically integrable conditional distributions in the posterior, parameter updates are draws from the conditional posterior directly since conditional posterior parameters are available using an ordinary Gibbs step (Geman and Geman, 1984). For the latent standardized time increments \( w_1, \ldots, w_{N-1} \) and the age at zero depth \( t_0 \), I used a Metropolis update step (Metropolis et al., 1953). For the \( \alpha \) hyperparameter or the \( \tau_w^2 \) hyperparameter (depending on the choice of prior on \( w \)), I used a slice sampler (Neal, 2003). Note that in all data analysis done here, I used the Normal prior on \( w \).

I began the chains near the expected posterior values for the model parameters, and convergence to the posterior was rapid. I burned in all chains for 1,000,000 samples. Following burn-in, I sampled 5,000,000 iterations of the Markov chain, thinning to 1,000 to save storage space. There was strong autocorrelation and poor
mixing of the $w$ parameters in the chain due to strong anti-correlation. I was able to improve the sampling efficiency of the chain using the method of Roberts et al. (2001), but although the sampling efficiency was improved, the computation time per sample was increased, resulting in no practical reduction in sampling time. Computation was on the order of minutes for either approach.

There are five model parameters that must be specified \textit{a priori} for the Bayesian chronology model presented here. For the Tibetan peat core and the prairie lake cores, the specified parameter values were:

- The parameter of the hyperprior on $\mu$, the mean age increment: $\tau_1 = 1/1000$ (all cases).

- The parameters of the hyperprior on $\tau_w^{-1}$, the age increment standard error: $\sigma_L = 0.001$, $\sigma_U = 100$ (all cases).

These are all intended to be at most weakly informative priors. The prior for $t_0$, the age at the top of the core, varied and depended on the information available. For the Tibetan peat core, it was assumed that the top of the core had a prior age that followed $t_0 \sim N(0, 25^2)$. For the prairie lakes data, priors for the age at the top of the core were chosen to be compatible with the reported chronology and $^{210}\text{Pb}$ dates:

- Chauvin: $t_0 \sim N(-40, 20^2)$
- Humboldt: $t_0 \sim N(0, 10^2)$
- Oro: $t_0 \sim N(80, 20^2)$
2.3 Results

Tibetan peat core

In the case of the Tibetan peat core, the posterior distributions for all model parameters were well-identified (Figure 2.2). In particular, it is evident that the posterior for \( w \) is not Normal - the age likelihoods have had a noticeable effect on the prior. The resulting uncertainties in the chronology are evident when the time series for the three proxies are plotted (Figure 2.3). In particular, times towards the end of the series and between approximately 2800 and 5800 years before present have high associated uncertainty, whereas times between 5800 and 6700 years before present are strongly constrained by the likelihoods of the \( ^{14}C \) data at 305 and 321 centimetres depth.

Canadian prairie lake cores

The Canadian prairie lake cores also had posterior distributions for chronology model parameters that were well-identified (Chauvin in Figure 2.4, Humboldt in Figure 2.5, and Oro in Figure 2.6). The lake sediment chronology data are in some sense more challenging to model than the Tibetan peat core chronology data, although the model described in this chapter did not have any issues. There appears to have been a lot of variation in sedimentation rate for Chauvin Lake in particular, as well as age likelihoods that would otherwise imply reversals in the age versus depth relationship. In the case of Humboldt Lake and Oro Lake, there are also apparent reversals in the age versus depth relationship, and also a lack of contrast in depths at which radiocarbon dates were available; in the case of Humboldt Lake, three out of six of the radiocarbon samples were clustered around a depth of 100 cm (Figure 2.5), and in the
2.3. RESULTS

Figure 2.2: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.
Figure 2.3: One thousand posterior time series of equal probability (light grey lines, over-plotted) and the mean posterior time series (white line). The top panel are the time series for the Greyscale measurements, the middle panel Humification measurements, and the bottom panel Ash Content.
case of Oro Lake, five out of seven of the radiocarbon samples were clustered around a depth of 400 cm (Figure 2.6). Modelling these chronologies using an interpolation approach would seem impossible, which is likely why Michels et al. used a curve fit to the calibrated ages.

The next step in a time series analysis of these data would be to obtain estimates of time series model parameters and/or spectra using the times from the sampled chronologies. The advantage of the Bayesian approach here is obvious since posterior parameter estimates and/or posterior spectrum estimates can be calculated directly from the posterior chronology samples, assuming independence between the environmental proxies and the core chronology. I present an analysis of this type in Chapter 4.

2.4 Comparison to other chronology methods

Trachsel and Telford (2017) (hereafter T&T) examine the performance of four recent chronology models using data from a set of three varved cores. Varves are regular laminations present in some cores that are annual cycles. Assuming that the varve counts accurately represent the passage of years and the true accumulation history of each core, T&T simulate the lab results of radiocarbon isotope ratios from an equally-spaced subset of layers in each core, for 6, 11, 21, 31, and 41 aged layers. The first and last layers in the core are always dated, although this is rarely the case in real-life applications. I will use the same approach to examine the performance of my Bayesian chronology model for an equally-spaced subset of 6, 21, and 41 layers.

The radiocarbon ages are obtained by:

1. *Decalibrating* the true ages to obtain a radiocarbon date, then
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.4: Chauvin Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.5: Humboldt Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.
Figure 2.6: Oro Lake: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_{w}^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted with the posterior age versus depth trajectories are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and three example age trajectories from posterior samples in light gray.
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

2. Drawing from a Gaussian distribution centered on the $^{14}C$ age with a standard deviation of 30 years, which is meant to mimic the lab age determination. The radiocarbon ages are drawn only once – there is no resampling. The results are thus non-repeatable without drawing the same radiocarbon ageing errors. This makes comparisons on measures of performance impossible without repeating the work of T&T. The lack of replication also limits to what extent their study can be generalized. However, comparisons based on the characteristics of the models can still be illuminating. For example: What is the inferred rate of accumulation at depths away from those with radiocarbon measurements? Or, how do the models behave when faced with abrupt changes in accumulation rate?

After the simulated radiocarbon ages are obtained, they are calibrated, and each chronology model is used to infer the chronology of the cores. Various measures of performance are computed in order to compare the performance of the different chronology models for the particular simulated dataset. Despite not being directly comparable, I have calculated three of these metrics here: The root mean squared error of prediction (RMSEP), the coverage probability for 2/3 credible regions at each depth, and the continuous ranked probability score (CRPS) computed from the definition in Hersbach (2000).

The three varved cores used by T&T, and again for this comparison, are:

1. Elk Lake (Bradbury et al., 1993).
3. Iceberg Lake (Loso, 2009).

The varve-derived (assumed true) ages are plotted in Figure 2.7. The age versus depth
relationships show some characteristic differences that are useful points of comparison within and between different modelling approaches. Iceberg Lake, for example, has a nearly constant rate of accumulation throughout the core, possibly due to its short length compared to the other cores. Holzmaar, on the other hand, has at least a few different intervals with obviously different accumulation rates. Finally, Elk Lake has some variation in accumulation rate throughout with a particularly low rate at the deepest depths in the core.

The four chronology models examined in T&T are:

1. OxCal (Bronk Ramsey, 2008).
2. Bchron (Haslett and Parnell, 2008).
4. CLAM (Blaauw, 2010).

The first three (1-3) are Bayesian models that aim to describe the deposition process, as is the case for the model described in this chapter. The final model (4) (Blaauw, 2010) uses smooth splines to interpolate between resampled points drawn from the calibrated age likelihoods. CLAM (presumably) stands for “classical age-modelling”, meaning that it favours a classic curve-fitting approach. It is a curve-fitting technique that is meant to address the complicated calibrated age likelihoods.

The data available for each of the three varved records differed, requiring a different approach for each. The Elk Lake data starts at depth of 0, with an absolute year of 1927. The data are the varve thicknesses. Proxy data for the core had a typical resolution of fifty years. The core spans approximately 10,223 years and has a cumulative thickness of 19,274.27mm, making 50 year intervals approximately 94.26mm
Figure 2.7: The varve-derived (assumed true) age versus depth relationships for the three cores used to compare chronology methods.
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

thick. There would be about 204 such intervals. Ignoring the varve data, and targeting a temporal resolution on the order of fifty years for proxy data, I aggregated to intervals with a thickness of 100mm for analysis.

The Holzmaar data starts at a depth of 121cm, corresponding to a varve age of 419 years BP (absolute year 1531). Rather than having individual varve thicknesses as was the case for Elk Lake, the data are reported in one-centimeter intervals, corresponding to 4-5 years each. As a result, no aggregation was necessary. There are 801 one-centimeter intervals.

The data for Iceberg Lake starts at a depth of 0, with absolute year of 1998. As was the case for Elk Lake, the data are reported as varve thicknesses. The core has 1,557 varves and a total thickness of 5,843.2mm. I targeted an approximate time resolution of five years. Taking approximately five-year intervals corresponds to intervals that are approximately 18.76mm thick. There would be about 311 such intervals. I aggregated the varve data to intervals that are 20mm thick, and removed the portion of the core that came since the year 1950.

There are five model parameters that must be specified \textit{a priori} for the Bayesian chronology model presented here. I used the same model priors for all analyses of the varved cores:

- The parameter of the hyperprior on $\mu$, the mean age increment: $\tau_1 = 1/1000$

- The parameters of the hyperprior on $\tau_w^{-1}$, the age increment standard error: $\sigma_L = 0.001, \sigma_U = 100$

- The prior on the age at the top of the core: $t_0 \sim N(0, 10000^2)$

These are all intended to be at most weakly informative priors.
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

2.4.1 Results

Posterior distributions of model parameters including posterior chronologies were obtained by MCMC sampling as before. In all cases, 500,000 samples were drawn and discarded as burnin before obtaining another 2,000,000 samples (for Elk Lake and Iceberg Lake) or 20,000,000 samples (for Holzmaar), thinning in all cases to 1,000 samples total. The resulting posterior distributions are plotted in Figures 2.8, 2.9, and 2.10 (Elk Lake); Figures 2.11, 2.12, and 2.13 (Holzmaar); and Figures 2.14, 2.15, and 2.16 (Iceberg Lake).

For Elk Lake, the RMSEP ranged from 48.6 years to 210.4 years for the particular simulated data generated here. The results from the current model and those presented in T&T are not based on the same simulated data, so direct comparisons are not justified. In T&T, the RMSEP for Elk Lake ranged from about 40 years in the most data-rich scenarios to about 120 years in the most data-poor scenarios. My results here are characteristically similar, although the RMSEP of 210.4 years in the six sample case stands out. It’s possible that this high RMSEP is highly dependent on the particular realization of the age data.

For Holzmaar, the RMSEP ranged from 82.9 years to 214.7 years for the particular simulated data generated here. In T&T, the RMSEP for Holzmaar ranged from about 50 years in the most data-rich scenarios to about 260 years in the most data-poor scenarios. As for Elk Lake, my results here are characteristically similar, although the RMSEP of 82.9 years in the most data-rich scenario is somewhat higher than the rest. In Figure 2.11, the most noticeable difference between the estimated chronology and the actual chronology appears at the top of the core. The difference appears to be due to the fact that the actual accumulation rate of the core is higher at shallower depths.
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.8: Elk Lake, 41 simulated dated intervals: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
Figure 2.9: Elk Lake, 21 simulated dated intervals: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
Figure 2.10: Elk Lake, six simulated dated intervals: Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
Figure 2.11: Holzmaar, 41 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_\mu^{-1}$, and the age increments $\omega$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.12: Holzmaar, 21 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $\mathbf{w}$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.13: Holzmaar, six simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_{w}^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), $2/3$ credible region coverage, and the continuous ranked probability score (CRPS).
2.4. COMPARISON TO OTHER CHRONOLOGY METHODS

Figure 2.14: Iceberg Lake, 41 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
Figure 2.15: Iceberg Lake, 21 simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau_w^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), $2/3$ credible region coverage, and the continuous ranked probability score (CRPS).
Figure 2.16: Iceberg Lake, six simulated dated intervals. Posterior distributions for the mean age increment $\mu$, the age increment standard error $\tau^{-1}$, and the age increments $w$ (left to right, top panels). Posterior age versus depth trajectories based on 1000 posterior samples (bottom panel). Also plotted are violin plots (in orange) of the likelihood from calibrated $^{14}$C dating, and the true age versus depth relationship in blue, obtained from varve information. The reported measures of performance are the root mean square error of prediction (RMSEP), 2/3 credible region coverage, and the continuous ranked probability score (CRPS).
than it is elsewhere in the core. The difference appears to be due to the compromise between the age likelihood and the average accumulation of the core, induced by the prior.

For Iceberg Lake, the RMSEP ranged from 12.6 to 27.4 years for the particular simulated data generated here. In T&T, the RMSEP for Iceberg Lake ranged from about 10 years in the most data-rich scenarios to about 35 in the most data-poor scenarios. Again, my results here are characteristically similar. In all scenarios, there was an obvious ambiguity in the age likelihoods for shallow depths. In this instance, the prior structure properly resolved this ambiguity and identified an average accumulation rate as having the highest posterior probability.

Coverage probabilities for 2/3 credible regions varied in the simulations conducted by T&T, from close to 30% to nearly 100%. In my simulations using the chronology model presented here, coverage probabilities ranged from 19% to 60%, in all cases underestimated.

The CRPS for the three simulated records had ranges: 23.0 to 124.8 (Elk Lake); 47.1 to 139.1 (Holzmaar); and 6.8 to 18.2 (Iceberg Lake). The CRPS ranges observed in T&T were approximately: 20 to 75 (Elk Lake); 25 to 150 (Holzmaar); and 7 to 30 (Iceberg Lake). These results are characteristically similar to the RMSEP comparison. For the most data-poor Elk Lake scenario, the CRPS was outside of the range observed in T&T. For the most data-rich Holzmaar scenario, the CRPS was somewhat larger than what was observed in T&T.
2.5 Discussion

The chronology model presented here was constructed with the four key features discussed in section 2.1.1 in mind: (1) Monotone chronology; (2) Robust to outliers; (3) Uses the complete dating likelihood; (4) Quantification of uncertainty can be propagated to time series analysis. By modelling the time it takes to deposit a standardized depth increment, and restricting these time increments to be positive, we are able to guarantee a monotonically increasing time series (feature 1). Note that some other approaches, in particular simple piecewise linear approaches, cannot guarantee monotonically increasing times and thus can be problematic for subsequent time series analysis. The Bayesian approach includes the entire likelihood of the dating estimates directly (feature 3), rather than a summary such as the expected value or median value of the dates. The approach also results in a posterior probability distribution for each of the latent ages which can be used in subsequent time series analysis of the environmental proxies (feature 4).

Chronology methods used in the literature tend to model core accumulation, often using piecewise linear formulations (e.g. Blaauw and Christen, 2005; Blaauw et al., 2007, in addition to those mentioned already). OxCal (Bronk Ramsey, 2008) treats the problem in a slightly different way: Rather than model the depth increments, a series of times is first defined, and interpolation between times is done uniformly or using a Poisson number of small depth increments, the thickness of which must be set by the user. This thickness parameter is generally not known and has a direct effect on the variability of the accumulation rate. Another approach equivalent to Bronk Ramsey’s where the number of accumulation increments is Poisson, but the increment size is random (and thus the age versus depth relationship is piecewise
linear), is called Bchron (Haslett and Parnell, 2008). However, Bchron does not allow for information on accumulation rate to be included, resulting in unrealistic accumulation rates in some cases (Blaauw and Heegaard, 2012). The Bacon model (Blaauw and Christen, 2011) also has linear segments. In Bacon, the rate of accumulation follows a Gamma autoregressive model, and a prior for the accumulation rate and autoregressive parameter must be specified, even though this information may not be readily available. The model structure and informative prior information allow Bacon to overcome the unrealistic accumulation rates of Bchron.

The problem of unknown and possibly unrealistic accumulation rates is handled automatically in the Bayesian hierarchical model presented here. According to Bayes’ rule, the posterior distribution for the time increments is a combination of the prior distribution and the dating likelihoods. Note that only the prior distribution family is defined \textit{a priori} in the hierarchical model; the prior distribution for the time increments is also a combination of hyperprior and likelihood. Thus, the hierarchical model is essentially a form of regularization with shrinkage of the individual time increments automatically controlled by the estimated prior distribution (Gelman et al., 2003; Carlin and Louis, 2008).

Estimators that exhibit shrinkage behaviour are well known to have reduced sensitivity to outliers (Box, 1980; Gelman et al., 2003). In the context of the current model, the shrinkage behaviour is evident when comparing the chronology obtained from the Bayesian hierarchical model to the chronology obtained from a cubic polynomial fit to the mean likelihoods and the chronology obtained from a piecewise linear connect-the-dots fit to the mean likelihoods (Figure 2.17). Examining the chronologies, it is clear that the Bayesian hierarchical model chronology shrinks the chronology
trajectory from the individual likelihoods towards the average chronology represented by the cubic polynomial fit. Considering the situation in which one of the likelihoods represents an outlier, the average chronology will move towards the outlier but only slightly due to the influence of the rest of the data, and the shrinkage behaviour will constrain the posterior chronology accordingly. Thus, the Bayesian hierarchical model will have limited sensitivity to outliers (feature 2), which can be evaluated on a case-by-case basis.

I replicated the analysis of T&T using the hierarchical chronology model presented here. Although not an entirely reasonable comparison due to the fact that measures of model performance are dependent on the particular realization of age samples, the comparison is nonetheless illuminating in some aspects. The performance of the current model in terms of RMSEP and CRPS was mostly in-line with what was observed for other popular chronology models. The 2/3 credible region coverage was universally low for the current model. Other models were a mixed bag, with many having credible regions that were too wide T&T. The underwhelming credible region coverage of the current model, particularly for data-poor scenarios, is a potential concern.

Without reference to competing models, the simulation/estimation exercise revealed some aspects of model behaviour that might not have been otherwise obvious. For example, in the case of Elk Lake and Iceberg Lake, at shallower depths and younger ages the model-based interpretation of the chronology follows the average accumulation rate based on the prior structure, even though the calibrated age likelihoods suggest a much faster accumulation rate at shallow depths. On the other
Figure 2.17: Posterior age versus depth trajectories based on 1000 posterior samples (black), plotted as residuals from a cubic polynomial fit to the mean likelihoods of the calibrated $^{14}$C dates. Also plotted are the violin plots (orange) of the likelihoods from calibrated $^{14}$C dating, a dashed zero-line (blue) representing the cubic polynomial fit, and a piecewise linear function connecting the means of the calibrated likelihoods. The Bayesian hierarchical trajectories can be seen as a compromise between the cubic polynomial fit and the piecewise linear function.
hand, this interpretation is not correct in the case of Holzmaar, where the accumulation rate at shallow depths is more rapid than the average for the entire core. In T&T, some models followed the age likelihoods closely and had the correct interpretation for Holzmaar, but some did not. Those models that follow age likelihoods closely seem less likely to have gotten the shallow ages right for Elk and especially Iceberg, but these results were not shown. In real-life cases where we don’t necessarily have radiocarbon samples at the shallowest or deepest depths in the record, extrapolation following the distribution of accumulation rates from the rest of the core seems like a reasonable assumption to make. Often, additional information about \( t_0 \) can be included in a prior, for example from \(^{210}\)Pb dating.

One downside of the hierarchical model is that it isn’t necessarily immediately responsive to rapid changes in sedimentation rate. This can be seen in the posterior chronology plots: in cases where the true accumulation rate changes abruptly, the model chronology tends to vary more slowly. It would be possible to include more than one population of \( w \) parameters for different sections of the core, which could help to resolve rapid changes in sedimentation rate, particularly if some information about the location of the change were available. This would be a worthwhile future work.

There was no inclusion of outliers in the simulation study. Outliers can by caused by an incorrect depth or deposition delay due to landscape processes, for example. This is one instance where the hierarchical compromise between the local information from the calibrated age likelihoods and overall time-to-accumulation prior distribution should result in a model that is robust to outliers. A comparison of the current model’s performance in the presence of outliers, as compared to the existing models
examined by T&T would be a valuable extension to the current work given that outliers are not uncommon. For instance, all three of the prairie lake cores had calibrated radiocarbon dates at nearly the same depth which had very different (non-overlapping) ages (Figures 2.4, 2.5 and 2.6).

Using standard Markov chain Monte Carlo (MCMC) techniques to sample the posterior resulted in inefficient sampling due to poor mixing of the chain for the time increments $w$. The $w$ were sampled using a Metropolis step, and the Metropolis algorithm requires some tuning of the proposal distribution to achieve efficient sampling. I attempted to improve the sampling efficiency by tuning the proposal distribution following Roberts et al. (2001), but the use of a multivariate proposal distribution reduced the speed per sample and did not improve overall sampling efficiency. There exist adaptive Metropolis algorithms such as Haario et al. (2001) that tune the proposal distribution as sampling progresses, and could result in more efficient sampling of the posterior.

In cases where they are available, varves can provide additional information on the accumulation and chronology of cores. Varved core data exhibit periodic banding, associated with processes that operate on an annual cycle, and this data should be included in the likelihood. A popular method that includes varve information is that of Bronk Ramsey (2008) – they suggest treating the detection of varves as independent Bernoulli trials, with the number of varves detected in a core following a Binomial distribution. However, the real number of varves is unknown, so a Negative Binomial distribution for the total number of varves given a particular number of detections would seem more appropriate. Keeping with the goal of estimating the time increments, if the core was varved but sectioned at independent (possibly regular)
2.6. CONCLUSION

intervals for climate proxy measurement, then the number of varve layer counts in each section could be included as a Negative Binomial likelihood with the true number of varve layers and the probability of detection being parameters. Although an enticingly simple starting point, there are at least two potential issues with this choice of likelihood for the varve layer data: (1) The identification of neighbouring layers is likely not the product of two independent Bernoulli trials – if a single layer is missed then the following layer is almost surely not; (2) The true number of varve layers and the probability of detection are likely both mediated by the rate of accumulation, which can vary over the core and could make identification of these parameters difficult. The sensitivity of results to the independent Bernouilli trial assumption would need to be explored before extending the Bayesian hierarchical approach to varved core records.

2.6 Conclusion

Time series analysis is complicated when the times themselves are unknown, as is the case for paleoenvironmental core data. In such cases, a chronology model that relates the unobserved times to depth in the core is a necessary first step for time series analysis. In this chapter, I present a Bayesian hierarchical approach that has several advantages compared to previous approaches. A preliminary analysis shows that the posterior distribution for model parameters is easily sampled, and that sampling efficiency gains are no doubt possible. The availability of the posterior distribution makes subsequent time series analysis straightforward.
Chapter 3

Rapid Computation of an Approximate Multitaper Statistic under Irregular Sampling

In the previous chapter, I described a method for inference of unknown times in the case of paleoclimate core records. The results come in the form of a set of draws from the Bayesian posterior distribution of age versus depth trajectories.

The ultimate objective is inference for the paleoclimate proxy records themselves, and the estimation of age versus depth trajectories is only a necessary first step in this objective. Often, inference amounts to simply plotting the proxy records and examining the general trajectory over time, as in for example the plots in Figure 2.3.

In some cases, examination of the mean value of the proxy over time is considered sufficient for answering the particular research question. However, it is plausible that the climate processes that generated the proxy time series have components that are periodic or that have an autocovariance structure. For inference concerning these components, I will take a spectral analysis approach.

There are several challenges that present themselves when attempting spectral analysis in the context of latent times. In this chapter, I describe a method for rapid
computation of the multitaper spectral statistic (Thomson, 1982) under irregular sampling. Using this method, it is possible to obtain multitaper spectral statistics for inference in the larger Bayesian latent times framework presented in the rest of this thesis. I also hope that it can be an easily-implemented improvement to the most common approach to spectrum estimation when times are known but irregular.

3.1 Models for time series

Consider a sequence of random variables

\[
\{X_t : t \in D_t\}
\]

and assume that the set of times \(D_t\) are strictly increasing. Assume that the random variables have a joint distribution with finite mean

\[
\mu_t = E[X_t], \quad t \in D_t
\]

and finite variance/covariance

\[
cov(X_t, X_r) = E[(X_t - \mu_t)(X_r - \mu_r)], \quad t, r \in D_t.
\]

A time series is a realization \(x\) drawn from \(X\). In some contexts, the \(\{D_t\}\) are also drawn from a distribution of observation times. A time series is regularly sampled if \(t_i - t_{i-1} = c\) for all \(i\), where \(c\) is a constant. Time series that are regularly sampled are the most common, particularly when the observation times can be controlled by the observer.
Of primary interest in the analysis of time series is the covariance of the random variables \( \{X_t\} \). This differentiates time series analysis from much of statistics in which random variables are often assumed to be independent or conditionally independent (spatial statistics has a similar interest in covariance structure). It also presents a problem for inference because typically we only obtain a single time series – that is, a single sample – making it impossible to characterize the covariance without simplifying assumptions about the covariance structure. Easily the most common assumption is that of stationarity, of which there are two flavours:

- **Strong** stationarity assumes that any finite collection of random variables \( \{X_{t_1}, \ldots, X_{t_m}\} \) has the same joint distribution as \( \{X_{t_1+s}, \ldots, X_{t_m+s}\} \), where \( s \in (-\infty, \infty) \) if the times are continuous, and \( s \in \mathbb{Z} \) for integer-valued discrete times.

- Assuming that the variance of each \( X_t \) is finite, **weak** or **second-order** stationarity assumes that \( E[X_t] = \mu \) for every \( t \) and also \( \text{cov}(X_t, X_{t+s}) \) does not depend on \( t \).

If \( \{X_t\} \) follow a multivariate Gaussian distribution of finite dimension, weak stationarity implies strong stationarity because the multivariate Gaussian is uniquely determined by its mean and covariance.

Assuming weak stationarity, we are interested in the autocovariance function

\[
R(s) = \text{cov}(X_t, X_{t+s}) = E[(X_t - \mu)(X_{t+s} - \mu)].
\]  

(3.1)

I will assume for now that we have the usual case of \( D_t = 0, 1, 2, \ldots, N - 1 \). If the
autocovariance has absolute summability

\[ \sum_{s=-\infty}^{\infty} |R(s)| < \infty, \tag{3.2} \]

then the spectral density function can be represented as

\[ S(f) = \sum_{s=-\infty}^{\infty} R(s) \exp(-2\pi i fs), \quad -\infty \leq f \leq \infty \tag{3.3} \]

which provides another approach to statistical inference via the spectral density or spectrum of the time series. In fact, the autocovariance and the spectrum are Fourier transform pairs, so that given the spectrum, the autocovariance can be represented as

\[ R(s) = \int_{-1/2}^{1/2} S(f) \exp(2\pi ifs) df. \tag{3.4} \]

This is the Einstein-Wiener-Khintchine theorem (Einstein, 1914; Wiener, 1930; Khintchine, 1934).

The spectrum has the following properties:

- \( S(f) \geq 0 \) (The spectrum is non-negative).
- \( S(f) = S(f + 1) \) (The spectrum is periodic and has period equal to one).
- \( S(f) = S(-f) \) (The spectrum is an even function for real-valued \( x \)).
- \( \sum_{s=-\infty}^{\infty} |R(s) \exp(-2\pi ifs)| < \infty. \)

Note that the presence of periodic components in the time series violates the assumption of absolute summability of the autocovariance function (equation 3.2). However,
in practice we never have a time series of infinite length, so the autocovariance function remains finite for a finite lag. Also, it is physically impossible for a periodic process in nature to exist forever, so the spectral representation of the covariance of natural processes is finite regardless.

The equivalence of the autocovariance $R(s)$ and the spectrum $S(f)$ means that either can be used for inference. However, the spectrum has an arguably simpler interpretation, and provides the opportunity to discover something unexpected (Tukey and Hamming, 1949). This second point is relevant in the context of analysis of paleoclimate series because of the plausibility that quasi-periodic signals are hidden in the time series. The spectrum is precisely tuned to detection of periodicities. Another strong point in favour of the spectrum compared to the autocovariance is that at suitably spaced frequencies, the distribution of spectrum estimates are or can be made approximately independent. This independence suggests that a product-likelihood be used in the Bayesian analysis of the spectrum in Chapter 4.

### 3.1.1 Direct spectrum estimators

Assume that we have a time series $x$, a realization of $X$, observed at a set of known and discrete times. Assume that the observation times are regularly spaced with $D_t = 0, 1, 2, \ldots, N − 1$. A naive estimator of the spectrum, known as the periodogram (Schuster, 1898), is

$$\hat{S}_P(f) = \frac{1}{N} \left| \sum_{t=0}^{N-1} x_t \exp(-2\pi i ft) \right|^2. \tag{3.5}$$

This estimator of the spectrum has some problems. It is inconsistent: The variance of the estimator does not tend to zero as $N$ goes to infinity. In fact, the variance of the estimator doesn’t decrease as the sample size increases. The spectral window
for the periodogram is Fejér’s kernel \( \mathcal{F}(f) = \frac{1}{N} \frac{\sin^2(N\pi f)}{\sin^2(\pi f)} \). This kernel has side lobes that contribute to the spectrum estimate at frequencies away from a given frequency. This effect is known as spectral leakage and can cause the periodogram to be badly biased (Thomson, 1982). For this reason, estimators that taper the data are preferred (Brillinger, 1981), as they have lower side lobes and reduced spectral leakage.

There are many data tapers that can be used to mitigate the spectral leakage problem (Harris, 1978). The zeroth-order discrete prolate spheroidal sequences (DPSS or Slepians) are sequences of length \( N \) that have the smallest possible spectral leakage outside of a defined frequency band of width \( 2W \). The tapered estimator of the spectrum, known as a direct estimate, is

\[
\hat{S}_D(f) = | \sum_{t=0}^{N-1} h_t x_t \exp(-2\pi i ft) |^2
\]  

(3.6)

where \( h_t \) is the data taper, which notably has the property that \( \sum_{t=0}^{N-1} h_t^2 = 1 \).

In the rest of this thesis, we will only be using the DPSS, or approximations thereof, as data tapers. The DPSS are usually denoted as \( v_t^{(k)}(N,W) \), where \( k \) is the order of the sequence. The higher-order DPSS can also have nearly minimal spectral leakage. In fact, the first \( K \approx 2NW \) have nearly minimal spectral leakage, so that

\[
\hat{S}_{k}^{(mt)}(f) = | \sum_{t=0}^{N-1} v_t^{(k)} x_t \exp(-2\pi i ft) |^2
\]  

(3.7)

for any \( k \approx 2NW \) will provide an estimate that is characteristically similar. Moreover, the estimates \( \hat{S}_{k}^{(mt)}(f) \) are approximately uncorrelated because the DPSS tapers are orthogonal (Thomson, 1982; Percival and Walden, 1993). This suggests averaging the \( \{ \hat{S}_{k}^{(mt)}(f) \} \) to not only control spectral leakage, but also to reduce the variance of the
combined estimate. The result is the multitaper spectrum estimate

$$\hat{S}^{(mt)}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{S}_k^{(mt)}(f),$$

(3.8)

where $K$ is the maximum order of DPSS taper chosen in forming the estimate. It is common to play it safe with respect to spectral leakage and use $K = [2NW] - 1$ or fewer DPSS tapers.

### 3.2 Irregular times

The methods of spectrum estimation mentioned in section 3.1.1 are for time series with observation times that are regular. The regular sampling case is the most common, particularly when observation times can be controlled when collecting the time series data. Methods for spectrum estimation under regular sampling can take advantage of the Fast Fourier Transform for computation, making computation extremely efficient.

There exists an irregular sampling version of the multitaper (Bronez, 1988), that involves computing generalized discrete prolate spheroidal sequences by finding the solution to a generalized matrix eigenvalue problem. However, in my experience the Bronez solution is extremely computationally expensive and numerically difficult to compute. These qualities make it ill-suited to the present problem in which a quick and reliable solution is necessary for MCMC sampling, for example. Others have developed solutions for otherwise regular series with gaps, (e.g. Fodor and Stark, 2000; Smith-Boughner and Constable, 2012) but these are not characteristic of the paleoclimate series examined in this thesis. Approaches that interpolate the autocovariance matrix using slotted estimators based on discretization of time lags are also possible.
(Stoica and Sandgren, 2006), but again would seem to be more suited to series that are otherwise regular or comprised of a small set of discrete time increments. A review of various parametric, non-parametric, and heuristic approaches is available in Babu and Stoica (2010).

Given the computational advantage afforded by the spectral estimators under regular sampling, a natural question is whether we can use methods of spectrum estimation that are designed for regularly sampled time series in cases where the times are irregularly sampled. A seemingly popular approach is to interpolate the irregularly sampled time series to a mesh of regular times, then proceed with spectrum estimation as though the interpolated time series were what was originally observed. However, the consequences of the interpolation on the spectrum estimate and the inferred characteristics of the underlying spectrum should be considered.

Lepage (2009) demonstrated that spurious peaks in the spectrum can occur as a result of interpolation of an irregularly sampled time series to regular spacing. Consider the following highly simplified example. Suppose we have a white noise process that follows

\[ X_t \sim N(0, 7^2) \]

for any \( t \). Draw a realization from this process \( x \) at observation times consistent with a simple age model for a paleoclimate core record, and in addition at observation times on a regular mesh for comparison. The simple age model is meant to be characteristic of “connect the dots” aging of core samples, and is piecewise linear (Figure 3.1). Time accumulates at the following rates in each piece:

- Eight per unit depth for 200 units of depth.
Figure 3.1: A simple age model to demonstrate the effect of interpolating an irregular series to a regular mesh before estimating the spectrum.

- Nine per unit depth for 300 units of depth.
- Seven per unit depth for 350 units of depth.
- Ten per unit depth for 150 units of depth.

The regular mesh had a time interval of five. Linear interpolation was used to interpolate the times from the piecewise linear time versus depth relationship to the regular mesh.

Figure 3.2 is a comparison of a multitaper spectrum estimate obtained by linear interpolation of the irregularly sampled series to the spectrum estimate that would have been obtained had the series been regularly sampled. The linear interpolation has had a noticeable effect on the overall shape of the spectrum. Although the underlying process is white, and should have resulted in a flat spectrum estimate, interpolation has reddened the estimate: high frequencies have a lower associated
3.2. IRREGULAR TIMES

Figure 3.2: A comparison of multitaper spectrum estimates for a white noise process sampled at a time interval of five (top panel), and the same realization sampled irregularly and interpolated to the same time interval of five (bottom panel), showing the effect of interpolation.

spectrum estimate, and low frequencies have a higher associated spectrum estimate.

Mann and Lees (1996) describe a popular method for estimation of background noise and signal detection in climatic time series. The method consists of partitioning the spectral representation of a series into a background autoregressive spectrum and a set of signals. The signals are then tested for significance by comparison to a null distribution. More details will be presented in Chapter 4, but here we might like to
know whether application of the Mann and Lees method would lead to false detection of signals in the interpolated case (as was seen in Lepage (2009)). Figure 3.3 shows the application of the Mann and Lees (1996) method to the interpolated spectrum estimate. Although it appears that the background spectrum does not do a good job of describing the spectrum estimate at higher frequencies, several low to moderate frequency signals are detected by the procedure.

Given that interpolation can have misleading effects on estimates of the spectrum, it seems that computational efficiency alone doesn’t justify interpolation to regular times. The Lomb-Scargle periodogram (Scargle, 1982) is widely used and considered to be a solution to the spectrum estimation problem for irregular sampling. However,
like the periodogram, it has statistical issues, namely:

1. It is inconsistent.

2. It has spectral leakage that is often worse than the periodogram.

3. It has spectral leakage that depends on the pattern of observation times and that is frequency-specific.

In Scargle (1982), the author is concerned primarily with the detection and identification of a single periodic component from a background noise process that is of no interest. Because of this narrow focus, Scargle argues that inconsistency and spectral leakage are not of particular concern, but that the estimator could be improved by applying a window function to reduce spectral leakage if desired. Here, I propose using interpolated discrete prolate spheroidal sequences as windows for the Lomb-Scargle periodogram, and averaging the resulting spectral statistics in the same way as for the regular-sampling multitaper. Although not theoretically optimal in the sense of Bronez (1988), the statistic can be made fast to compute using a non-equispaced Fast Fourier Transform (nfft). Furthermore, it has demonstrably low spectral leakage in the paleoclimate series from cores examined in this thesis.

The properties of the multitaper spectral likelihood for the evenly sampled case do not necessarily transfer to the irregular sampling case. In particular, the spectral window under irregular sampling varies as a function of the actual sample times and differs as a function of frequency (Scargle, 1982). In the next section I describe an approximate multitaper spectral statistic based on the \( \text{nfft} \), the Lomb-Scargle periodogram, and the DPSS tapers.
3.3 The multitaper Lomb-Scargle statistic

In this section I present the Lomb-Scargle periodogram and describe how it can be computed using a fast algorithm. Next, I define a multitaper version of the Lomb-Scargle periodogram and demonstrate that, for a set of times drawn from the posterior distribution of chronologies in Chapter 2, it has spectral windows that exhibit limited spectral leakage.

3.3.1 The Lomb-Scargle periodogram and fast computation

The Lomb-Scargle periodogram estimator (Scargle, 1982) is

\[
P_{LS}(f) = \frac{1}{2} \left( \frac{\left( \sum_{i=1}^{N} x_i \cos \omega (t_i - \tau_{LS}) \right)^2}{\sum_{i=1}^{N} \cos^2 \omega (t_i - \tau_{LS})} + \frac{\left( \sum_{i=1}^{N} x_i \sin \omega (t_i - \tau_{LS}) \right)^2}{\sum_{i=1}^{N} \sin^2 \omega (t_i - \tau_{LS})} \right)
\]  (3.9)

where \( \tau_{LS} \) is defined by

\[
\tan 2\omega \tau_{LS} = \frac{\sum_{i=1}^{N} \sin 2\omega t_i}{\sum_{i=1}^{N} \cos 2\omega t_i}
\]  (3.10)

and \( N \) is the length of the time series, \( \omega = 2\pi f \) is the angular frequency, \( x_i \) is the \( i^{th} \) observation, and \( t_i \) is the sampling time of the \( i^{th} \) observation. The Lomb-Scargle estimator has the following properties:

1. The estimator is invariant to time-translation. That is, if the \( t_i \) are replaced with \( t_i - t^* \) where \( t^* \) is a constant, the estimator is unchanged.

2. The estimator has the same distribution for a Gaussian white noise process as the regularly-sampled periodogram.
3. The estimator is equivalent to a least-squares fit of a sinusoid at a given frequency $f$ to the data $x$ in the time domain.

Leroy (2012) points out that the $nfft$ can be used to compute the Lomb-Scargle periodogram using the decomposition of Press and Rybicki (1989). Following Press and Rybicki, the Lomb-Scargle periodogram can be represented using the following sums:

$$s_h = \sum_{i=1}^{N} x_i \sin \omega t_i$$  
$$c_h = \sum_{i=1}^{N} x_i \cos \omega t_i$$  \hspace{1cm} (3.11)

$$s_2 = \sum_{i=1}^{N} \sin 2\omega t_i$$  
$$c_2 = \sum_{i=1}^{N} \cos 2\omega t_i$$  \hspace{1cm} (3.12)

If the sums in equations 3.11 and 3.12 can be computed rapidly, then the Lomb-Scargle periodogram can be computed rapidly as well using the following terms:

$$A_1 = c_h \cos \omega \tau_{LS} + s_h \sin \omega \tau_{LS}$$  
$$A_2 = \frac{N}{2} + \frac{1}{2} c_2 \cos 2\omega \tau_{LS} + \frac{1}{2} s_2 \sin 2\omega \tau_{LS}$$  \hspace{1cm} (3.13)

$$B_1 = s_h \cos \omega \tau_{LS} - c_h \sin \omega \tau_{LS}$$  
$$B_2 = \frac{N}{2} - \frac{1}{2} c_2 \cos 2\omega \tau_{LS} - \frac{1}{2} s_2 \sin 2\omega \tau_{LS},$$  \hspace{1cm} (3.14)

where $\cos 2\omega \tau_{LS}$ and $\sin 2\omega \tau_{LS}$ can be computed from equation 3.12, and $\cos \omega \tau_{LS}$ and $\sin \omega \tau_{LS}$ can be computed using half-angle formulae from $\cos 2\omega \tau_{LS}$ and $\sin 2\omega \tau_{LS}$. The Lomb-Scargle periodogram is thus expressed

$$P_{LS}(f) = \frac{1}{2} \left( \frac{A_1^2}{A_2} + \frac{B_1^2}{B_2} \right).$$  \hspace{1cm} (3.15)

To compute the sums in equations 3.11 and 3.12, I used the $nfft$ library (Keiner et al., 2009). The $nfft$ library provides a fast Fourier transform for irregularly sampled
data. It is important that the transform be speedy because ultimately we will be doing inference in situations where the times are not only irregular but unknown. In these situations, we will not have a single set of irregular sampling times, but many possible sets of irregular sampling times for a single time series analysis. Monte Carlo sampling of the posterior distribution of model parameters will require repeated sampling of timing model parameters in addition to time series model parameters, placing some pressure on the time required for evaluation at each Monte Carlo step.

The nfft library defines the nonequispaced discrete Fourier transform (NDFT) as

\[ x_{t_j^*} = \sum_{k=-L/2}^{L/2-1} \tilde{x}_k e^{-i2\pi t_j^* k} \]  

(3.16)

where the \( \tilde{x} \) are input complex Fourier coefficients, and \( x_{t_j^*} \) are the time series as observed at times \( t_j^* \), standardized so that the times are in the interval \([-1/2, 1/2]\).

Writing the NDFT in matrix notation, we have

\[ \mathbf{x} = \mathbf{A} \tilde{\mathbf{x}} \]  

(3.17)

\[ A_{jk} = e^{-i2\pi kt_j^*}. \]  

(3.18)

To make the nfft library useful for time series analysis, we need to consider the \( \mathbf{x} \) as an input and the Fourier representation as an output. The adjoint (conjugate transpose) of the matrix \( \mathbf{A} \) multiplied by the time series vector \( \mathbf{x} \) is

\[ \tilde{\mathbf{z}} = \mathbf{A}^H \mathbf{x} \]  

(3.19)
3.3. THE MULTITAPER LOMB-SCARGLE STATISTIC

which is

$$\tilde{z}_k = \sum_{j=0}^{N-1} x_t^* e^{i2\pi kj}. \quad (3.20)$$

Thus, the adjoint NDFT performs the necessary Fourier transform of the input time series \(x\), except that it seems to return the complex conjugate of the usual transform. In fact, because of the centering of the times to the interval \([-1/2, 1/2)\), there is a further rotation of \(k\pi\) due to the \(-1/2\) offset. This additional rotation requires a further transformation of \(\tilde{z}_k\)

$$\tilde{z}_k^* = \begin{cases} 
\overline{\tilde{z}_k} & k \text{ is even} \\
-\tilde{z}_k & k \text{ is odd}
\end{cases} \quad (3.21)$$

This subtlety is not mentioned by Leroy (2012).

3.3.2 Multitapering the Lomb-Scargle periodogram

As described by Scargle (1982), the Lomb-Scargle periodogram suffers from even worse spectral leakage than the regularly sampled periodogram. In the current application, this spectral leakage presents a problem because we would like the spectral statistics to be independent so that they can be included in a product likelihood for Bayesian inference. In the regularly sampled case, the multitaper statistic based on the DPSS limits the spectral leakage to a negligible amount, so that multitaper spectral statistics spaced far enough apart in frequency are nearly uncorrelated. Scargle (1982) suggests that the situation in the irregular sampling case can also be improved through use of data tapering.

My approach for multitapering using the DPSS for irregular time series is to first
compute the discrete prolate spheroidal sequences at a mesh corresponding to an average sampling interval $\Delta t = T/N$, where $T$ is $t_{N-1} - t_0$ and $N$ is the number of times in the series. The DPSS on the regular mesh are denoted as before: $v^{(k)}_t(N,W)$, where $k$ is the order of the sequence. The taper weights at intermediate points corresponding to the irregular times are then obtained by interpolation using a cubic spline. Denote these interpolated weights as $v^{*(k)}_t(N,W)$. This approximation is best when the set of times is close to regular – for example, there are no long gaps – which we assume is true for climatic time series of the kind considered here. Finally, the interpolated weights for each DPSS order are renormalized so that $\sum_{i=t_1}^{t_N} \left( v^{*(k)}_i(N,W) \right)^2 = 1$ for each $k$. The resulting multitaper statistics are then used as in the regular sampling case described previously.

The multitaper Lomb-Scargle statistic (mtLS) is a compromise between the optimal quadratic estimator of Bronez (1988), which is difficult and costly to compute, and the un-tapered Lomb-Scargle, which can be made fast to compute but does not achieve minimum spectral leakage out-of-band. Because the spectral leakage performance of the mtLS depends on the observation times $D_t$ as well as the frequency band, the spectral pseudo-windows should be examined across a range of frequencies for a given set of irregular times for which the statistic is to be computed.

To close out this section, I will revisit the earlier simplified example of an irregularly sampled white noise process, interpolated to regular sampling (Section 3.2). We saw that interpolation to regular sampling induced an overall shape to the spectrum estimate whereas the spectrum estimate would have been flat had the process been regularly sampled. The mtLS spectrum estimate is able to overcome the irregular sampling without introducing any shape or other spurious features (Figure 3.4). In
The multitaper spectrum estimate for a white noise process sampled irregularly. The spectrum estimate is comparable to the estimate that would have been obtained under regular sampling (Figure 3.2, top).

the next section, I examine spectral pseudowindows in the case of observation times drawn from the posterior distribution of chronologies obtained in Chapter 2.

3.4 Application to paleoclimate time series

In this section, I compute the mtLS statistics for the Tibetan peat core data that was first introduced in Chapter 2. I begin by examining some pseudowindows for both the mtLS statistics as well as for the non-tapered Lomb-Scargle statistics to demonstrate the differences between the two. The mtLS statistic shows less leakage compared to the non-tapered version.

Checking pseudowindows

In the regular sampling case, the spectral leakage that results from tapering with the low-order DPSS is minimized. When using the multitaper Lomb-Scargle for irregular
times, the DPSS are interpolated, and the resulting spectral leakage depends on both the particular set of irregular times and the frequency band. Although the spectral leakage properties of the mtLS should be close to optimal for nearly regular sampling, it is a good idea to check the pseudowindows for a range of bands to avoid errors in interpretation of the mtLS statistics. Scargle (1982) examines pseudowindows and shows that in some cases, spectral leakage can be significant even at remote frequencies. This suggests that checking pseudowindows is good practice in all cases of spectrum estimation under irregular sampling, whether using the mtLS, Lomb-Scargle, or the method of Bronez.

In order to check pseudowindows, it suffices to create an artificial signal at a given frequency in-band, sample it at the same times as the original irregular series, and examine the resulting spectral statistics. This reveals the pattern of spectral leakage out of band. The pseudowindows for one thousand posterior chronologies from the Tibetan peat core (Yu et al., 2006), from Chapter 2, for frequencies of 0.01, 0.03, and 0.05 cycles per year show mild out-of-band leakage to remote frequencies (Figure 3.5). In particular, there is some leakage to the zero frequency. For comparison, the same pseudowindows for the Lomb-Scargle statistic are plotted in Figure 3.6.

A close-up comparison of the pseudowindows (Figure 3.7) reveals differences between the non-tapered Lomb-Scargle and the mtLS. The Lomb-Scargle has much higher variance, although there are frequencies out-of-band that have very low response, there are also frequencies that show noticeably higher spectral leakage compared to the mtLS.
3.4. APPLICATION TO PALEOCLIMATE TIME SERIES

Figure 3.5: Three mtLS pseudowindows for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year. The parameters of the underlying DPSS were $N = 649$, $W = 0.0046225$, and $k = 5$. 
Figure 3.6: Three Lomb-Scargle (non-tapered) pseudowindows for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year.
Figure 3.7: A comparison of the pseudowindows for the mtLS (in blue) and non-tapered Lomb-Scargle (in gray) for the Tibetan peat core, for chronologies drawn from the posterior distribution of chronologies (Figure 2.2). The pseudowindow for each chronology is plotted, for frequencies of 0.01 (top), 0.03 (middle), and 0.05 (bottom) cycles per year. The parameters of the underlying DPSS were $N = 649$, $W = 0.0046225$, and $k = 5$. 
Multitaper Lomb-Scargle spectral statistics

I computed the mtLS spectral statistics for the Tibetan peat core proxy data, given each of one thousand posterior chronology samples from the chronology analysis in Chapter 2 (Figure 3.8). There were three proxies collected at each layer in the core; Greyscale and Humification are believed to be related to the strength of the summer monsoon, whereas Ash content is believed to be related to the strength of the winter monsoon (Yu et al., 2006).

I subtracted the mean of each proxy prior to computing the spectral statistics but did not detrend. As a result, all three proxies had a large amount of variance at the lowest frequency bin. Besides the first frequency bin, the statistics of the three proxies only appear to show distributions different than the background noise shape at frequencies smaller than about 0.0028 cycles per year for Greyscale, and 0.0018 cycles per year for Humification and Ash content (Figure 3.8). Whether these apparent differences from the overall shape represent signals will be explored in Chapter 4. Frequencies larger than about 0.0104 cycles per year did not show any noticeable features suggestive of anything but background noise.

3.5 Application in cases of irregular but known times

The mtLS spectral statistic will form the basis of a product likelihood for a Bayesian analysis of the spectrum in Chapter 4. It is hoped that the mtLS could also be used as an approximate direct estimator of the spectrum in cases where the observation times are known but irregular. This situation arises often in astronomical observations, and astronomers are avid users of the Lomb-Scargle spectrum estimator.

Like climate science, astronomy is a field in which periodic processes play an
3.5. APPLICATION IN CASES OF IRREGULAR BUT KNOWN TIMES

Figure 3.8: mtLS spectral statistics for the Tibetan peat core data: Greyscale (top panel), Humification (middle panel), and Ash Content (bottom panel). The frequencies are jittered, and are in units of cycles per year. Each point represents chronology drawn from the posterior distribution of chronologies (Figure 2.2).
important role. For example, helioseismology studies solar oscillations as a way of inferring the inner workings of the Sun (Christensen-Dalsgaard, 2002). These inner workings include things such as structure, rotation, composition, dynamics and evolution. Asteroseismology is the equivalent study of stars other than the Sun (Aerts et al., 2010). The Kepler spacecraft has a highly sensitive photometer which can measure the dimming caused by planets transiting in front of a star. Using measurements of total flux, the astronomer can not only detect transiting planets, but also unravel stellar properties such as whether red giant stars are burning helium in their cores (Bedding et al., 2011). This inference is based on the spacing of line components in the spectrum of the time series of flux, which relies on accurate identification of signals from noise in the spectrum estimate. Another method of inference involves measurement of the line-of-sight velocity using Doppler shifts in the light spectrum of a star. Although these measurements can be taken using ground-based telescopes, the exposure times are long and as a result sample sizes can be limited (e.g. Lillo-Box et al., 2014a).

In this section, I compute the mtLS for a time series of relative flux measurements measured using the Kepler spacecraft for the red giant star Kepler-91 (Figure 3.9). The star has a transiting planet Kepler-91b with an orbital period of approximately 6.25 days (Lillo-Box et al., 2014b), in addition to other periodic variations in flux indicative of stellar processes. The data are sampled irregularly due to the orbit of the spacecraft around the Sun. There are also some longer gaps in the series (Figure 3.10). For this preliminary analysis, the series was analyzed in its entirety, across these longer gaps.
3.5. APPLICATION IN CASES OF IRREGULAR BUT KNOWN TIMES

Figure 3.9: Relative flux data for the red giant star Kepler-91.

Multitaper Lomb-Scargle estimates

I obtained estimates of the spectrum of relative flux from the data in Figure 3.9 using three frequency bandwidth values (corresponding to NW = 4, 10, or 20) and the mtLS, and also using the Lomb-Scargle periodogram for comparison (Figure 3.11). The increasing number of tapers with increasing bandwidth has a noticeable effect on the variance of the estimate. In the extreme case of the Lomb-Scargle, signals appear to be less obvious than for even the narrow-bandwidth mtLS. Comparing the three mtLS spectrum estimates, the one with a time-bandwidth parameter of 10 appears to retain the relative amplitude of any signals present while providing a good separation relative to the noise in the estimate. Although there has been some work done on the choice of bandwidth parameter under some assumptions about the underlying true spectrum (Haley and Anitescu, 2017), there is no universally applicable method for
Figure 3.10: Distribution of first time differences $\Delta t$ for the red giant flux data in Figure 3.9. Most data are sampled nearly regularly, but there are also some longer gaps.
determining the optimal choice for this parameter.

The mtLS estimate of the spectrum appears to confirm previous estimates of the orbital period of the transiting planet Kepler-91b (Figure 3.12b). In particular, there is a peak in the spectrum estimate at a period that matches previous estimates, as well as a series of harmonics, which is to be expected due to the shape of the light curve attenuation as the planet passes in front of the star. There are also a number of apparent signals or line components in the frequency band from eight to twelve cycles per day, approximately 92-140 µHz (Figure 3.12c). These are presumably due to stellar processes of interest to asteroseismologists. The high frequency portion of the mtLS estimate (Figure 3.12d) should not be trusted, for reasons explained in the next paragraph.

Checking pseudowindows

The apparently high spectrum values at high frequency are potentially due to leakage from elsewhere in the spectrum. To investigate this possibility, I generated synthetic signals across the frequency range and calculated the response of the mtLS statistic for each. The frequency response shows a pattern of aliasing beyond about ten cycles per day (Figure 3.13). This pattern was verified by computing the Lomb-Scargle periodogram of the same synthetic signals using independent computer code (Ruf, 1999). The aliasing effect is a result of the particular pattern of times.

3.6 Discussion

Unlike many branches of statistics in which observations are assumed to be realizations of independent or conditionally independent random variables, time series
Figure 3.11: Four spectrum estimates using the Kepler-91 flux data from Figure 3.9. The top panel is the usual Lomb-Scargle periodogram. The second panel is the mtLS spectrum estimate with NW=4 and K=7 tapers. The third panel is the mtLS spectrum estimate with NW=10 and K=19 tapers. The bottom panel is the mtLS spectrum estimate with NW=20 and K=39 tapers. The progression shows the tradeoff between frequency resolution and estimator variance.
3.6. DISCUSSION

Figure 3.12: The mtLS spectrum estimate with NW=10 and K=19 tapers. Panel (a) is the spectrum estimate across all frequencies. Panel (b) is the low-frequency portion of the spectrum; lines show the fundamental frequency and harmonics of the previously identified transiting planet. Panel (c) is the band from 8-12 cycles per day, showing many apparent signals. Panel (d) is the highest-frequency portion of the spectrum. This portion of the spectrum should be interpreted with caution due to the possibility of leakage or aliasing from distant parts of the spectrum.
Figure 3.13: Frequency response of the mtLS for the Kepler-91 time series. Frequencies above about 10 cycles per day show pseudo-aliasing to higher frequencies.

Analysis is preoccupied with the dependence of sequences of random variables. Unfortunately for time series analysts, the random variables underlying the observations are fleeting, and it is impossible to obtain repeated measures of a random variable corresponding to a moment in time. Instead, methods of time series analysis make simplifying assumptions about the nature of the underlying random variables and their covariance structure. The assumption of stationarity, for example, allows measurement to begin at any point in time, because if the assumption holds the joint distribution of random variables depends only on their relative position in time. The stationarity assumption allows the covariance to be summarized as a function of time.
3.6. DISCUSSION

The autocovariance can be obtained from the spectrum if desired, meaning that spectral statistics or spectrum estimates can be used instead of estimates of the autocovariance. Inference in the spectral domain is convenient because many natural processes are approximately periodic, and show up as signals in the spectrum of a time series. Moreover, spectral statistics can be made to be approximately independent using data tapers, so that the total likelihood of a set of spectral statistics can be approximated using a product likelihood (Chapter 4).

For irregularly sampled time series, the spectrum estimate of Scargle (1982) is widely used. Scargle was motivated by the case of a single periodic component embedded in a white noise process, and for this type of process the Lomb-Scargle estimator has its place. However, the spectral leakage properties of the estimator can be extremely poor, and if the underlying spectrum is not of the type envisioned by Scargle, then it is possible to be misled by the result. Scargle suggests tapering the time series to control spectral leakage before computing his periodogram, a suggestion that I take to the same conclusion as Thomson (1982) by multitapering using interpolated DPSS tapers.

The mtLS statistic matches the Thomson multitaper under regular sampling. Deviations from regular sampling introduce frequency-specific differences that depend on the particular set of observation times. For a given set of times, careful examination of the pseudowindows is recommended. If there appear to be strong signals or a large dynamic range in the spectrum, frequencies with high associated spectral statistics should be examined for their leakage properties because they are liable to contribute to spectral statistics at other frequencies in the spectrum. Unlike in the
case of regular sampling, leakage (or aliasing) to distant frequencies is possible for
irregular sampling. In particular, high-frequency parts of the spectrum may not be
trustworthy and, if judged to be suspect, should be ignored.

In the example of the red giant star data, there appear to be many signals present,
and the overall shape of the spectrum at high frequencies seems to be affected by a
kind of aliasing. This limits the effective Nyquist frequency to something much smaller
than $1 / 2 \Delta t = N / 2T$. One possibility is that the aliasing is due to regular gaps in
the series. If this is the case, an mtLS spectrum estimate could be computed for each
contiguous portion of the series, and averaged as in a non-overlapped Welch estimate
(Welch, 1967). Although this doesn’t resolve the higher frequency components of the
spectrum estimate, if the aliasing is removed it makes interpretation of the estimate
more straightforward. Regardless, the technique of separating time series according
to the location of long gaps could be helpful in certain cases where the long gaps
make the high frequency portion of the spectrum estimate unreliable.

There are obvious extensions to the current mtLS as an estimator of the spectrum,
based on analogies to the use of the multitaper estimator in the regular sampling case.
The eigencoefﬁcients (Thomson, 1982) are the complex-valued result of taking the
discrete Fourier transform of the DPSS-windowed time series. The eigencoefﬁcients
are the basis of several extensions to the multitaper method in the regular sampling
case. The Thomson F-test (Thomson, 1982) is a test for the presence of periodic
components that is based on a regression estimate of the eigencoefﬁcients. The variance
explained by the regression is compared to the residual variance of the estimate
to form the F-test. In principle a similar test could be constructed for the mtLS, but
this might require and adjustment to the method of computation described in section
3.3.1. A targeted approach in which frequencies of interest are tested for periodicity might be the most efficient if the fast computing advantage is lost. Coherence, the frequency-domain analogue of correlation, would also be a useful tool in cases where an estimate of the relationship between two or more time series is of interest. Jackknife error estimates for the spectrum estimate, periodic F-test, and coherence are also possible, and would provide insight into the uncertainties and biases associated with each (Thomson, 2007).

3.7 Conclusion

In this chapter, I described a multitapered version of the popular Lomb-Scargle estimator of the spectral density of an irregularly sampled time series. The mtLS was shown to have limited spectral leakage, but because the leakage depends on the particular pattern of observation times, examination of the pseudowindows is necessary in any particular application. As an estimator, the mtLS was applied to a time series of photometry data and displayed the reduced variance and spectral leakage characteristic of the multitaper estimator for regularly spaced data. However, examination of the frequency response (pseudowindows) revealed aliasing and a much smaller effective Nyquist frequency than might have been expected. The mtLS pseudowindows for the posterior chronology of the Tibetan peat core from Chapter 2 had no such aliasing, and had limited spectral leakage. Thus, the mtLS spectral statistic will be modelled using a product likelihood in the Bayesian analysis of the paleoclimate spectra in Chapter 4.
Chapter 4

Bayesian Spectral Inference for Paleoclimate Time Series

Mann and Lees (1996) describe a popular method for estimation of background noise and signal detection in climate time series. As of this writing, the paper has been cited 817 times according to Google Scholar. The method, mentioned in Chapter 3, consists of partitioning the spectral representation of a series into a “red” noise background and a set of signals. The signals themselves are not modelled parametrically – rather, they are treated non-parametrically and their significance is assessed by comparison to percentiles of the red noise distribution. The trick is to obtain an estimate of the red noise spectrum that is robust to the presence of signals, in order to properly separate signals from noise and their effects on the shape of the spectrum. Mann and Lees argue that an autoregressive model of order 1 is justified based on: (1) the physics of the climate system; (2) interpretability of the characteristic decay time; and (3) a desire to avoid interpreting signals as noise, which the authors argue can become problematic for higher-order autoregressive models.
4.1. A BAYESIAN MODEL FOR THE SPECTRA OF CLIMATE PROCESSES

Like Mann and Lees, my goal here is to separate signal from (red) noise. However, I will treat both the signal and the noise as parametric, obtaining posterior distributions for model parameters in a Bayesian paradigm. The likelihood will be evaluated in the spectral domain using the multitaper spectral statistic from Chapter 3. The multitaper statistic is desirable in this scenario because it limits the amount of out-of-band bias at a given frequency, at the expense of reduced frequency resolution. This means that, for adequately spaced frequencies, it is not unreasonable to treat multitaper statistics as independent and compute a product likelihood. The independence makes computing the spectral likelihood convenient compared to the time-domain likelihood, in which we assume a dependence structure for observations made at different times.

4.1 A Bayesian model for the spectra of climate processes

In this section I present a model for the spectrum of a climate process that is in the spirit of Mann and Lees (1996) and to a large extent Thomson et al. (2001). In particular, the spectrum is modeled as a background autoregressive noise process with additional signal components. The spectral likelihood is based on a mixture of central (noise) and non-central (signal) $\chi^2$ distributions, with the spectral statistics based on the multitaper Lomb-Scargle statistic presented in Chapter 3. A spike-and-slab prior (Ishwaran and Rao, 2005) is used to identify the probability that the spectrum has a signal at each frequency.
4.1.1 The spectral likelihood

In Mann and Lees, the authors assume that the ratio of the multitaper statistic to the parametric noise spectrum is distributed as a *standardized* chi-squared distribution with $2K$ degrees of freedom. By standardized, I mean that if a random variable $X$ follows a $\chi^2_{2K}$ distribution, then the ratio of the multitaper statistic to the parametric model spectrum is a random variable $Y = X/2K$. This assumption is credited to Tukey (1949), Percival and Walden (1993), and Gilman et al. (1963). For the spectral likelihood, a collection of random variables $2KY_f \sim \chi^2_{2K}$ at adequately spaced ($\sim 2W$) frequencies will be approximately independent, so that the total spectral likelihood can be taken as the product of the likelihoods at each frequency. This independence arises due to the independence of the eigencoefficients in the multitaper statistic, which are (approximately) independent due to being (approximately) uncorrelated complex Gaussian random variables. Thus, because the multitaper statistic is a function of the independent eigencoefficients, it too is independent for sufficiently spaced frequencies.

For a white noise process, the spectral likelihood is described by the product of standardized chi-squared random variables as in Mann and Lees (1996). If there is a signal at a given frequency, however, the spectral likelihood will be described by a standardized noncentral chi-squared random variable. Because the chi-squared distribution is a special case of the noncentral chi-squared distribution, I will take the spectral likelihood to be a product of noncentral chi-squared random variables, and place special meaning on the case where the noncentrality parameter is zero – the no-signal or ordinary chi-squared case.
4.1. A BAYESIAN MODEL FOR THE SPECTRA OF CLIMATE PROCESSES

4.1.2 The mean value function

All time series considered here are centered prior to analysis so that they have zero mean. Otherwise, I don’t include a model for the mean value function of the time series here. There is a rich literature dealing with models for mean value functions in time series, including trends, changepoints, and so on. However, much can be learned by graphical display of the time series data, and this has been a popular approach in the paleoclimate literature. What is less evident is whether “wiggles” in the time series represent periodic signals. Answering this question is what I focus on here.

4.1.3 The (red) noise spectrum

For simplicity, I restrict the class of possible parametric noise spectra to be autoregressive (AR) and low-order. An autoregressive process of order $p$ is denoted $AR(p)$, and its spectral distribution function is

$$S(f) = \frac{\sigma_p^2}{|1 - \sum_{j=1}^{p} \phi_j e^{-i2\pi fj}|^2}$$  \hspace{1cm} (4.1)

(Percival and Walden, 1993, eq. 392b). There are $p$ autoregressive parameters $\phi$ and the variance parameter $\sigma_p^2$, for a total of $p + 1$ parameters.

We require a prior distribution for each of the $p + 1$ autoregressive parameters. It might be desirable to formulate a prior for the $\phi$ that restricts the process to be stationary. In order for a value of $\phi$ to imply a stationary $AR(p)$ process, the roots of the polynomial equation $1 - \sum_{j=1}^{p} \phi_j z^j = 0$ must lie outside the unit circle (see for example Percival and Walden, 1993, pp. 392). Here, I will restrict $p = 1$. For $p = 1$, the condition $|\phi_1| < 1$ results in a stationary autogressive process. For
a red noise process with \( p = 1 \), the condition \( 0 \leq \phi_1 < 1 \) results in a stationary autoregressive process. For \( p = 2 \), an autoregressive process is stationary if and only if three conditions are satisfied (Cryer and Chan, 2008):

1. \( \phi_1 + \phi_2 < 1 \)
2. \( \phi_2 - \phi_1 < 1 \)
3. \( |\phi_2| < 1 \)

Extension to values of \( p > 1 \) is beyond the scope of this work, and will require careful consideration when constructing prior probability distributions for \( \phi \). For example, Tanaka and Komaki (2008) present a non-informative prior for the parameters of an AR(2) process. In practice, the process in question might not be stationary, and it might not be appropriate to place such a restriction on the parametric noise spectrum.

For an AR(1) process, the autocorrelation function decays exponentially as a function of lag. The characteristic decay time \( t \) is the average lifetime of a system undergoing exponential decay. In the case of an AR(1) process, it is related to the autoregressive parameter \( \phi_1 \) by

\[
t = \frac{1}{\log \phi_1},
\]

and provides an alternative way to characterize the AR(1) process.\(^1\)

4.1.4 The signals

In the absence of a signal, the multitaper statistics at each frequency spaced at a distance of \( 2W \) are assumed to follow a chi-squared distribution with \( 2K \) degrees of freedom.

\(^1\)Mann and Lees (1996) say that: “For periodicities much larger than \( t \), the spectrum behaves like a white spectrum.” It is unclear what is meant by this statement.
4.1. A BAYESIAN MODEL FOR THE SPECTRA OF CLIMATE PROCESSES

freedom. If there is a signal, the multitaper statistic will no longer follow a central chi-squared distribution; rather, it will follow a noncentral chi-squared distribution with non-centrality parameter \( \xi_i^2 \) (Thomson et al., 2001). Thus, I take the spectral likelihood to be noncentral chi-squared distributed, since the case where \( \xi_i^2 = 0 \) corresponds to the ordinary chi-squared distribution (lack of signal). Each of these \( \xi_i^2 \) parameters requires a prior distribution.

Suppose that at each frequency spaced at a distance of \( 2W \) there are two possibilities:

1. There is no signal at the frequency. The value of \( \xi_i^2 = 0 \).

2. There is a signal at the frequency. The value of \( \xi_i^2 > 0 \).

This suggests an additional set of parameters \( \varrho_i \) that codes whether there is a signal at a given frequency. The idea is to assume that, conditional on the value of \( \varrho_i \), \( \xi_i^2 \) follows one of two prior distributions:

- \( \varrho_i = 0 \to \xi_i^2 = 0 \)
- \( \varrho_i = 1 \to \xi_i^2 \sim f_{\xi^2 | \varrho=1}(\zeta) \), where \( \zeta \) is a vector of hyperparameters.

This concept has been described as a spike and slab prior in the context of variable selection in linear models (Mitchell and Beauchamp, 1988). For a linear model with coefficients \( \beta^{SS} \), and in which it is expected that a subset of the coefficients are effectively zero, the spike and slab prior can be used to shrink some coefficients towards zero while having little or no influence on the remaining non-zero coefficients. Ishwaran and Rao (2005) develop a framework for spike and slab priors that I modify slightly here. In particular, consider the formulation of Ishwaran and Rao in which the
4.1. A BAYESIAN MODEL FOR THE SPECTRA OF CLIMATE PROCESSES

Figure 4.1: The prior hierarchy for linear model coefficients from Ishwaran and Rao (2005), equation 4 therein.

\[
P(\beta_{SS} | \mathcal{J}_j, \tau_j^2) \overset{\text{ind}}{\sim} N(0, \mathcal{J}_j \tau_j^2) \\
P(\mathcal{J}_j | v_0, w^{SS}) \overset{\text{iid}}{\sim} (1 - w^{SS}) \delta_{v_0}(\cdot) + w^{SS} \delta_1(\cdot) \\
P(\tau_j^{-2} | a_1, a_2) \overset{\text{iid}}{\sim} \text{Gamma}(a_1, a_2) \\
P(w^{SS}) \sim \text{Uniform}[0, 1]
\]

authors write the prior hierarchy for \( \beta^{SS} \) as in Figure 4.1. In the prior hierarchy, \( \delta_{v_0} \) and \( \delta_1 \) are discrete probability measures at \( v_0 \) (a small value) and 1, respectively, and the symbols \( \overset{\text{ind}}{\sim} \) and \( \overset{\text{iid}}{\sim} \) mean independently distributed and independent and identically distributed, respectively.

The key features of the prior hierarchy of Ishwaran and Rao are:

1. Separation of coefficients into those near zero (spike) and those which are non-zero (slab).

2. A continuous bimodal distribution on the hypervariance \( \mathcal{J}_j \tau_j^2 \) which helps when sampling the posterior.

3. The \( w^{SS} \) parameter which can be interpreted as the proportion of non-zero coefficients.

Adapting the spike and slab approach to the current problem, we are interested in a prior hierarchy for the non-centrality parameters \( \xi^2 \). I specify the following prior
hierarchy:

\[ P(\xi_i^2|\varrho_i, \tau_i^2) \sim N_{1/2}((1 - \varrho_i)v_0\tau_i^2 + \varrho_i\tau_i^2) \]
\[ P(\tau_i^{-2}|\alpha_\tau, \beta_\tau) \sim \text{Gamma}(\alpha_\tau, \beta_\tau) \]
\[ P(\varrho_i|p_\varrho) \sim \text{Bern}(p_\varrho) \]
\[ P(p_\varrho) \sim \text{Beta}(1, 1), \]

where \( N_{1/2} \) is the half-normal distribution, \( \sim \), and \( \sim_{\text{ind}} \) are defined as before, and \( \alpha_\tau \) and \( \beta_\tau \) are specified a priori. The prior hierarchy describes a mixture model, and the \( \varrho \) parameters are included to code for each member of the mixture. This aids both sampling using Markov chain Monte Carlo methods and also interpretation of the posterior in terms of identifying which frequencies have signals.

Using this prior hierarchy, the posterior distribution of \( \varrho \) parameters can be used to compute the probability of a signal at each frequency, and in combination with the posterior distribution of the \( \xi^2 \) parameters, the distribution of the magnitude of the non-zero signal at each frequency. The parameter \( p_\varrho \) is the proportion of frequencies having a signal.

In the absence of informative prior information, a natural choice of prior for \( \varrho \) is independent and identically distributed Bernoulli distributions. If there is some prior information available, for instance previous indication of signals at certain frequencies, it can be included by specification of the Bernoulli parameter \( p_\varrho \) for those frequencies, or by specification of an informative prior for \( p_\varrho \). If there is no prior information, the set of Bernoulli parameters could be given a common prior, for example a weakly informative beta distribution (as in equation 4.3). Similarly, we could choose \( P(\xi^2|\varrho = 1) \) to represent prior knowledge of the magnitude of signals typically present in climate
series.

4.1.5 Sampling from the posterior distribution

The objective of this section is to compute the posterior distribution for spectral model parameters $\theta$ given the time series $x$ and times $T$, which may be observed or may be only known up to a distribution. In the case where the times are observed, the posterior distribution in question is

$$P(\theta|x,T) = P(\theta|S(x,T)) \propto P(S(x,T)|\theta)P(\theta),$$

(4.4)

where $S(x,T)$ are the spectral statistics.

In Chapter 2, I described a method for obtaining samples from the posterior distribution of possible chronologies based on data other than the time series $x$. This results in a posterior distribution for $T$ from which we can draw samples. The marginal posterior distribution for $\theta$ can be obtained by marginalizing the joint distribution of $\theta$ and $T$ over $T$:

$$P(\theta|x) = \int P(\theta,T|x)dT$$

$$= \int P(\theta|T)P(T|x)dT$$

(4.5)

$$= \int P(\theta|S(x,T))P(T)dT$$

Rather than evaluating the integral in equation 4.5 directly to obtain the target posterior distribution for $\theta$, the posterior distribution is computed by sampling first from the marginal distribution for $T$, then from the conditional posterior distribution
for $\theta$ (equation 4.4) given the sampled value for $T$ (Gelman et al., 2013).

For the specified model, the posterior can be written as

$$
P(\sigma_p^2, \phi_1, \xi^2, \varrho, \tau^2, p_\varrho | S(x, T)) \propto P(S(x, T) | \sigma_p^2, \phi_1, \xi^2, \varrho, \tau^2, p_\varrho) P(\sigma_p^2, \phi_1, \xi^2, \varrho, \tau^2, p_\varrho) \quad (4.6)
$$

$$
= P(S(x, T) | \sigma_p^2, \phi_1, \xi^2) P(\sigma_p^2, \phi_1) P(\xi^2 | \varrho, \tau^2) P(\varrho | p_\varrho) P(\tau^2) P(p_\varrho).
$$

Due to the hierarchical structure of the prior, I obtained samples from the posterior distribution using a Metropolis-within-Gibbs approach (Carlin and Louis, 2008; Gelman et al., 2003). In a Gibbs sampler (Geman and Geman, 1984), draws are made in turn from the conditional distribution of a subvector of the parameter vector given the remainder of the parameters. The resulting Markov Chain approximates draws from the target joint distribution of parameters. If a conditional distribution is not available, so that a Gibbs step is not possible, an alternative is to draw samples using a slice sampler (Neal, 2003), or according to the Metropolis algorithm (Metropolis et al., 1953).

Working through the conditional distributions for each parameter in equation 4.6:

- The $\sigma_p^2$ and $\phi_1$ parameters are updated together with a Metropolis step. The prior for $\sigma_p^2$ was chosen to be Uniform$(\sigma_{p,L}, \sigma_{p,U})$, where $\sigma_{p,L}$ and $\sigma_{p,U}$ are fixed, and the prior for $\phi_1$ was chosen to be Uniform$(0, 1)$.

- The $\xi^2$ parameters are updated using a slice sampler. The prior is as specified in equation 4.3.

- The $\varrho$ parameters are updated using a Gibbs step. The conditional posterior distribution for $\varrho_i$ is a Bernouilli draw in which $P(\varrho_i = 1) = \frac{p_i p_\varrho}{p_i p_\varrho + q_i (1 - p_\varrho)}$, where

$$
p_i = \frac{1}{1 + \sum_j (1 - p_j)}.
$$
where \( p_i = N_{1/2}(\xi_i^2, \tau^2) \) and \( q_i = N_{1/2}(\xi_i^2, v_0 \tau^2) \).

- The \( \tau^2 \) parameters are updated using a Metropolis step. The independent and identically distributed prior is as in equation 4.3.

- The \( p_\theta \) parameter is updated using a Gibbs step.

Each subvector of parameters is sampled in turn to obtain samples from the posterior conditional on \( T \). A number of samples are obtained for a given draw from \( P(T) \), and this process is repeated for subsequent draws from \( P(T) \), keeping the number of samples the same for each \( T \) drawn. All samples are then combined to obtain the marginal posterior for \( \theta \) as in equation 4.5.

### 4.2 Analysis of paleoclimate series

In this section, I apply the Bayesian model described in Section 4.1 to the paleoclimate series from Chapter 2, specifically the Tibetan peat core data and the Canadian prairie lakes data. Samples from the posterior distribution are used to estimate the posterior probability of a signal within each frequency band defined by the mtLS statistics, with the non-centrality parameters providing an indication of the magnitude of signals if they exist. Examinations of mtLS statistics at a frequency resolution that is higher than the DPSS bandwidth are used to suggest which frequencies in-band might contribute to the overall signal.
4.2. ANALYSIS OF PALEOCLIMATE SERIES

4.2.1 Tibetan peat series

The Tibetan peat series (Yu et al., 2006) are three climate proxy time series taken from a single core. The three series share a common chronology, whose posterior distribution was obtained in Chapter 2. The core was sectioned into 649 one-centimeter intervals and measurements were made on each section. These were:

- **Greyscale**: A measure of the visible darkness of the layer, thought to be related to the strength of the summer monsoon in the region.

- **Humification**: A measure of the amount of decomposition of the layer, thought to also be related to the strength of the summer monsoon.

- **Ash Content**: The amount of ash deposited in the layer, thought to be related to the strength of the winter monsoon.

**Specification of priors**

The Bayesian spectral model in Section 4.1 requires specification of some prior parameters. In all cases, priors were chosen to be at most weakly informative. For the priors on the variance of the non-centrality parameters, I followed Ishwaran and Rao (2005) and specified: $\alpha_r = 5$, $\beta_r = 50$, and $\nu_0 = 0.005$. For the priors on the noise spectrum variance, I specified the bounds on the uniform priors so that they did not restrict the posteriors:

- **Greyscale**: $\sigma_{p,L} = 0.1$ and $\sigma_{p,U} = 100$

- **Humification**: $\sigma_{p,L} = 10^{-9}$ and $\sigma_{p,U} = 1$

- **Ash Content**: $\sigma_{p,L} = 0.1$ and $\sigma_{p,U} = 1000$
4.2. ANALYSIS OF PALEOCLIMATE SERIES

Posterior distributions

Samples were obtained from the posterior distribution of model parameters using the methodology in Section 4.1.5. This involved alternating draws from the posterior distribution of chronologies and the conditional posterior distribution of spectral parameters. Specifically, three Markov chains were used, with starting parameter values drawn from prior distributions. First, a chronology was drawn. Following an initial burn-in of 10,000 samples, 500 samples were drawn from the posterior distribution of spectral parameters for each chain and thinned to 50 samples (150 samples total). Then, a new chronology was drawn from the posterior distribution of chronologies, and a further 100 samples were burned before drawing another 500 samples thinned to 50. This process of sampling the chronology and then sampling the spectral parameters was repeated one hundred additional times, for a total of 15,150 samples.

The non-centrality parameters are of particular interest because they indicate the magnitude of a potential “signal” in a particular frequency band. Due to the spike-and-slab prior, those bands that follow the noise spectrum have a posterior non-centrality parameter near zero, with a correspondingly high posterior probability of $\varrho = 0$. The majority of frequency bands follow the posterior noise spectrum, with only a handful of low-frequency bands having a high probability of $\varrho = 1$ (Figures 4.2, 4.3, and 4.4).

The posterior distributions of the autoregressive noise spectrum parameters suggest that the spectra of the three peat core time series are only slightly “red” (Figure 4.5). This characterization is based on posterior distributions for $\phi_1$ that are concentrated at values less than about 0.4 for the Greyscale and Ash Content series, and
Figure 4.2: Posterior distribution of non-centrality parameters for the Greyscale series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
Figure 4.3: Posterior distribution of non-centrality parameters for the Humification series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
Figure 4.4: Posterior distribution of non-centrality parameters for the Ash Content series from the Tibetan peat core. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated—frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
Figure 4.5: Posterior distribution of autoregressive noise spectrum parameters from the Tibetan peat core. The variance parameters for the three series are: (a) for Greyscale; (b) for Humification; (c) for Ash Content. Similarly, the $\phi_1$ parameters for the three series are: (d) for Greyscale; (e) for Humification; (f) for Ash Content.

less than about 0.6 for the Humification series. Despite the use of the mtLS statistic to limit spectral leakage, it is possible that some small amount of leakage from low frequencies to high frequencies contributed to the results seen in the posterior distributions of noise spectrum parameters.

Given that some of the non-centrality parameters had posterior distributions that were different than zero (Figures 4.2, 4.3, and 4.4), it is natural to wonder which frequencies in particular might be responsible within a given band of width $2W$. I computed the mtLS spectral statistics for each of the three series from the Tibetan
4.2. ANALYSIS OF PALEOCLIMATE SERIES

peat core at high frequency resolution, for each of the 1000 posterior chronology samples obtained in Chapter 2 (Figure 4.6). The Greyscale and Humification series are both thought to be measures of the strength of the summer monsoon, and appear to show similar spectral features at low frequencies. In particular, there appear to be features at periods of approximately 4000, 2200, and 7500-8000 years. The Ash Content series is thought to be a measure of the strength of the winter monsoon, and appears to show fewer distinct spectral features, except possibly two at periods of approximately 2700 and 1200 years. It could be speculated that the delivery of ash to the region depends not only on the strength of the winter monsoon but also on a steady source of ash. If there was variability in the availability of ash this could lead to less well-defined signals in the spectral statistics.

4.2.2 Prairie lake series

The Canadian prairie lake series are lake sediment records collected from three lakes (Michels et al., 2007). The three series belong to three separate records each with their own chronology. The posterior distributions for each chronology were obtained in Chapter 2. The three lakes studied were:

- Chauvin Lake in eastern Alberta,
- Humboldt Lake in central Saskatchewan,
- Oro Lake in southern Saskatchewan.

Diatoms are algae whose fossils can be used to infer an index of lake salinity using the relative abundances of different species. Based on records of diatom assemblages, an index of salinity was inferred for each lake as a function of depth. Using this method,
Figure 4.6: Frequency-wise mtLS spectral statistics for the series from the Tibetan peat core. The spectral statistics have been calculated at high frequency resolution in order to examine characteristics within-band. The heavy line is the mean spectral statistic across 1000 posterior chronology samples, and finer lines are the 2.5 and 97.5 percentiles. Gray lines delineate the $2W$ intervals.
there is evidence of shifts in drought conditions in the Canadian prairies. Salinity index data are available at a resolution of 2 cm for Chauvin and Oro Lake, and a resolution of 1 cm for Humboldt Lake.

**Specification of priors**

The Bayesian spectral model in Section 4.1 requires specification of some prior parameters. As for the peat core analysis, priors were chosen to be at most weakly informative. For the priors on the variance of the non-centrality parameters, I again followed Ishwaran and Rao (2005) and specified: \( \alpha = 5, \beta = 50, \) and \( v_0 = 0.005. \)

For the priors on the noise spectrum variance, I specified the bounds on the uniform priors so that they did not restrict the posteriors: \( \sigma_{p,L} = 0.1 \) and \( \sigma_{p,U} = 3. \)

**Posterior distributions**

As for the peat core analysis above, samples were obtained from the posterior distribution of model parameters using the methodology in Section 4.1.5. This involved alternating draws from the posterior distribution of chronologies and the conditional posterior distribution of spectral parameters. In identical fashion to the peat core analysis, three Markov chains were used, with starting parameter values drawn from prior distributions. The same sampling protocol was used: First, a chronology was drawn. Following an initial burn-in of 10,000 samples, 500 samples were drawn from the posterior distribution of spectral parameters for each chain and thinned to 50 samples (150 samples total). Then, a new chronology was drawn from the posterior distribution of chronologies, and a further 100 samples were burned before drawing another 500 samples thinned to 50. This process of sampling the chronology and then
sampling the spectral parameters was repeated one hundred additional times, for a total of 15,150 samples.

For the prairie lakes, the majority of frequency bands follow the posterior noise spectrum, with only a handful of low-frequency bands having a high probability of $\varrho = 1$ (Figures 4.7, 4.8, and 4.9). This was especially true for Oro Lake, suggesting that the relative uncertain chronology played a role in obscuring any potential signals.

The posterior distributions of the autoregressive noise spectrum parameters suggest that the spectra of the three lake sediment core time series are moderately “red” (Figure 4.10); the posterior distributions for $\phi_1$ have values ranging from near zero up to about 0.6 for Chauvin and Humboldt Lake, and about 0.8 for Oro Lake. There remains a fair amount of uncertainty in $\phi_1$.

Some of the non-centrality parameters had posterior distributions that were different than zero (Figures 4.7, 4.8, and 4.9). I computed the mtLS spectral statistics for each of the three lake salinity series at high frequency resolution, for each of the 1000 posterior chronology samples obtained in Chapter 2, to see if there is a suggestion of any signals within-band (Figure 4.11). For Chauvin Lake, there appears to be features at periods of approximately 2000, 6700, and 1100 years. For Humboldt Lake, there appears to be features at periods of about 8000 and 2000 years. For Oro Lake, there appears to be features at periods of about 4400 and 2400 years. Many of these features appear to be broad in the high resolution spectral statistics, so that they may be common among the three lakes or in two of the three lakes. In particular, the $\sim$2000 year feature in both Chauvin and Humboldt Lake stands out as being possibly in common. Other features at higher frequencies may be indicative of periodic phenomena, but timing uncertainty makes interpretation and identification
Figure 4.7: Posterior distribution of non-centrality parameters for the salinity index series from Chauvin Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
Figure 4.8: Posterior distribution of non-centrality parameters for the salinity index series from Humboldt Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
Figure 4.9: Posterior distribution of non-centrality parameters for the salinity index series from Oro Lake. Plotted as blue bars are the probability that the non-centrality parameter is zero. Plotted as orange violin plots are the distribution of non-zero non-centrality parameters. Frequencies are in cycles per year. The frequency range has been truncated – frequencies higher than those displayed all have zero non-centrality parameter with probability near 1.
4.3 Discussion

In this Chapter, I followed the lead of Mann and Lees (1996) and chose an AR(1) process as the baseline or noise spectrum. The recommendation of an AR(1) process was based on three considerations:

1. The physics of the climate system,
Figure 4.11: Frequency-wise mtLS spectral statistics for the series from the prairie lakes data. The spectral statistics have been calculated at high frequency resolution in order to examine characteristics within-band. The heavy line is the mean spectral statistic across 1000 posterior chronology samples, and finer lines are the 2.5 and 97.5 percentiles. Gray lines delineate the $2W$ intervals.
2. The interpretation of the characteristic decay time $t$, and

3. An aversion to interpreting signals as being part of the noise spectrum.

The first consideration is one that should not be taken lightly; given that we are presently interested in a model for climate processes, it is important that a noise process suited to the particular application be used. The second consideration is perhaps less important if our primary interest is in the characterization of signals rather than the characterization of the noise process. It is a boon to have a more easily interpreted model for the system, but only if the model is appropriate. The third consideration is statistical and could be investigated further. In Thomson et al. (2001), an AR(15) process combined with a smoothed running quantile estimate was used to model the noise spectrum. Thomson et al. advise increasing the number of AR parameters until the range of the intermediate spectrum (the ratio of the multitaper spectrum and the AR spectrum) no longer delineates rapidly as a function of the AR order. This approach is reminiscent of the “scree plot” approach or stopping rule in deciding the number of principal components or underlying factors in multivariate statistical analysis (Cattell, 1966; Horn, 1965; Jackson, 1993).

Thomson et al. (2001) further post-whitens the spectrum by estimating the local 5th-10th percentile of the spectral statistics, since these low percentiles are relatively robust to the presence of signals. Using smoothed local estimates of these percentiles, an empirical noise spectrum is estimated and subsequently removed, so that the remaining variance at each frequency can be compared to a central chi-squared null distribution for detection of signals. It might be possible to incorporate this kind of frequency-local noise estimate into the Bayesian modelling approach specified in this chapter by including a parametric function of frequency instead of the autoregressive
noise process. The choice of function might be dependent on context, but smoothly varying functions such as splines would be one option.

Although the mtLS statistic has limited spectral leakage, in cases where there is much dynamic range in the spectrum, there can be some contribution at remote frequencies from strong signals. For example, the mtLS statistics show a slight upward trend at high frequencies possibly due to leakage from low frequencies. Whereas an AR(1) noise spectrum is unable to capture this shape, a smooth function would be able to track the shape and as a result provide an improved model for the non-signal portion of the spectrum and the process variance $\sigma_p^2$. There are reasons to prefer an autoregressive noise spectrum over a smooth function such as a spline. Whereas the autoregressive process has a well-known and well-studied time domain representation, which may even have special meaning in some contexts, the smooth function approach might not be as easily interpreted, and could not be justified based on a priori knowledge of the process.

It is hardly surprising that few non-centrality parameters were found to have posterior distributions away from zero in the series examined. Uncertainty in the chronologies leads to uncertainty in the mtLS statistics (e.g. Figure 3.8), and ultimately impacts the high-frequency portion of the mtLS statistics the most. This is precisely the point: Rather than ignore uncertainty in the times, my goal here was to account for the posterior distribution of times when obtaining the posterior distribution of spectral model parameters. In this way, we avoid being overconfident when performing spectral inference by assuming that times are known when in fact they are not. As a consequence, we can be confident that our spectral inference represents an honest assessment of the major sources of uncertainty, and can move to issues of
model choice and so on.

The paleoclimate analyses in this chapter all had strong signals in the lowest frequency bands. Other than subtracting the mean of each series, I did not include a model for the general trend in the time domain other than that provided by the spectral model. It is worth investigating models for the general trend in the time domain, especially in cases where there is a suspected regime shift or discontinuity, or in cases where there is an overall trend in one direction. A mean value function of this type might be included in the same way as the noise spectrum, by “dividing out” the spectral representation of the process.

The choice of frequency bandwidth parameter when computing the mtLS statistics determines the spacing of the frequency mesh over which the spectral model is parameterized because spectral statistics outside of the band are approximately uncorrelated and assumed independent when computing the product likelihood. In Figures 4.6 and 4.11, I plotted frequency-wise mean and percentile mtLS statistics at a high frequency resolution to illuminate which frequencies in-band might be contributing to any signal present. I subsequently identified several frequencies that had apparent peaks. These peaks include uncertainty due to the posterior uncertainty in the chronology, but care should be taken in their interpretation because they may be within-band and are thus not independent. In particular, wiggle in the mean value of the spectral statistics that are within the credible regions are best ignored. Still, these examinations can be suggestive if comparing to a priori findings from other studies. In fact, if certain frequencies are of interest a priori, then the frequency mesh for the mtLS statistics can be specified to include those frequencies of interest, including an appropriate bandwidth so that the mtLS statistics can still form the basis of a product
likelihood. The usual tradeoff of between frequency resolution and variance of the likelihood applies.

4.4 Conclusion

In this chapter, I describe a new approach for Bayesian modelling of the spectrum of a time series process, with particular application to paleoclimate time series generated from core data. The approach combines ideas from the well-known and widely applied work of Mann and Lees (1996) and Thomson et al. (2001) with recent developments in “spike-and-slab” modelling of high-dimensional parametric models (Ishwaran and Rao, 2005). Uncertainty in the chronology of the core is propagated to the posterior distribution of spectral model parameters through marginalization, assuming that the time series process depends on the chronology only through the likelihood. Thus, the posterior distribution of spectral model parameters includes uncertainties due to both the process of interest as well as the chronology.
Chapter 5

General Summary and Recommendations

This thesis describes a cohesive framework for spectral analysis of paleoclimate records in which the times are unknown and must be inferred. It presents three key contributions, presented as three stand-alone chapters.

In Chapter 2, I detailed a new approach for modelling the chronology of time series data collected from cores extracted from a sedimentary record. The novelty of the approach was in the inclusion of of a group-level prior distribution for the time-to-accumulation whose parameters were given a hyperprior distribution intended to be at most weakly informative. The hierarchy results in time-to-accumulation parameters for each sampling interval that are a compromise between the local information available from the radiocarbon age likelihoods and the overall age versus depth trajectory of the core. This type of compromise has intuitive appeal, especially given that any particular radiocarbon age is prone to unquantified errors in the field or the lab. It also avoids having to specify an informative prior for unknown accumulation parameters as is the case in some other approaches. Chronology model parameters were sampled from the posterior using an MCMC sampler, so that posterior chronologies could be carried through to the spectral analysis in Chapter 4.
The posterior distribution of the chronology in hand, the next step is the analysis of the proxy record. In Chapter 3, I presented a multitaper spectral statistic based on the Lomb-Scargle periodogram (the mtLS), which had demonstrably reduced spectral leakage compared to the Lomb-Scargle. Because the spectral window varies as a function of frequency and depends on the particular pattern of observation times, performance is not guaranteed and the spectral pseudowindows should be checked prior to interpretation of the statistic as an estimator or otherwise. As an estimator, the mtLS was able to replicate an exoplanet detection from a Kepler photometry time series. As a statistic, the low spectral leakage was used to justify a product likelihood in the Bayesian spectral model of Chapter 4.

Previous work by Mann and Lees (1996) and Thomson et al. (2001) were the motivation for the Bayesian spectral model in Chapter 4. To our knowledge, the Bayesian spectral model is a new approach for spectral inference, incorporating a baseline noise spectrum that is an AR(1) process, and a spike-and-slab prior to separate frequency bands that follow the noise spectrum from frequency bands that contain a signal. Although the Bayesian spectral model could be used for time series with known and either regular or irregular times, a strength of the Bayesian approach is the ability to marginalize over the timing uncertainty represented by the posterior chronology distribution from Chapter 2. The resulting posterior distribution for the spectral model parameters are thus able to include the uncertainty arising not only from the time series process but also from the chronology. When applied to actual paleoclimate series from sedimentary records, moderate to high frequency bands are interpreted as noise, and only low frequency signals are interpreted as such. Given the uncertainty present in the chronology, this is not unexpected.
Recommendations for future work are presented in each chapter, but I will mention some further recommendations here. The first recommendation is to explore the use of more sophisticated MCMC techniques to improve posterior sampling efficiency. The MCMC sampling techniques used in this thesis were adequate but the compute times could be improved if the techniques are to be used by paleoclimatologists. It is also preferable if the MCMC sampling technique did not require tuning, so that the process of generating samples from the posterior was nearly automatic.

The second recommendation is to investigate relaxing the assumption that the accumulation process giving rise to the chronology and the climate processes giving rise to the proxy data are independent. The way in which this would be done depends on context – one possibility might be to have a latent process that affects both the time-to-accumulation and the spectrum of the climate proxy.

Finally, a model for combined inference in the case of multiple proxies within a single core such as the Tibetan peat core from Yu et al. (2006), or multiple proxies from different cores such as the Canadian prairie lakes data from Michels et al. (2007). Again, the appropriate model would be dependent on context, but possibilities include a latent climate process or modelling the coherence between proxies. Coherence is the frequency-domain analogue of correlation, and if a coherence statistic could be computed from the mtLS for two series, would enable identification of signals in common between two proxies.
Bibliography


