LINEAR TRENDS, PERIODICITIES, AND EXTREMES

by

JAECHOUL LEE

(Under the direction of Robert Lund)

ABSTRACT

Time series regression models with periodically autocorrelated errors are studied for purposes of studying United States extreme temperatures trends.

First we study properties of ordinary and generalized least squares estimates in a simple linear regression model with stationary, but autocorrelated, errors. Our primary focus is on deriving explicit expressions for the variances of the regression parameter estimates for some common time series autocorrelation structures, including a first order autoregression and general moving-averages. We review large sample properties of the parameter estimates and suggest improvements to confidence intervals that are used in practice. An example where the variance of the trend slope estimate actually decreases with increasing autocorrelation in the errors is presented.

We then move to the study of linear trends in monthly maximum and minimum temperatures observed during the past 150 years in the contiguous 48 United States. Regression and extreme value trend modeling methods are compared and contrasted. Issues of temporal autocorrelations, periodicities, extreme value modeling, station relocations (changepoints), and spatial smoothing algorithms arise. The seasonal aspect of the analysis allows for the issue of uniformity of temperature change over varying season to be investigated. Spatial contour maps of the estimated trends are drawn for each of the four seasons and the entire year. Finally, a conjecture on why the Southeastern United States is cooling is put forth.

LINEAR TRENDS, PERIODICITIES, AND EXTREMES

by

JAECHOUL LEE

B.A., Korea University, Korea, 1994
M.S., Korea University, Korea, 1998

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003
LINEAR TRENDS, PERIODICITIES, AND EXTREMES

by

JAECHOUL LEE

Approved:

Major Professor: Robert Lund

Committee: William P. McCormick
Mary Meyer
Jaxk Reeves
Lynne Seymour

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003
ACKNOWLEDGMENTS

Foremost I am extremely grateful to Dr. Robert Lund who not only served as my major professor but also encourage me throughout my Ph.D. program at UGA. His support, patience, and guidance was essential to the completion of this dissertation and he gave me valuable insights and ideas on academic research. Thank you again, Dr. Lund!

My sincere thanks go to Dr. Lynne Seymour for her kind help and technical suggestions. I also would like to express my gratitude to the other members of my advisory committee: Dr. Ishwar Basawa, Dr. William McCormick, Dr. Mary Meyer, Dr. Jaxk Reeves, and Dr. Robert Taylor for reading previous drafts of this dissertation and providing many valuable comments that improved the content of this dissertation. I also thank the staff of our Statistics Department: Molly Rema, Connie Durden, Daphne Smith, David Gavisk, and Jesse Bowling for their paperwork help and technical assistances related to my study.

Furthermore, I would like to thank some of my friends who contributed to this dissertation and influenced my life. QiQi Lu is thanked for her help in drawing beautiful maps; Mi-Ja Woo gave technical computing help. Special thanks go to Dr. Jeong-gun Park and his family for their love to me during the first three years of my stay in the USA.

I want to thank my family for their consistent love, especially my father, Joon-Hee Lee, and mother Myung-Ja Choi for their continuing spiritual support for me.

Last I would like to thank God. This dissertation also is Yours.
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acknowledgments</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>References</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Literature Review</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Regression Models with Autocorrelated Errors</td>
<td>4</td>
</tr>
<tr>
<td>2.2</td>
<td>Extreme Value Modeling</td>
<td>8</td>
</tr>
<tr>
<td>2.3</td>
<td>References</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Revisiting Simple Linear Regression with Autocorrelated Errors</td>
<td>13</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>Variances of the Estimates</td>
<td>15</td>
</tr>
<tr>
<td>3.3</td>
<td>Efficiency Comparisons</td>
<td>17</td>
</tr>
<tr>
<td>3.4</td>
<td>Asymptotic Properties and Confidence Intervals</td>
<td>21</td>
</tr>
<tr>
<td>3.5</td>
<td>Non-Monotonicity of Variances in Autocorrelation</td>
<td>26</td>
</tr>
<tr>
<td>3.6</td>
<td>References</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>Trends in Extreme United States Temperatures</td>
<td>31</td>
</tr>
</tbody>
</table>
4.1 Monthly Maximum and Minimum Temperatures at Lewiston, Maine . 36
4.2 Station Locations . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
4.3 Lewiston ARMA(3,2) Residuals and Sample Autocorrelations . . . . 46
4.4 Comparison of Regression and GEV Linear Trends . . . . . . . . . . 49
4.5 Boxplot of Trend Estimates by Season . . . . . . . . . . . . . . . . . 51
4.6 Boxplot of Trend $z$-Scores by Season . . . . . . . . . . . . . . . . . 52
4.7 Smoothed Winter Trends . . . . . . . . . . . . . . . . . . . . . . . . . 56
4.8 Smoothed Spring Trends . . . . . . . . . . . . . . . . . . . . . . . . . 57
4.9 Smoothed Summer Trends . . . . . . . . . . . . . . . . . . . . . . . . . 58
4.10 Smoothed Fall Trends . . . . . . . . . . . . . . . . . . . . . . . . . . . 59
4.11 Smoothed Annual Trends . . . . . . . . . . . . . . . . . . . . . . . . . 60
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Simulation Length Comparison of the Two Methods</td>
<td>25</td>
</tr>
<tr>
<td>4.1</td>
<td>Monthly Lewiston Regression Estimates</td>
<td>45</td>
</tr>
<tr>
<td>4.2</td>
<td>Seasonal Lewiston Regression Estimates</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>Monthly Lewiston GEV Estimates</td>
<td>47</td>
</tr>
<tr>
<td>4.4</td>
<td>Seasonal Lewiston GEV Estimates</td>
<td>48</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Temperatures and their changes greatly affect our lives. Life would change considerably if temperatures change even slightly. A hot spell in the summer, for example, could decrease corn yields (Mearns et al. 1984); while a deep freeze during the winter could kill many citrus trees (Miller and Glantz 1988). Thus, understanding trends in extreme temperatures is of fundamental importance. In this dissertation, we study linear trends in monthly maximum and minimum temperatures in the contiguous 48 United States over the past century. This dissertation also contributes research on the general topic of time series, regression, and extreme value modeling.

In Chapter 2, we introduce a statistical regression model for analyzing time trends in temperatures. Our exposition there focuses on mean behavior rather than extremes. A brief review of historical and recent trend studies is presented there. Extreme value modeling approaches are also narrated.

Chapter 3 presents the mathematical framework of our analysis. In particular, this chapter compares properties of ordinary and generalized least squares parameter estimates in a simple linear regression model when the regression errors are autocorrelated, but stationary. The properties derived here provide useful insight into the model that we adopt in Chapter 4, which is essentially a periodic extension of a simple linear regression that allows fixed shifts for change points and seasonal trend slopes. Explicit expressions for the variances of the regression parameter estimates are derived for some common autocorrelation structures which include a first order
autoregression and general moving-averages. We discuss large sample properties of
the regression parameter estimates and suggest improvements to confidence inter-
vals that are used in practice. Finally, we present a counterintuitive example where
a more accurate trend slope estimate is obtained as the correlation in the regression
errors increases.

In Chapter 4, a periodic time series regression model is introduced to quantify
trends in monthly extreme United States temperatures. The time series regression
techniques are developed and compared with extreme value trend modeling tech-
niques. We consider a case study of monthly minimum temperatures at Lewiston,
Maine, for illustrative purposes. Linear monthly trends and their standard errors
are estimated for 528 United States locations. Seasonal and annual trends and their
standard errors are also obtained to summarize a general pattern of extreme temper-
ature change. The estimated trends are then spatially smoothed by longitude and
latitude with the weighted head-banging algorithm. Contour maps of the four sea-
sons and the entire year are presented. Warming and cooling areas are identified and
conjectures about the patterns observed are made. Practical implications regarding
global warming are also discussed.

Finally, Chapter 5 proposes a future plan of study. The topics include efficiency
comparison of ordinary least squares and generalized least squares estimates in time
series regression models with periodically autocorrelated errors. Interdisciplinary
research topics of our interest are also listed.

1.1 References

ture events: changes in their probabilities with changes in mean temperature,
Chapter 2

Literature Review

2.1 Regression Models with Autocorrelated Errors

A statistical model used in many global temperature trend studies is the simple linear regression

\[ X_t = \mu + \alpha t + \epsilon_t, \quad t = 1, 2, \ldots, n, \]  

(2.1)

where \( \{\epsilon_t\} \) is mean zero random error. Here, \( X_t \) denotes the observed temperature at time \( t \). Due to autocorrelations in temporally adjacent temperatures, \( \{\epsilon_t\} \) is not necessarily white noise.

The ordinary least squares estimates for the parameters \( \alpha \) and \( \mu \) above, denoted by \( \hat{\alpha}_{\text{OLS}} \) and \( \hat{\mu}_{\text{OLS}} \), are

\[ \hat{\alpha}_{\text{OLS}} = \frac{\sum_{t=1}^{n}(t - \bar{t})(X_t - \bar{X})}{\sum_{t=1}^{n}(t - \bar{t})^2} \quad \text{and} \quad \hat{\mu}_{\text{OLS}} = \bar{X} - \hat{\alpha}_{\text{OLS}} \bar{t}, \]  

(2.2)

where \( \bar{t} = (n + 1)/2 \) and \( \bar{X} = n^{-1} \sum_{t=1}^{n} X_t \).

The model (2.1) has been applied in studying temperature data by many climate authors investigating global warming issues. Roden (1966) estimates linear trends for Western United States temperatures and concludes that climatic changes are not significant in non-urban regions, but that temperatures in urban localities have significantly increased. Jones and Kelly (1983) present spatial maps of linear trends of Northern Hemisphere annual mean temperatures for the three periods (the 1917–1939 strong warming period, the 1940–1964 cooling period, and the 1965–1980 warming period), which are in general agreement with Hansen and Lebedeff’s (1987)

As temperature data is strongly serially correlated in time $t$, the usual regression assumption that $\{\epsilon_t\}$ is white noise is violated. Hence, $\hat{\mu}_{\text{OLS}}$ and $\hat{\alpha}_{\text{OLS}}$ are not necessarily best linear unbiased estimates (BLUE). However, assuming that the errors $\{\epsilon_t\}$ are covariance stationary in time with mean zero and a spectral density that is bounded away from zero, it is well known from Grenander (1954) that the OLS estimators in (2.2) are asymptotically as efficient as generalized least squares estimators.

A controversial statistical issue lies with computation of standard errors for the estimators (cf. Fomby 2002). Detailed properties of the OLS estimators are derived in Chapter 3.

Periodicities are present in most climate data (not only in temperatures). A random sequence $\{Y_t\}$ with finite second moments is called a periodic time series with period $T$ if

$$E[Y_{t+T}] = E[Y_t] \quad (2.3)$$

and

$$\text{Cov}(Y_{t+T}, Y_{s+T}) = \text{Cov}(Y_t, Y_s) \quad (2.4)$$

for all integers $t$ and $s$. The period $T$ is taken as the smallest positive integer satisfying both (2.3) and (2.4). A survey of periodic time series and their properties is given by Lund and Basawa (1999) and Lund and Seymour (2002).

A more general version of the regression model in (2.1) allowing for periodicity in the mean response, as well as periodicities in the autocovariances of $\{\epsilon_t\}$, is now
introduced. This is the simple regression model

\[ X_{nT+\nu} = \mu_{\nu} + \alpha_{\nu}(nT + \nu) + \epsilon_{nT+\nu}, \quad (2.5) \]

where \( X_{nT+\nu} \) is the observation taken during month \( \nu \) of year \( n \). Here, \( \nu \) is a monthly index satisfying \( 1 \leq \nu \leq T \), and \( T = 12 \) is the period of the series. In (2.5), \( \alpha_{\nu} \) is the linear trend for month \( \nu \), the intercept term \( \mu_{\nu} \) is the month \( \nu \) location parameter, and \( \{\epsilon_t\} \) is assumed to be a periodic time series with mean zero (\( \mathbb{E}[\epsilon_t] = 0 \)) and period \( T \). The model in (2.5) is more appropriate for analyzing monthly data. The OLS estimates for \( \alpha_{\nu} \) and \( \mu_{\nu} \) can be worked out as

\[
\hat{\alpha}_{\nu} = \frac{\sum_{n=0}^{d-1}(nT + \nu - \bar{t}_\nu)X_{nT+\nu}}{\sum_{n=0}^{d-1}(nT + \nu - \bar{t}_\nu)^2} \quad \text{and} \quad \hat{\mu}_{\nu} = \bar{X}_{\nu} - \hat{\alpha}_{\nu}\bar{t}_\nu, \quad (2.6)
\]

where \( \bar{t}_\nu = d^{-1}\sum_{n=0}^{d-1}(nT + \nu) \) and \( \bar{X}_{\nu} = d^{-1}\sum_{n=0}^{d-1}X_{nT+\nu} \) are the average time-of-observation and observation during month \( \nu \). The notation here uses \( d \) as the number of years of data. For simplicity of exposition, we take \( d \) as a whole multiple of \( T \).

However, the model in (2.5) does not satisfactorily describe changepoints. As most temperature recording stations move locations due to sociological reasons, change their recording instruments, or change temperature gauge types, a regression model considering such “changepoints” (formally defined as a station instrument, shelter, and/or location change) is needed. Fortunately, the times of the changepoints are known in this data.

Suppose that a station has experienced \( k \) different regimes over the observation record (i.e., there has been \( k - 1 \) changepoints) and that time of all site changes is known. A modification of (2.5) that allows for changepoints is

\[ X_{nT+\nu} = \mu_{\nu}^{(i)} + \alpha_{\nu}(nT + \nu) + \epsilon_{nT+\nu}, \quad (2.7) \]

where again \( X_{nT+\nu} \) is the observed temperature during month \( \nu \) of year \( n \). Here, \( \mu_{\nu}^{(i)} \) is the intercept term during month \( \nu \) while the station is in regime \( i \), for \( 1 \leq i \leq k \).
This is allowed to vary by regime and month. This model is used by Lund et al. (2001). For US monthly averages temperatures over the past century, Lund et al. (2001) estimate linear trends for about 300 stations and their standard errors, and display geographical patterns of temperature change across the contiguous 48 United States. They found that cooling is taking place in the Southeast and warming in the Northeast, the West, and the Northern Midwest.

A possible drawback of the model in (2.7) involves the large number of parameters. Specifically, $T$ location parameters $\mu^{(i)}_\nu$ are “spent” each time the station experiences a changepoint. Suppose that the site changes occur at the ordered times $1 < \tau_1 < \tau_2 < \cdots < \tau_{k-1} < N$, where $N = dT$ is the total number of monthly series values. The following model allows for a mean shift in temperatures at each changepoint time, where the changepoint shift effect is uniform over the months:

$$X_{nT+\nu} = \mu_\nu + S_{nT+\nu} + \alpha_\nu(nT + \nu) + \epsilon_{nT+\nu},$$

(2.8)

where $S_{nT+\nu}$ is the mean shift changepoint effect satisfying

$$S_{nT+\nu} = \begin{cases} 
0, & \text{if } 1 \leq nT + \nu < \tau_1, \\
\Delta_2, & \text{if } \tau_1 \leq nT + \nu < \tau_2, \\
\vdots & \\
\Delta_k, & \text{if } \tau_{k-1} \leq nT + \nu \leq N.
\end{cases}$$

(2.9)

The mean shift during the first regime is taken as $\Delta_1 = 0$ as a baseline. This is needed for parameter identifiability. The OLS estimates for the parameters in (2.8) and (2.9) cannot be explicitly evaluated in a simple closed form. Part of our research involves computation and standard error evaluation for $\hat{\alpha}_\nu$ in the above regression model.

The model in (2.8) and (2.9) is key in our trend study of monthly extreme temperatures. With such a periodic model, one can quantify seasonal temperature
changes. There is little reason to expect that temperatures are changing uniformly across the seasons; our methods can address this issue.

2.2 Extreme Value Modeling

As the temperatures we examine are extremes, one should also consider classical approaches for extreme value modeling. Likelihood methods for fitting generalized extreme value (GEV) distributions or related families are well-studied. The GEV family of distributions focuses on the behavior of

$$M_n = \max\{Y_1, Y_2, \ldots, Y_n\},$$

where \(\{Y_i\}_{i=1}^n\) is a sequence of independent and identically distributed random variables. The distribution of \(M_n\) can be approximated by a member of the GEV family with distribution function

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - m}{v} \right) \right]^{-1/\xi} \right\},$$

defined on \(\{z : 1 + \xi(z - m)/v > 0\}\), where \(-\infty < m < \infty, v > 0,\) and \(-\infty < \xi < \infty\) are location, scale, and shape parameters, respectively. This family includes Gumbel, Fréchet, and Weibull families. The case \(\xi = 0\) is interpreted as a limit as \(\xi \to 0\) in (2.11), leading to the Gumbel family. The cases \(\xi > 0\) and \(\xi < 0\) correspond to Fréchet and Weibull families.

Suppose that \(\{M_{n,1}, M_{n,2}, \ldots, M_{n,m}\}\) is a series of block maxima from data blocked into sequences of \(n\) observations. The blocking is frequently made by a time period of one year/season/month, in which case \(n\) is the number of observations in a year/season/month and the block maxima are taken as annual/seasonal/monthly maxima. Then the GEV distribution can be used to describe the series of block maxima. For an early application to environmental work, see Horowitz (1980). A recent summary of detailed analysis procedures is introduced in Beirlant et al. (1996).
and Coles (2001). Numerous applications to insurance and hydrology are given by Reiss and Thomas (2001).

A more comprehensive approach is to use the joint distribution of the $r$ largest/smallest values (Smith 1986 and Tawn 1988), instead of just the maxima/minima. We often use the largest/smallest $r$ order statistics

$$\tilde{M}_n^{(r)} = (M_n^{(1)}, M_n^{(2)}, \ldots, M_n^{(r)})$$

(2.12)

for each of several blocks, where $M_n^{(k)}$ is the $k$th largest/smallest order statistic of independent and identically distributed sequences $\{Y_1, Y_2, \ldots, Y_n\}$. Application to athletic records, fitted by the GEV distribution and estimated through maximum likelihood approaches, is presented in Robinson and Tawn (1995). However, this approach is not possible in our context as we only have maximum/minimum ($r = 1$) for each month.

The blocking and thresholding peaks over threshold is another alternative for studying extremes. Extreme values can be regarded as those of the $Y_i$ observations exceeding some high/low threshold $w$. Then, for the large/small enough $w$, an approximate conditional distribution function of $Y - w$ given $Y > w$ is obtained as

$$H(y) = 1 - \left(1 + \frac{\xi y}{\tilde{v}}\right)^{-1/\xi},$$

(2.13)

where $\tilde{v} = v + \xi(w - m)$. The support set of the distribution is $\{y : y > 0 \text{ and } (1 + \xi y/\tilde{v}) > 0\}$.

The family of distributions in (2.13) is called the generalized Pareto family. The threshold method considers the magnitude of exceedances over a high/low threshold. After determining a threshold, maximum likelihood estimates can be computed numerically from the likelihood function from (2.13). Smith (1989) applies this approach to estimate trends in hourly ozone data from Houston, TX. For more applications, see Walshaw (1994) and Grady (1992).
General extreme value models and their analysis methods are summarized nicely in Coles (2001). There, several applications to environmental data and a brief introduction to point process methods for extreme values are given.

2.3 References


Chapter 3

Revisiting Simple Linear Regression with Autocorrelated Errors

3.1 Introduction

This chapter studies the ordinary and generalized least squares parameter estimates in the simple linear regression model

\[ X_t = \mu + \alpha t + \epsilon_t \]  

(3.1)

when the errors \( \{\epsilon_t\} \) are mean zero and stationary in time \( t \) with autocovariance \( \gamma(h) = \text{cov}(\epsilon_t, \epsilon_{t+h}) \) at lag \( h \).

The model in (3.1) has a fundamental role in many statistical analyses (cf. Grenander 1954, Harvey and Phillips 1979, Zinde-Walsh and Galbraith 1991, Choudhury et al. 1997 and 1999). There is little loss of generality in considering (3.1) over the more general regression \( X_i = \mu + \alpha t_i + \epsilon_i \); hence, we assume that \( t_i = i \) from now on.

Ordinary least squares estimates, which have minimal variance when \( \{\epsilon_t\} \) is uncorrelated, are frequently computed in lieu of weighted least squares estimates when in fact \( \{\epsilon_t\} \) is autocorrelated. This chapter rehashes aspects of this classic problem. By being very patient with algebraic manipulations, we obtain explicit structures for the variances of the parameter estimates under some autocorrelation structures commonly encountered in time series. We also point out some new facts along the way.
The ordinary least squares estimates of $\mu$ and $\alpha$, denoted by $\hat{\mu}_{OLS}$ and $\hat{\alpha}_{OLS}$, have the simplistic explicit form

$$
\hat{\alpha}_{OLS} = \frac{\sum_{t=1}^{n}(t - \bar{t})(X_t - \bar{X})}{\sum_{t=1}^{n}(t - \bar{t})^2}
$$

(3.2)

and

$$
\hat{\mu}_{OLS} = \bar{X} - \hat{\alpha}_{OLS} \bar{t},
$$

(3.3)

where $\bar{t} = (n + 1)/2$ and $\bar{X} = n^{-1} \sum_{t=1}^{n} X_t$ are the time and observation averages.

The denominator of (3.2) can be explicitly evaluated as

$$
\sum_{t=1}^{n}(t - \bar{t})^2 = \frac{n(n+1)(n-1)}{12}.
$$

(3.4)

The ordinary least squares estimates have the general linear models form

$$
\begin{pmatrix}
\hat{\mu}_{OLS} \\
\hat{\alpha}_{OLS}
\end{pmatrix} = (D_n' D_n)^{-1} D_n' \bar{X}_n.
$$

(3.5)

where $\bar{X}_n = (X_1, \ldots, X_n)'$ (’ denotes matrix transpose) is the observation vector and $D_n$ is the design matrix

$$
D_n = \begin{pmatrix}
1 & 1 \\
1 & 2 \\
\vdots & \vdots \\
1 & n
\end{pmatrix}.
$$

(3.6)

The estimates $\hat{\mu}_{OLS}$ and $\hat{\alpha}_{OLS}$ are unbiased for any mean zero $\{\epsilon_t\}$; however, the ordinary least squares estimates will not have the smallest variance amongst all unbiased estimates unless $\{\epsilon_t\}$ is white noise (uncorrelated with a constant variance).

The minimum variance unbiased estimates of $\mu$ and $\alpha$ that are linear in $X_1, \ldots, X_n$, denoted by $\hat{\mu}_{GLS}$ and $\hat{\alpha}_{GLS}$ to emphasize their interpretation in a generalized least squares paradigm (also called Gauss-Markov, weighted least squares,
or BLUE estimates) are
\[
\begin{pmatrix}
\hat{\mu}_{\text{GLS}} \\
\hat{\alpha}_{\text{GLS}}
\end{pmatrix}
= (D_n'\Gamma_n^{-1}D_n)^{-1}D_n'\Gamma_n^{-1}\tilde{X}_n.
\] (3.7)

In (3.7), $\Gamma_n$ denotes the covariance matrix of $\tilde{X}_n$ which is tacitly assumed invertible for each $n \geq 1$.

The generalized least squares estimates are also unbiased for any mean zero $\{\epsilon_t\}$ and have a smaller variance than the ordinary least squares estimates: $\text{var}(\hat{\mu}_{\text{GLS}}) \leq \text{var}(\hat{\mu}_{\text{OLS}})$ and $\text{var}(\hat{\alpha}_{\text{GLS}}) \leq \text{var}(\hat{\alpha}_{\text{OLS}})$. The ordinary and weighted least squares estimates are equal if and only if the columns of $D_n$ span the same linear subspace as the columns of $\Gamma_nD_n$ (cf. Bloomfield and Watson 1975).

### 3.2 Variance of the Estimates

The variance of the ordinary and weighted least squares estimates can be obtained in matrix form by taking variances in (3.5) and (3.7):
\[
\text{var} \begin{pmatrix}
\hat{\mu}_{\text{OLS}} \\
\hat{\alpha}_{\text{OLS}}
\end{pmatrix}
= (D_n'\Gamma_n^{-1}D_n)^{-1}(D_n'\Gamma_nD_n)(D_n'\Gamma_n^{-1}D_n)^{-1}
\] (3.8)

and
\[
\text{var} \begin{pmatrix}
\hat{\mu}_{\text{GLS}} \\
\hat{\alpha}_{\text{GLS}}
\end{pmatrix}
= (D_n'\Gamma_n^{-1}D_n)^{-1}.
\] (3.9)

There is a wealth of literature comparing the variances in (3.8) and (3.9), with several interesting efficiency bounds arising (cf. Grenander 1954, Gurland 1954, Watson 1955 and 1967, Zyskind 1967, Knott 1975, and Chipman 1979 for a historical sample).

As $\hat{\mu}_{\text{OLS}}$ and $\hat{\alpha}_{\text{OLS}}$ are linear combinations of $X_1, \ldots, X_n$, more explicit expressions for the individual parameter variances can be derived. In particular, some bookkeeping with (3.2) and (3.3) will yield
\[
\text{var}(\hat{\alpha}_{OLS}) = \frac{\gamma(0) + 2 \sum_{j=1}^{n-1} w_j \gamma(j)}{\sum_{t=1}^{n} (t - \bar{t})^2} \quad (3.10)
\]

and
\[
\text{var}(\hat{\mu}_{OLS}) = \frac{1}{n} \left[ 4 \left( \frac{2n+1}{2n-2} \right) \gamma(0) + 2 \sum_{j=1}^{n-1} \left( 1 - \frac{j}{n} + \frac{3(n+1)}{n-1} w_j \right) \gamma(j) \right], \quad (3.11)
\]

where the weights \( \{w_j\} \) are
\[
w_j = \frac{\sum_{t=1}^{n-j} (t - \bar{t})(t + j - \bar{t})}{\sum_{t=1}^{n} (t - \bar{t})^2}, \quad 0 \leq j \leq n - 1. \quad (3.12)
\]

Straightforward manipulations with (3.12) provide the explicit form
\[
w_j = \frac{(1 - j/n)(1 - 2j/n - 2j^2/n^2 - 1/n^2)}{(1 + 1/n)(1 - 1/n)}, \quad 0 \leq j \leq n - 1. \quad (3.13)
\]

The covariance of the ordinary least squares parameter estimates is
\[
\text{cov}(\hat{\mu}_{OLS}, \hat{\alpha}_{OLS}) = -\frac{\bar{t}}{\sum_{t=1}^{n} (t - \bar{t})^2} \left[ \gamma(0) + 2 \sum_{j=1}^{n-1} w_j \gamma(j) \right]. \quad (3.14)
\]

We note here some properties of \( \{w_j\} \) for later use. The Cauchy-Schwarz inequality provides \( |w_j| \leq w_0 = 1 \) for each \( j \). Taking a limit in (3.13) shows that \( \lim_{n \to \infty} w_j = 1 \) for each fixed \( j \). Finally, it is easy to algebraically verify that
\[
\sum_{j=1}^{n-1} w_j = -\frac{1}{2} \quad (3.15)
\]
for all \( n \geq 2 \).

**Example 3.1.** Suppose that \( \{\epsilon_t\} \) is the \( q \)th order moving average; specifically,
\[
\epsilon_t = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}, \quad (3.16)
\]

where \( \{Z_t\} \) is mean zero white noise with variance \( \text{var}(Z_t) \equiv \sigma^2 \). Then
\[
\gamma(h) = \sigma^2 \sum_{i=0}^{q} \theta_i \theta_{i-h}, \quad h \geq 0, \quad (3.17)
\]
where the convention $\theta_0 = 1$ is adopted. Note in particular that $\gamma(h) = 0$ when $|h| > q$. Combining (3.17) and (3.10) gives

$$\text{var}(\hat{\alpha}_{\text{OLS}}) = \sigma^2 \left[ \sum_{i=0}^{q} \theta_i^2 + \frac{2}{n(n^2-1)} \sum_{j=1}^{q} (n-j)(n^2-2jn-2j^2-1) \sum_{i=j}^{q} \theta_i \theta_{i-j} \right] \sum_{t=1}^{n} (t-t)^2.$$

(3.18)

An analogous expression for $\text{var}(\hat{\mu}_{\text{OLS}})$ can be obtained by combining (3.17) and (3.11); we omit the algebraic details.

Some mileage can be made with the explicit expressions obtained above. For example, differentiating (3.10) with respect to $\gamma(h)$ for a fixed $h \geq 1$ gives

$$\frac{\partial \text{var}(\hat{\alpha}_{\text{OLS}})}{\partial \gamma(h)} = \frac{2w_h}{\sum_{t=1}^{n} (t-t)^2}.$$

(3.19)

Hence, $\text{var}(\hat{\alpha}_{\text{OLS}})$ increases with increasing $\gamma(h)$ when $w_h > 0$. In particular, $\text{var}(\hat{\alpha}_{\text{OLS}})$ for a $q$th order moving-average increases with increasing $\gamma(h)$, $1 \leq h \leq q$, whenever $n^2 - 2qn - 2q^2 - 1 \geq 0$ (this inequality holds whenever $n \geq 3q$). We return to monotonicity of parameter estimate variances in Section 4.

3.3 Efficiency Comparisons

The model in (3.1) with autocorrelated errors has been the subject of many theoretical investigations. In particular, Grenander (1954) shows that if $\{\epsilon_t\}$ has a spectral density that is strictly positive at all frequencies, then

$$\lim_{n \to \infty} \frac{\text{var}(\hat{\alpha}_{\text{OLS}})}{\text{var}(\hat{\alpha}_{\text{GLS}})} = 1 \quad \text{and} \quad \lim_{n \to \infty} \frac{\text{var}(\hat{\mu}_{\text{OLS}})}{\text{var}(\hat{\mu}_{\text{GLS}})} = 1.$$

(3.20)

Hence, it is mathematically justifiable to use ordinary least squares estimates in lieu of generalized least squares estimates in substantial generality with a large sample. As Gaussian likelihood estimates have the same asymptotic properties as generalized least squares estimates and can be quite demanding to compute, there is little motivation to work in a likelihood paradigm.
A sufficient condition for the existence of a spectral density is the summability of autocovariances; that is, if
\[ \sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty, \]  

(3.21) then \( \{\epsilon_t\} \) has the continuous spectral density \( f \) defined at frequency \( \lambda \in [0, \pi) \) by
\[ f(\lambda) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \gamma(h)e^{-ih\lambda} \]  

(3.22) (cf. Brockwell and Davis 1991, Chapter 4). Grenander (1954) is an excellent early reference for mathematical generalities regarding (3.20). The stated efficiency in (3.20) holds for a fixed autocovariance structure of \( \{\epsilon_t\} \); if one fixes a sample size \( n \) and compares the variances in (3.20) over varying \( \gamma(\cdot) \), then, not surprisingly, ordinary least squares can be made arbitrarily inefficient relative to generalized least squares.

The next two examples will explicitly compute \( \hat{\mu}_{\text{GLS}} \) and \( \hat{\alpha}_{\text{GLS}} \) in first order autoregressions and moving-averages. The first example explicitly links ordinary and generalized least squares estimates.

**Example 3.2.** Suppose that \( \{\epsilon_t\} \) is a causal first order autoregression satisfying the difference equation
\[ \epsilon_t = \phi\epsilon_{t-1} + Z_t, \]  

(3.23) where \( \{Z_t\} \) is mean zero white noise with \( \text{var}(Z_t) \equiv \sigma^2 \). Causality of the model implies that \( |\phi| < 1 \). In this case, \( \hat{\alpha}_{\text{GLS}} \) is a linear combination of \( \hat{\alpha}_{\text{OLS}} \) and the endpoint estimate \( \hat{\alpha}_{\text{EP}} = (X_n - X_1)/(n - 1) \); specifically,
\[ \hat{\alpha}_{\text{GLS}} = \frac{w_{\text{OLS}}\hat{\alpha}_{\text{OLS}} + w_{\text{EP}}\hat{\alpha}_{\text{EP}}}{w_{\text{OLS}} + w_{\text{EP}}} , \]  

(3.24) where
\[ w_{\text{OLS}} = (1 - \phi)^2n(n + 1) \quad \text{and} \quad w_{\text{EP}} = 6\phi[(1 - \phi)(n - 1) + 2] \]  

(3.25)
(cf. Bloomfield and Nychka 1992). For the location parameter $\mu$, one can obtain

$$
\hat{\mu}_{\text{GGLS}} = \frac{(1-\phi)n}{(1-\phi)n + 2\phi} \bar{X} + \frac{2\phi}{(1-\phi)n + 2\phi} \left( \frac{X_1 + X_n}{2} \right) - \hat{\alpha}_{\text{GGLS}}. \quad (3.26)
$$

The relative efficiency of ordinary and generalized least squares trend estimates can be explicitly computed in this case; in particular, Chipman (1979) expends considerable effort to derive the explicit relation

$$
\frac{\text{var}(\hat{\alpha}_{\text{GGLS}})}{\text{var}(\hat{\alpha}_{\text{OLS}})} = \frac{n^2(n+1)^2(n-1)(1-\phi^2)(1-\phi)^4}{f(\phi,n)g(\phi,n)}, \quad (3.27)
$$

where

$$
g(\phi, n) = (n-3)(n-2)\phi^2 - 2(n-3)(n+1)\phi + n(n+1) \quad (3.28)
$$

and

$$
f(\phi, n) = -6\phi^{n+1}[(n-1)\phi - (n+1)]^2 - (n^3 - n)\phi^4 + 2(n^2 - 1)(n-3)\phi^3 
+ 12(n^2 + 1)\phi^2 - 2(n^2 - 1)(n+3)\phi + (n^3 - n). \quad (3.29)
$$

For $\phi \in (0, 1)$, Chipman (1979) shows that $\text{var}(\hat{\alpha}_{\text{GGLS}})/\text{var}(\hat{\alpha}_{\text{OLS}}) \geq 0.753763$ for any $\phi$ and sample size $n$. Knott (1975), Watson (1955 and 1967), Grenander (1954), and Bloomfield and Watson (1975) explore efficiency bounds in more general cases.

Equations (3.24), (3.25), and (3.26) can be used to establish pointwise convergence of the ordinary and weighted least squares estimates: $\hat{\alpha}_{\text{OLS}} - \hat{\alpha}_{\text{GGLS}} \rightarrow 0$ and $\hat{\mu}_{\text{OLS}} - \hat{\mu}_{\text{GGLS}} \rightarrow 0$ as $n \rightarrow \infty$ for any fixed $\phi$ with probability one. This upgrades the convergence of estimates to the almost sure mode; this result is therefore stronger than that of (3.20).

The relative inefficiency of ordinary least squares with varying $\phi$ is also seen; in particular, (3.25) shows that

$$
\lim_{\phi \uparrow 1} \frac{w_{\text{EP}}}{w_{\text{OLS}} + w_{\text{EP}}} = 1, \quad (3.30)
$$

for any fixed $n$. \quad \square
A semi-explicit expression can be obtained for \( \hat{\alpha}_{\text{GLS}} \) in generality. In particular, some detailed computations from the normal equations will produce

\[
\hat{\alpha}_{\text{GLS}} = \frac{\sum_{t=1}^{n} \left[ \sum_{r=1}^{n} \left( m_r, m_{r,t} - m_{r,t} \right) (r - \bar{t}) \right] X_t}{\sum_{t=1}^{n} \left[ \sum_{r=1}^{n} \left( m_r, m_{r,t} - m_{r,t} \right) (r - \bar{t}) \right] (t - \bar{t})} \tag{3.31}
\]

and

\[
\hat{\mu}_{\text{GLS}} = \frac{\sum_{t=1}^{n} \left[ \sum_{j=1}^{n} \left( \sum_{r=1}^{n} \left( m_r, m_{j,r} - m_{j,r} \right)(r - \bar{t}) \right) (j - \bar{t}) \right] X_t}{\sum_{t=1}^{n} \left[ \sum_{r=1}^{n} \left( m_r, m_{r,t} - m_{r,t} \right) (r - \bar{t}) \right] (t - \bar{t})} - \hat{\alpha}_{\text{GLS}} \bar{t}, \tag{3.32}
\]

where \( m_{r,t} \) is the \((r, t)th\) entry of \( \Gamma_n^{-1} \). The dot notation indicates column or row summation:

\[
\begin{align*}
m_{., t} &= \sum_{r=1}^{n} m_{r,t}, \\
m_{r,.} &= \sum_{t=1}^{n} m_{r,t}, \\
m_{., .} &= \sum_{t=1}^{n} \sum_{r=1}^{n} m_{r,t}. \tag{3.33}
\end{align*}
\]

One should compare the structure of (3.31) and (3.32) to that in (3.2) and (3.3).

**Example 3.3.** Suppose that \( \{\epsilon_t\} \) is an invertible first order moving average satisfying

\[
\epsilon_t = Z_t + \theta Z_{t-1}, \tag{3.34}
\]

where \( \{Z_t\} \) is mean zero white noise with variance \( \text{var}(Z_t) \equiv \sigma^2 \). Invertibility of the model implies that \(|\theta| < 1\). In this case, \( \Gamma_n \) is tridiagonal (cf. Brockwell and Davis 1991) and one can obtain the explicit expression

\[
m_{r,t} = (-\theta)^{t-r} \frac{(1 - \theta^{2r})(1 - \theta^{2(n+1-t)})}{(1 - \theta^2)(1 - \theta^{2(n+1)})} \tag{3.35}
\]

for \( r \leq t \) and \( m_{r,t} = m_{t,r} \) for \( r > t \). Using (3.35) in (3.33), we obtain

\[
\begin{align*}
m_{., t} &= \frac{(1 - (-\theta)^{n+1-t})(1 - (-\theta)^t)}{(1 + \theta)^2(1 + (-\theta)^{n+1})}, \tag{3.36} \\
m_{r,.} &= \frac{n(1 + \theta)(1 + (-\theta)^{n+1}) + 2\theta(1 - (-\theta)^n)}{(1 + \theta)^3(1 + (-\theta)^{n+1})}, \tag{3.37}
\end{align*}
\]
and \( m_{-r} = m_{r} \). This explicitly identifies \( \hat{\alpha}_{\text{GLS}} \) and \( \hat{\mu}_{\text{GLS}} \) for a first order moving-average \( \{\epsilon_t\} \).

3.4 Asymptotic Properties and Confidence Intervals

We now return to the general setting of (3.1) and consider asymptotic properties of \( \hat{\mu}_{\text{OLS}} \) and \( \hat{\alpha}_{\text{OLS}} \). Suppose that (3.21) holds. Applying the dominated convergence theorem in (3.10) and (3.11) gives

\[
\lim_{n \to \infty} \left( \frac{n(n+1)(n-1)}{12} \right) \text{var}(\hat{\alpha}_{\text{OLS}}) = 2\pi f(0)
\]

(3.38)

and

\[
\lim_{n \to \infty} n\text{var}(\hat{\mu}_{\text{OLS}}) = 8\pi f(0).
\]

(3.39)

Hence, asymptotic variances of ordinary least squares estimates are proportional to \( f(0) \); in particular, the spectral density appears only at frequency zero. Bloomfield and Nychka (1992) provide an excellent discussion of this point and its ramifications to global warming trend studies.

Asymptotic normality of the parameter estimates can also be established. In particular, Theorem 9.1.1 of Fuller (1996) can be used to extract the following result.

**Theorem 3.1.** Suppose that \( \{\epsilon_t\} \) has the causal linear representation

\[
\epsilon_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j},
\]

(3.40)

where \( \{Z_t\} \) are independent and identically distributed with mean zero and variance \( \sigma^2 \in (0, \infty) \) and that \( \sum_{j=0}^{\infty} |\psi_j| < \infty \). Then distributional convergence to normality takes place as \( n \to \infty \):

\[
\begin{pmatrix} n^{1/2} & 0 \\ 0 & (\sum_{t=1}^{n}(t-\bar{t})^2)^{1/2} \end{pmatrix} \begin{pmatrix} \hat{\mu}_{\text{OLS}} \\ \hat{\alpha}_{\text{OLS}} \end{pmatrix} \overset{D}{\to} N \left( \begin{pmatrix} \mu \\ \alpha \end{pmatrix}, 2\pi f(0) \begin{pmatrix} 4 & -3^{1/2} \\ -3^{1/2} & 1 \end{pmatrix} \right).
\]

(3.41)
It is worth noting that every causal autoregressive moving-average series with independent innovations satisfies the assumptions in Theorem 3.1; hence, the result is quite general.

Large-sample confidence intervals for the regression parameters can be extracted from the above asymptotic results. In particular,

$$\hat{\mu}_{\text{OLS}} \pm z_{\alpha/2} \left( \frac{8\pi f(0)}{n} \right)^{1/2} \quad \text{and} \quad \hat{\alpha}_{\text{OLS}} \pm z_{\alpha/2} \left( \frac{2\pi f(0)}{\sum_{t=1}^{n} (t - \bar{t})^2} \right)^{1/2}$$  \hspace{1cm} (3.42)

are approximate \((1 - \alpha) \times 100\%\) large sample confidence intervals for the regression parameters. Here, \(z_{\alpha/2}\) denotes the customary \(1 - \alpha/2\)th quantile of the standard normal distribution. The spectral density at frequency zero is needed to apply (3.42) in practice. Estimation of \(f(0)\) can be difficult (cf. Bloomfield and Nychka 1992).

For small samples from a Gaussian series with known autocovariances, the confidence intervals have the \(z\) form

$$\hat{\mu}_{\text{OLS}} \pm z_{\alpha/2} \text{var}(\hat{\mu}_{\text{OLS}})^{1/2} \quad \text{and} \quad \hat{\alpha}_{\text{OLS}} \pm z_{\alpha/2} \text{var}(\hat{\alpha}_{\text{OLS}})^{1/2},$$  \hspace{1cm} (3.43)

where the variances in (3.43) are computed from (3.10) and (3.11).

In the more practical setting where the autocovariances in \(\{\epsilon_t\}\) are unknown, one can substitute estimates of \(\gamma(\cdot)\) into (3.10) and (3.11) and use the \(t\)-based interval

$$\hat{\alpha}_{\text{OLS}} \pm t_{\alpha/2} \text{var}(\hat{\alpha}_{\text{OLS}})^{1/2}.$$  \hspace{1cm} (3.44)

Here, \(t_{\alpha/2}\) denotes the \((1 - \alpha/2)\)th quantile of the \(t\) distribution. The appropriate degrees of freedom to use in the \(t\) distribution, however, is an unresolved issue. Nychka et al. (2000) consider this problem with the first order autoregressive \(\{\epsilon_t\}\) in Example 3.2, but with \(\phi\) unknown. For inferences involving the trend slope \(\alpha\), they suggest replacing the customary \(n - 2\) degrees of freedom with the equivalent degrees of freedom \(n_e - 2\), where \(n_e\) is

$$n_e = n \left( \frac{1 - \hat{\rho}(1) - 0.68/n^{1/2}}{1 + \hat{\rho}(1) + 0.68/n^{1/2}} \right),$$  \hspace{1cm} (3.45)
and

\[ \hat{\rho}(1) = \frac{\sum_{t=1}^{n-1} R_t R_{t+1}}{\sum_{t=1}^{n} R_t^2} \overset{\text{def}}{=} \hat{\phi}_{\text{NND}} \]  

(3.46)

is the non-negative definite lag one sample autocorrelation of the residual series \( \{R_t\} = \{X_t - \hat{\mu}_{\text{OLS}} - \hat{\alpha}_{\text{OLS}} t\} \). The justification for the general form of (3.45) stems from an asymptotic correction involving (3.10): use \( \gamma(h) = \sigma^2 \phi^h/(1 - \phi^2) \) for \( h \geq 0 \), the dominated convergence theorem (\( |w_j| \leq 1 \) for all \( j \)), and \( w_j \to 1 \) as \( n \to \infty \) for each fixed \( j \) to get

\[ \lim_{n \to \infty} \frac{\gamma(0)/\sum_{t=1}^{n}(t - \bar{t})^2}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{1 - \phi}{1 + \phi}. \]  

(3.47)

In the left hand side of (3.47), the numerator is the variance of \( \hat{\alpha}_{\text{OLS}} \) when \( \{\epsilon_t\} \) is uncorrelated with variance \( \gamma(0) \) and the denominator is the variance of \( \hat{\alpha}_{\text{OLS}} \) computed from (3.10) under the AR(1) model. Hence, we interpret \( n_e \) as the number of independent observations with variance \( \gamma(0) \) (see Mitchell et al. 1966 for further discussion). Nychka et al. (2000) justify the \( 0.68/n^{1/2} \) term in (3.45) entirely on simulation, stating that the term helps correct for bias in the sampling distribution of \( \hat{\phi} \). Nychka et al. (2000) also point out that likelihood methods do not significantly improve upon least squares methods.

For a fixed \( n \), the moment results derived earlier can be used to improve accuracy. From (3.47), a better estimate of \( n_e \) for a finite sample size \( n \) is

\[ n_e = n \left( 1 + 2 \sum_{j=1}^{n-1} w_j \phi^j \right)^{-1}. \]  

(3.48)

Using (3.13) in (3.48) and performing a tedious algebraic computation gives

\[ n_e = n \left[ 1 + \frac{2\phi}{n(n + 1)(n - 1)(1 - \phi)^{\frac{3}{2}}} \left( \sum_{i=0}^{n+2} \gamma_i \phi^i \right) \right]^{-1}. \]  

(3.49)
where the only non-zero $\gamma_i$ in (3.49) are

\begin{align*}
\gamma_0 &= (n + 1)(n - 1)(n - 3), \\
\gamma_1 &= -3[n(n^2 - 1) - 2(n^2 + 1)], \\
\gamma_2 &= 3(n + 1)(n - 1)^2, \\
\gamma_3 &= -n(n + 1)(n - 1), \\
\gamma_n &= -3(n + 1)^2, \\
\gamma_{n+1} &= 6(n + 1)(n - 1), \\
\gamma_{n+2} &= -3(n - 1)^2.
\end{align*}

As noted in Nychka et al. (2000), bias in the estimates of $\phi$ and $\sigma^2$ are important to take into account with small $n$. Tjøstheim and Paulsen (1983) suggest using the bias correction $\hat{\phi}_{BC} = \hat{\rho}(1) + n^{-1}[1 + 4\hat{\rho}(1)]$ (round $\hat{\phi}_{BC}$ to 1 or $-1$ in cases where this estimate falls outside the causal range $|\phi| < 1$). Fuller (1996, Chapter 6) discusses bias corrections of general autocorrelation estimates.

A simulation was conducted to further compare the above confidence intervals. Table 3.1 summarizes the results. There, the width of the Nychka et al. (2000) interval

$$\hat{\alpha}_{OLS} \pm t_{\alpha/2} \left( \frac{\sum_{t=1}^{n} R_t^2 / (n_e - 2)}{\sum_{t=1}^{n} (t - \bar{t})^2} \right)^{1/2},$$

with $n_e$ as in (3.45) was compared to the width of the interval in (3.44) with $n_e$ as in (3.48) and the standard error

$$\text{var}(\hat{\alpha}_{OLS})^{1/2} = \left( \frac{\hat{\sigma}_{BC}^2}{1 - \hat{\phi}_{BC}^2} \right) \frac{1 + 2 \sum_{j=1}^{n-1} w_j \hat{\phi}_{BC}^j}{\sum_{t=1}^{n} (t - \bar{t})^2},$$

where

$$\hat{\sigma}_{BC}^2 = \frac{\sum_{t=2}^{n} (R_t - \hat{\phi}_{BC} R_{t-1})^2}{n - 3}$$

is a biased corrected estimate of $\sigma^2$. 


Table 3.1: Simulation Length Comparison of the Two Methods

<table>
<thead>
<tr>
<th>n</th>
<th>φ</th>
<th>EXACT Method Length</th>
<th>ADJUSTED Method Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>0.00</td>
<td>0.113 (0.930)</td>
<td>0.134 (0.968)</td>
</tr>
<tr>
<td>25</td>
<td>0.25</td>
<td>0.150 (0.920)</td>
<td>0.187 (0.950)</td>
</tr>
<tr>
<td>25</td>
<td>0.50</td>
<td>0.225 (0.901)</td>
<td>0.594 (0.937)</td>
</tr>
<tr>
<td>25</td>
<td>0.75</td>
<td>0.385 (0.857)</td>
<td>5.423 (0.912)</td>
</tr>
<tr>
<td>25</td>
<td>0.90</td>
<td>0.491 (0.787)</td>
<td>12.814 (0.879)</td>
</tr>
<tr>
<td>50</td>
<td>0.00</td>
<td>0.039 (0.940)</td>
<td>0.044 (0.966)</td>
</tr>
<tr>
<td>50</td>
<td>0.25</td>
<td>0.052 (0.935)</td>
<td>0.057 (0.956)</td>
</tr>
<tr>
<td>50</td>
<td>0.50</td>
<td>0.078 (0.927)</td>
<td>0.092 (0.951)</td>
</tr>
<tr>
<td>50</td>
<td>0.75</td>
<td>0.161 (0.905)</td>
<td>0.836 (0.943)</td>
</tr>
<tr>
<td>50</td>
<td>0.90</td>
<td>0.303 (0.852)</td>
<td>8.324 (0.925)</td>
</tr>
<tr>
<td>100</td>
<td>0.00</td>
<td>0.0136 (0.945)</td>
<td>0.0148 (0.963)</td>
</tr>
<tr>
<td>100</td>
<td>0.25</td>
<td>0.0182 (0.943)</td>
<td>0.0195 (0.957)</td>
</tr>
<tr>
<td>100</td>
<td>0.50</td>
<td>0.0274 (0.940)</td>
<td>0.0301 (0.956)</td>
</tr>
<tr>
<td>100</td>
<td>0.75</td>
<td>0.0560 (0.930)</td>
<td>0.0726 (0.958)</td>
</tr>
<tr>
<td>100</td>
<td>0.90</td>
<td>0.142 (0.900)</td>
<td>3.073 (0.952)</td>
</tr>
<tr>
<td>500</td>
<td>0.00</td>
<td>0.00121 (0.950)</td>
<td>0.00126 (0.958)</td>
</tr>
<tr>
<td>500</td>
<td>0.25</td>
<td>0.00162 (0.948)</td>
<td>0.00167 (0.955)</td>
</tr>
<tr>
<td>500</td>
<td>0.50</td>
<td>0.00243 (0.949)</td>
<td>0.00253 (0.957)</td>
</tr>
<tr>
<td>500</td>
<td>0.75</td>
<td>0.00488 (0.947)</td>
<td>0.00523 (0.960)</td>
</tr>
<tr>
<td>500</td>
<td>0.90</td>
<td>0.0124 (0.942)</td>
<td>0.0152 (0.970)</td>
</tr>
</tbody>
</table>

For each sample size and correlation parameter, 100000 time series with AR(1) Gaussian errors were simulated. Confidence intervals were evaluated at 95% level. The simulation performance of the confidence intervals are invariant over $\mu$, $\alpha$, and $\sigma^2$. So we chose $\mu = 0$, $\alpha = 1$, and $\sigma^2 = 1$. Average confidence interval lengths of the two methods, referred to as “EXACT” for the (3.48) and (3.52) based interval and as “ADJUSTED” for the (3.51) based interval, were compared with empirical coverage probabilities in parentheses. Under the true coverage probability of 0.95, the standard error of coverage probability is $\sqrt{(0.95)(1-0.95)/100000} = 0.00069$. 
Table 3.1 compares average confidence interval lengths (two times \( \pm \) error margins) of the two methods with empirical probability coverages in parentheses. A 95% level of significance was used; \( \{ \epsilon_t \} \) was simulated as Gaussian. One hundred thousand entries were conducted for each table entry; hence, uncertainty due to simulation is minimal. (The standard error of coverage probability is about 0.0007.) The parameters chosen for the simulation were \( \mu = 0, \alpha = 1, \) and \( \sigma^2 = 1. \) The sampling properties are invariant (mathematically) over the choice of \( \mu \) and \( \alpha. \)

Table 3.1 shows that the average length of the moment corrected interval in (3.48) and (3.52) is smaller than that of the (3.51)-based interval, dramatically so when \( n \) is small and \( \phi \) is close to unity. The (3.51)-based interval performs poorly in some of the simulations because of the large frequency at which (3.45) returns the physical impossibility \( n_e < 2. \) In such cases, we set \( n_e = 3, \) making the (3.51)-based interval seem better than it truthfully is. Equivalent degrees of freedom computed in (3.48) do not suffer from such structural defects. We note that the empirical probability of coverage of the (3.51)-based interval is typically closer to its target value of 95%. However, both methods return reasonable coverages that become more accurate with increasing \( n \) and smaller \( \phi. \) Overall, there is again some mileage to be gained via exact calculations.

3.5 Non-Monotonicity of Variances in Autocorrelation

Consider estimation of \( \mu \) in (3.1) when \( \alpha = 0. \) The ordinary least squares estimate is then \( \hat{\mu}_{\text{OLS}} = \bar{X} \) and

\[
\text{var}(\hat{\mu}_{\text{OLS}}) = n^{-1} \left[ \gamma(0) + 2 \sum_{k=1}^{n-1} (1 - k/n) \gamma(k) \right].
\] (3.54)

It follows from (3.54) that \( \text{var}(\hat{\mu}_{\text{OLS}}) \) increases with increasing autocovariance; specifically, if \( \gamma^{(1)} \) and \( \gamma^{(2)} \) are stationary autocovariance functions with \( \gamma^{(1)}(h) \geq \gamma^{(2)}(h) \) for all \( h \geq 0, \) then \( \text{var}(\hat{\mu}^{(1)}_{\text{OLS}}) \geq \text{var}(\hat{\mu}^{(2)}_{\text{OLS}}), \) where \( \bar{X}^{(1)} \) and \( \bar{X}^{(2)} \) denote sample
averages of size \( n \) from series with autocovariance functions \( \gamma^{(1)} \) and \( \gamma^{(2)} \), respectively. For example, the first order autoregression in Example 3.2 has \( \gamma(h) = \sigma^2 \phi^h/(1 - \phi^2) \) for \( h \geq 0 \); (3.54) now gives

\[
\text{var}(\hat{\mu}_{\text{OLS}}) = \frac{\sigma^2}{n(1 - \phi^2)(1 - \phi)^2} \left[ 1 - \frac{2\phi}{n} - \phi^2 + \frac{2\phi^{n+1}}{n} \right], \quad (3.55)
\]

which is nondecreasing for \( \phi \in [0, 1) \) (this monotonicity does not hold for \( \phi \in (-1, 0) \)).

The general principle is that variances increase with increasing positive autocorrelation. As the following example shows, the above monotonicity, perhaps rather surprisingly, is not true for the general model in (3.1).

**Example 3.4.** Consider estimation in (3.1) under two autocovariance structures: \( \gamma^{(1)}(h) = I\{0\}(h) \) and \( \gamma^{(2)}(h) = (1 - \rho^2) + \rho^2 I\{0\}(h) \), where \( I_A(h) \) denotes the indicator function over the set \( A \) and \( \rho \in (0, 1) \). Here, \( \gamma^{(1)} \) is a white noise autocovariance where ordinary least squares estimates have minimal variance and \( \gamma^{(2)} \) is the autocovariance of the stationary series

\[
\epsilon_t = (1 - \rho^2)^{1/2}S + \rho W_t, \quad (3.56)
\]

where \( S \) is a mean zero unit variance random shift that is assumed uncorrelated with the mean zero unit variance white noise \( \{W_t\} \). We have scaled both series to make \( \gamma^{(1)}(0) = \gamma^{(2)}(0) = 1 \) for a common basis of comparison. It should be noted that \( \gamma^{(2)}(h) \not\rightarrow 0 \) as \( h \rightarrow \infty \), nor is \( \gamma^{(2)} \) absolutely summable in the sense of (3.21).

In this setting, \( \hat{\alpha}_{\text{OLS}} = \hat{\alpha}_{\text{GLS}} \); (3.56) merely shifts means (albeit randomly) and rescales the noises in (3.1) and \( \hat{\alpha}_{\text{OLS}} \) is invariant over mean shifts in \( \{X_t\} \) and white noise rescalings. Applying (3.10) gives

\[
\text{var}(\hat{\alpha}_{\text{OLS}}) = \frac{1 + 2(1 - \rho^2)\sum_{j=1}^{n-1} w_j}{\sum_{t=1}^{n}(t - \bar{t})^2} = \frac{\rho^2}{\sum_{t=1}^{n}(t - \bar{t})^2}, \quad (3.57)
\]
where the second equality follows from (3.15). Equation (3.57) yields the perhaps surprising conclusion that

\[ \text{var}(\hat{\alpha}_{OLS}) \leq \frac{1}{\sum_{t=1}^{n}(t - \bar{t})^2}, \]  

(3.58)

the right hand side of (3.58) being the variance of the ordinary least squares trend estimate under independent and identically distributed errors with a unit variance.

As \( \hat{\alpha}_{OLS} = \hat{\alpha}_{GLS} \) the non-monotonicity of variance in autocorrelation is not attributable to use of ordinary least squares (in lieu of weighted least squares). Mathematically, the paradox is explained as follows. Let \( \{W_t\} \) be mean zero unit variance white noise and define two mean zero error sequences \( \{\epsilon_{t,1}\} = \{(1 - \rho^2)^{1/2}S + \rho W_t\} \) and \( \{\epsilon_{t,2}\} = \{\rho W_t\} \). Observe that \( \{\epsilon_{t,1}\} \) has larger autocovariances than \( \{\epsilon_{t,2}\} \) at all non-zero lags. However, the linear trends computed in (3.1) with errors \( \{\epsilon_{t,1}\} \) and \( \{\epsilon_{t,2}\} \) are identical. Hence, a larger error autocovariance is producing the same trend estimate variance. As the autocovariance structure of \( \{\epsilon_{t,1}\} \) arises in split-plot design inference (cf. Chapter 11 of Christensen 2002), the above construct should not be dismissed as totally pathological. □

3.6 References


Chapter 4

Trends in Extreme United States Temperatures

4.1 Introduction

Climate changes have enormous impacts on ecological systems. Human health and behavior, agriculture, forests, wildlife, and water resources are all heavily influenced by temperatures and their changes. This chapter presents an empirical analysis of maximum/minimum temperatures trends observed in the United States over the last 150 years from a modern statistical viewpoint.

In particular, we examine extreme temperatures in the form of monthly highs and lows (a monthly maximum, for example, is the largest of all daily high temperatures during the month), as these quantities have fundamental ramifications if even slightly changed. One merely envisions a Floridian orange grove subject to freezing spells, or deaths from heatwaves in Chicago where air conditioning is sparse, to appreciate the impact of temperature extremes. Scientific reasons to study monthly extremes are also numerous. Global warmists theorize that wintertime low temperatures should be increasing more rapidly than other seasonal temperatures (cf. Callendar 1961; Madden and Ramanathan 1980). Only a seasonally structured analysis can confirm or deny such a property. It is also plausible that climate change is more easily detected during summer months when temperature variability is low, than during winter months when temperature variability is comparatively higher. Public policy is also influenced by temperature extremes. For instance, more frequent snowstorms require more frequent road plowings. Should a geographic area be proven cooling
due to reforestation, as we conjecture later for the Southeastern United States, it may prove advantageous to mimic reforestation policies elsewhere.

Statistically, we examine linear trend estimates and realistic standard errors of their uncertainty. A simple time-homogeneous linear regression for such purposes is

\[ X_t = \mu + \alpha t + \epsilon_t, \]  

(4.1)

where \( \{\epsilon_t\} \) is mean zero random error and \( X_t \) is the observed temperature at time index \( t \). Whereas splines and/or other non-linear regression response functions are easily considered, such structures are typically more difficult to interpret, and perhaps more importantly, seldom add additional insight beyond (4.1) in resolving whether a fixed station is warming, cooling, or neither. Indeed, when regression function response form issues are viewed with the structural feature of changepoints in mind, there is little compelling reason to discard a linear paradigm. On a practical level, Lund et al. (2001) fit quadratic and sinusoidal components to several United States series of monthly averaged temperatures and conclude statistically that, by and large, such non-linear components do not add significant explanatory structure at individual stations.

Essentially, our focus is on estimates of \( \alpha \) and their uncertainty margins. The statistical estimation method of \( \alpha \) is not crucial; however, accurate (or at least realistic standard errors) are more difficult to obtain (cf. Bloomfield 1992; Bloomfield and Nychka 1992; Wikle and Cressie 2000; Lund et al. 2001; Fomby 2002). Such error margins are needed to make reliable trend inferences. This point is either summarily dismissed or not adequately addressed in many early studies in the climate literature.

Equation (4.1) has served as the basis of many trend studies of annual averaged temperatures. Bloomfield (1992) and Bloomfield and Nychka (1992) are good statistical starting points. In the climate literature, Roden (1966) estimates linear trends for Western United States stations and concludes that temperature changes are not
significant in rural regions, but that urban temperatures have significantly increased.
Jones and Kelly (1983) compute linear trends of Northern Hemisphere annual mean
temperatures for the three periods 1917–1939 (strong warming), 1940–1964 (cooling),
classic global temperature trend patterns. In addition Jones et al. (1986a, b) com-
pare annual trends estimates for the Northern and Southern Hemispheres with linear
trends from other study periods. Their results are updated through 1987 in Jones
(1988).

At a fixed station, extreme temperatures are serially correlated in time, but
not as heavily as corresponding series of monthly averages. This is not unexpected
as, intuitively, one ‘freak observation’ can make an extreme. Thus, whereas \( \{\epsilon_t\} \) in
(4.1) cannot be taken as white noise, autocorrelations in \( \{\epsilon_t\} \) are not as heavy as
those in series of annual averages. Nonetheless, maximum and minimum tempera-
tures are extremes, and extremes are more difficult to distributionally model than
averages. Extreme value modeling methods are well developed (cf. Gumbel 1958;
Horowitz 1980; Leadbetter et al. 1983; Smith 1986 and 1989; Robinson and Tawn
1995; Embrechts et al. 1999) and the physically justifiable approach of generalized
extreme value (GEV) modeling should be investigated. A good recent reference on
the topic is Coles (2001). Later, we compare a GEV-based likelihood fitting approach
with a regression approach.

Periodicities are also present in monthly extremes: winter temperatures are colder
and more variable than summer temperatures. We use care to incorporate seasonal
aspects into the analysis. This allows us to address uniformity of climate change
across varying season. Seasonal standard errors give insight to the related issue of
whether climate change is easier to detect during summer months, when temperature
variability is comparatively low.
An important feature of temperature series lies with changepoints. Station recording locations can move, the temperature shelter may be altered, or temperature recording gauges are frequently changed. Each changepoint can induce temperature mean shifts of several degrees. It is not uncommon for a fixed station to see five or more changepoints during a century of operation. Lund et al. (2001) demonstrate that changepoint information is the single most critical issue in obtaining an accurate temperature trend estimate at a fixed station. Whereas a law of large number averaging of changepoints effects to zero is expected over larger geographical regions, the appreciation of changepoints at an individual station is well understood in climatic circles (cf. Thompson 1984; Gullet et al. 1991; and Vincent 1998). Fortunately, the times of the changepoints are known and documented in this data. This allows us to account for changepoint effects in a straightforward manner.

After trend estimates and standard errors are obtained for each station, a spatial smoothing of the trends by longitude and latitude will be conducted. We select the head-banging algorithm of Hansen (1991) and Mungiole et al. (1998) for this purpose. Head-banging smoothers are especially adept at preserving edges in rough spatial fields while downweighting outliers. We use a weighted version of the head-banging algorithm with weights set inversely to standard error.

The rest of this chapter proceeds as follows. Section 4.2 discusses the data and its quality. Section 4.3 narrates the statistical methods used in the analysis. Section 4.4 uses the Lewiston, Maine station as a case study illustrating the general methods. Section 4.5 then proceeds with the spatial analysis of all stations.
4.2 The Data

Our data source for US temperatures is taken from the National Climatic Data Center’s (NCDC) United States Historical Climatology Network (USHCN) located at


This data contains 1221 stations located throughout the contiguous 48 United States and, as of this writing, is updated for observations through December of 2000.

Figure 4.1 shows time series plots of the Lewiston, Maine monthly extremes. This station will be analyzed in detail in Section 4.4. The Lewiston series begins in January of 1887 and contains 114 years of monthly data. As is common with many temperate zone environmental series, the Lewiston extremes display periodic patterns: wintertime temperatures (the troughs in Figure 4.1) are cooler and more variable than summertime temperatures (the peaks in Figure 4.1). The seasonal cycle in variability is graphically evident when comparing the year-to-year jaggedness of the summer peaks (smaller) in Figure 4.1 to their winter trough counterparts (larger). One may see an increasing time-trend in the data, but we postpone statistical quantification of this issue until Section 4.4.

Missing data is an issue. Ad-hoc screening of the individual station records was undertaken to ensure reliable findings based on a reasonably complete record. In particular, stations missing more than 2.5% of their total period of record, or stations starting observation after 1925, were discarded. After these data quality restrictions were imposed, 528 stations remained. Figure 4.2 graphically plots the spatial location of these stations and shows good spatial coverage over the contiguous 48 United States. Each station has at least 75 years of monthly data with at most 2.5% of the record missing. The first observation in each series is taken in January and the last observation is recorded in December of 2000. The longest data record is from
Figure 4.1: Monthly Maximum and Minimum Temperatures at Lewiston, Maine
Charleston City, SC (130 years) and the shortest length occurs at 4 stations (75 years).

Erroneous data entries are not problematic. Indeed, this data has been rigorously quality checked via reference station comparisons. We select the time-of-observation adjusted subset of the data so that all observations are converted to a midnight-to-midnight observation time scale. As our main focus is on temperature change, any subset of observations would serve our purpose as long as the type of data analyzed remains consistent across varying station.
The missing observations that remain are infilled with a regression model as discussed in the next section. Spline and/or time series infilling techniques are difficult to implement across varying station as almost every conceivable pattern of missing observations surfaces amongst the 528 stations. Whereas the analysis below could take into account missing data (all summations are merely over present observations), a regression infilling scheme works well, is easy to implement, and most importantly, minimizes bookkeeping that will have little influence on overall trends.

The length of the series at a fixed station is denoted by \( N = dT \) where \( d \) is the number of years of data and \( T = 12 \) is the natural period for monthly data. The exposition below takes \( d \) as a whole positive integer. As the first observation at each station always occurs during January (the starting year varies) and the last during December of 2000, this is not an issue.

4.3 Methods

In this section, we consider the statistical analysis of a series \( \{X_t\} \) from a fixed station. This series can be either monthly maximums or monthly minimums.

4.3.1 Regression Methods

The ordinary least squares (OLS) estimate for the trend \( \alpha \) in (4.1) has the familiar explicit form

\[
\hat{\alpha}_{\text{OLS}} = \frac{\sum_{i=1}^{n} (t - \bar{t})(X_t - \bar{X})}{\sum_{i=1}^{n} (t - \bar{t})^2},
\]

where \( \bar{t} = (n + 1)/2 \) and \( \bar{X} = \frac{1}{n} \sum_{t=1}^{n} X_t \) are average time and observations, respectively. The OLS location parameter estimate is \( \hat{\mu}_{\text{OLS}} = \bar{X} - \hat{\alpha}_{\text{OLS}} \bar{t} \). When \( \{\epsilon_t\} \) is autocorrelated (not white noise), \( \hat{\mu}_{\text{OLS}} \) and \( \hat{\alpha}_{\text{OLS}} \) are not the best linear unbiased estimates. However, under vast generality (\( \{\epsilon_t\} \) is time stationary with a
spectral density that is bounded away from zero), OLS estimates are asymptotically as efficient as generalized least squares estimates (cf. Grenander 1954). In short, OLS estimates are reasonable in time series settings.

There is little a priori reason to believe that temperatures are changing uniformly across all months. To investigate such structure, we formulate a periodic version of (4.1) with period $T$:

$$X_{nT+\nu} = \mu_\nu + \alpha_\nu (nT + \nu) + \epsilon_{nT+\nu}. \quad (4.3)$$

Here, the notation $X_{nT+\nu}$ denotes the observation during the $\nu$th month of year $n$ ($1 \leq \nu \leq T = 12$). The first year of observations is indexed as year $n = 0$ so that the first January observation is indexed as month 1. The notations $\{X_t\}$, $\{\epsilon_t\}$, etc. will be freely interchanged with $\{X_{nT+\nu}\}$, $\{\epsilon_{nT+\nu}\}$, etc. below; the latter notations being used when emphasis on seasonality is paramount. In (4.3), $\{\epsilon_{nT+\nu}\}$ is a mean zero periodic time series of error disturbances with period $T = 12$.

The regression in (4.3) does not take into account changepoints. Suppose that the station has seen $k$ different regimes over the observation record (i.e., there are $k - 1$ changepoints) and that the changepoints occur at the known month indices $1 < \tau_1 < \tau_2 < \cdots < \tau_{k-1} < N$. A modification of (4.3) that allows for a mean shift at each changepoint time is

$$X_{nT+\nu} = \mu_\nu + S_{nT+\nu} + \alpha_\nu (nT + \nu) + \epsilon_{nT+\nu}, \quad (4.4)$$

where the changepoint factor at month $t$ is $S_t = \Delta_j$ if $\tau_{j-1} \leq t < \tau_j$ for $1 \leq j \leq k$ ($\tau_0 = 1$ and $\tau_k = N + 1$). We take $\Delta_1 = 0$ as a baseline so that there are $k - 1$ changepoint mean shifts to estimate. As our primary interest lies with $\hat{\alpha}_\nu$ and its standard error, the location parameters $\mu_\nu$, $1 \leq \nu \leq T$, the changepoint mean shifts $\Delta_2, \ldots, \Delta_k$, and all autocovariance parameters of $\{\epsilon_t\}$ are viewed as nuisance parameters.
The OLS parameter estimates of $\mu_\nu$, $\alpha_\nu$, $1 \leq \nu \leq T$, and $\Delta_2, \ldots, \Delta_k$ cannot be explicitly derived in closed form akin to (4.2). Nonetheless, some simplification of computation can be made. The general linear models form of (4.4) is

$$\vec{X} = D\vec{\beta} + \vec{\epsilon},$$

(4.5)

where $\vec{X} = (X_1, X_2, \ldots, X_{dT})'$ is the $N = dT$ dimensional observation vector, $D$ is an $N \times (2T + k - 1)$ dimensional design matrix consisting of columns for monthly indicators, site change effects indicators, and times, $\vec{\beta} = (\mu_1, \ldots, \mu_T, \Delta_2, \ldots, \Delta_k, \alpha_1, \ldots, \alpha_T)'$ is the $(2T + k - 1)$ dimensional parameter vector, and $\vec{\epsilon} = (\epsilon_1, \ldots, \epsilon_{dT})'$ is an $N$-dimensional error vector whose covariance matrix will be denoted as $\Gamma = E[\vec{\epsilon}\vec{\epsilon}']$.

The OLS estimate of $\vec{\beta}$ takes the familiar general linear models form

$$\hat{\vec{\beta}} = (D'D)^{-1}D'\vec{X};$$

(4.6)

$\hat{\vec{\beta}}$ is unbiased for any mean zero $\{\epsilon_t\}$. The variance/covariance matrix of $\hat{\vec{\beta}}$ is

$$\text{var}(\hat{\vec{\beta}}) = (D'D)^{-1}D'TD(D'D)^{-1}.$$  

(4.7)

Some simplification in the computations can be made to isolate on $\vec{\alpha} = (\alpha_1, \ldots, \alpha_T)'$. Partition the design matrix $D$ into $D_1$ and $D_2$, where $D_1$ is the $N \times (T + k - 1)$ subcolumn matrix of $D$ whose columns consist of the $T$ monthly indicators and $k - 1$ site-change indicators and $D_2$ is the remaining $N \times T$ dimensional part of $D$ whose columns refer to the time index. The OLS estimate of $\vec{\alpha}$ can be bookkept as

$$\hat{\vec{\alpha}} = (D_2'M_1D_2)^{-1}D_2'M_1\vec{X},$$

(4.8)

where $M_1 = I_N - D_1(D_1'D_1)^{-1}D_1'$ for an $N$ dimensional identity matrix $I_N$. The covariance matrix of $\hat{\vec{\alpha}}$ is

$$\text{var}(\hat{\vec{\alpha}}) = (D_2'M_1D_2)^{-1}D_2'M_1\Gamma M_1D_2(D_2'M_1D_2)^{-1},$$

(4.9)
since $M_1$ is symmetric.

To consolidate the number of graphics presented, we summarize the 12 months into the four seasons Winter, Spring, Summer, and Fall. Winter is taken as December, January, and February; Spring as March, April, and May; Summer as June, July, and August; and Fall as September, October, and November. A Spring trend is computed as

$$\hat{\alpha}_{\text{SPR}} = \hat{\alpha}_3 + \hat{\alpha}_4 + \hat{\alpha}_5. \quad (4.10)$$

All trend estimates will be reported in degrees Celsius per Century; hence, one need not worry with denominator scalings in (4.10) to convert to annual accounting. Estimated trends for Winter, Summer, and Fall, denoted by $\hat{\alpha}_{\text{WIN}}$, $\hat{\alpha}_{\text{SUM}}$, and $\hat{\alpha}_{\text{FAL}}$, have analogous forms to (4.10). Standard errors are obtained through (4.9). For example, the Spring trend has variance

$$\text{var}(\hat{\alpha}_{\text{SPR}}) = L_{\text{SPR}}' \text{var}(\hat{\alpha}) L_{\text{SPR}}, \quad (4.11)$$

where the $12 \times 1$ dimensional Spring months indicator $L_{\text{SPR}} = (0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0)'$. Standard errors for the other seasons have an analogous structure.

Perhaps the most important trends we consider are yearly trends:

$$\hat{\alpha}_{\text{YR}} = \sum_{\nu=1}^{T} \hat{\alpha}_{\nu} = \hat{\alpha}_{\text{WIN}} + \hat{\alpha}_{\text{SPR}} + \hat{\alpha}_{\text{SUM}} + \hat{\alpha}_{\text{FAL}} \quad (4.12)$$

The variance computation of $\hat{\alpha}_{\text{YR}}$ is as above except that $L_{\text{YR}} = I_T$ is a $T$ dimensional vector whose entries are all unity. Use of the above variance expressions requires the covariance matrix $\Gamma$ of $\{\epsilon_t\}_{t=1}^N$. Elaborate schemes for estimating $\Gamma$ are not necessary as the standard error is not that sensitive to small changes in autocovariance estimates, but one should not use empirical (non-negative definite) sample averages at large lags either. A tradeoff of computational feasibility and physical reasonability employs a periodic function multiplying a stationary series: $\epsilon_t = \sigma(t)A_t$,
where \(\sigma(\cdot)\) is a positive deterministic periodic sequence with period \(T\) and \(\{A_t\}\) is a causal and invertible ARMA\((p,q)\) time series with mean zero and autocovariance \(\gamma_A(h)\) at lag \(h\). The restriction of \(\gamma_A(0) = 1\) is imposed for parameter identifiability.

Let \(R_j = \{\tau_{j-1}, \tau_{j-1} + 1, \ldots, \tau_j - 1\}\) denote the set of times when the observation record experienced regime \(j\) for \(1 \leq j \leq k\). A simple moment estimate of \(\sigma(\nu)\) is obtained by summing squared residuals:

\[
\hat{\sigma}(\nu) = \sqrt{\frac{1}{d-1} \sum_{j=1}^{k} \sum_{n: nT + \nu \in R_j} \left[ X_{nT + \nu} - \hat{\mu}_\nu - \hat{S}_{nT + \nu} - \hat{\alpha}_\nu(nT + \nu) \right]^2.}
\] (4.13)

The series \(\{A_t\}\) is now estimated via

\[
\hat{A}_{nT + \nu} = \frac{X_{nT + \nu} - \hat{\mu}_\nu - \hat{S}_{nT + \nu} - \hat{\alpha}_\nu(nT + \nu)}{\hat{\sigma}(\nu)}.}
\] (4.14)

Observe that \(\{\hat{A}_t\}\) has a sample mean of zero and an approximately unit variance. An ARMA\((p,q)\) model is then fitted to \(\{\hat{A}_t\}\) to obtain an estimate of \(\gamma_A(\cdot)\) which we denote by \(\hat{\gamma}_A(h)\) at lag \(h\). The ARMA model having the smallest AICC statistic over all possible orders satisfying \(p + q \leq 5\) was selected (cf. Brockwell and Davis 1991 for standard ARMA model fitting). Again, such autocovariance estimates are not overly critical.

The individual elements of \(\Gamma\) are hence estimated via

\[
\hat{\text{cov}}(X_{nT + \nu}, X_{mT + \kappa}) = \hat{\sigma}(\nu)\hat{\sigma}(\kappa)\hat{\gamma}_A(|(n - m)T + \nu - \kappa|)
\] (4.15)

for \(1 \leq \nu, \kappa \leq T\). A standard error of \(\hat{\alpha}\) is now obtained through (4.9).

### 4.3.2 Extreme Value Methods

Extreme value trend modeling techniques for time-stationary series focus on the generalized extreme value (GEV) cumulative distribution function

\[
G(z) = \exp \left\{- \left[ 1 + \xi \left( \frac{z - m}{\nu} \right) \right]^{-1/\xi} \right\},
\] (4.16)
for marginal distributions. The support set of the GEV family is taken as \( \{ z : 1 + \xi (z - m)/\nu > 0 \} \), where \( \nu \) and \( \xi \) satisfy the constraints \( \nu > 0 \), and \( \xi < 1/2 \) (needed for finite variances). The Gumbel, Fréchet, and Weibull distributions are contained in the GEV family. For mathematical formalism, the case \( \xi = 0 \) is interpreted as a limit as \( \xi \rightarrow 0 \) in (4.16).

When \( \xi < 1 \), the mean of a random variate \( Y \) having the GEV distribution is

\[
E[Y] = m + \frac{\nu}{\xi} (\Gamma(1 - \xi) - 1).
\]  

To allow for time changes in the distribution of the series being studied, the GEV location and scale parameters are allowed to vary in time:

\[
m_{nT+\nu} = \mu_{\nu} + S_{nT+\nu} + \alpha_{\nu}(nT + \nu),
\]

where \( \mu_{\nu} \) and \( S_t \) are seasonal mean and changepoint parameters, respectively, akin to those discussed in the regression subsection, and

\[
u_{nT+\nu} = \kappa_0 + \kappa_1 \cos \left( \frac{2\pi(\nu - \eta)}{T} \right),
\]

where \( \kappa_0, \kappa_1, \) and \( \eta \) are free parameters. We take the advice of Coles (2001) and keep \( \xi \) constant in time \( t \): \( \xi_{nT+\nu} \equiv \xi \). These parametrizations capture periodicity and changepoint features with reasonable parsimony. The one-parameter cosine wave in (4.19) and constant \( \xi \) work well in for modeling a wide range of periodic extremes; the interested reader is referred to Coles (2001) for other environmental examples.

The log-likelihood of \( \{ X_t \} \) can be numerically maximized to obtain maximum likelihood estimates of \( \alpha_{\nu} \) and \( \mu_{\nu} \) for \( 1 \leq \nu \leq T, \) \( \Delta_2, \ldots, \Delta_k, \) \( \xi, \) and \( \kappa_0, \kappa_1, \) and \( \eta \) as well as their standard errors. Such a likelihood is based upon independence. Whereas such properties do not hold in this setting, the techniques of blocking and/or thresholding heavily correlated extremes (cf. Smith 1989) do not add additional precision here. Estimates of trends for the four seasons and their standard errors are
obtained analogously to the methodology discussed in the regression section. The
GEV performance will be compared with regression methods in Section 4.4.

4.3.3 Spatial Smoothing

For each station, season, and extreme type (maximum or minimum), a trend estimate
and its standard error are obtained. To graphically summarize these figures, we
smooth them by longitude and latitude with the weighted head-banging approach
described in Mungiole et al. (1998).

Briefly, the weighted head-banging algorithm is a nonparametric spatial smoother
that is especially adept with rough (highly variable) spatial fields. Head-banging
techniques make local averages by classing stations into many sets of neighboring
triples. Head-banging methods are effective in preserving edge features in the spatial
field while simultaneously downweighting the effect of outliers. The algorithm variant
we use weights the estimated trends inversely to their standard error; hence, more
questionable trends have less influence on the end trend estimate. The head-banging
algorithm is attractive in contiguous United States settings where most ‘geographical
boundaries’ (the Appalachian, Sierra, and Rocky Mountain Ranges and the Atlantic
and Pacific Ocean coasts) are oriented more North-South. Here, station triples are
constructed in a more North-South orientation.

Application of the head-banging algorithm yields a spatially smoothed trend
estimate at each station in the study. Such averaging implicitly accounts for spatial
correlations in the trend estimate field. As an end step, the Geographic Information
Systems (GIS) plotting package was applied to the head-banging smoothed trends.
The end-product is aesthetic and preserves the general structure in the raw trends.
Kriging and hierarchical Bayesian smoothings of the trends could be considered,
but are more elaborate, less compatible with the GIS plotting package, and most
importantly, do not offer better insight into the spatial trend structure of the field.
4.4 Lewiston, Maine Temperatures

This section studies the Lewiston, Maine monthly extremes as a case study. We concentrate on Lewiston minimum temperatures; maximum temperatures follow a similar pattern. The Lewiston record contains 114 years of monthly data spanning January 1887 — December 2000. Five changepoints are noted in the NCDC’s data log: February 1906, July 1960, August 1961, June 1964, and November 1989.

Monthly regression-based trend estimates and their standard errors are listed in Table 4.1 below. The $z$-scores are the trend estimates divided by their standard errors and are provided for rough intuition. The figures in Table 4.1 suggest that Lewiston minimum temperatures are indeed warming, and doing so quite uniformly over the seasons. Notice that the standard errors and estimated trends are larger during winter months. For this station, an ARMA($3,2$) model was fitted for $\{\hat{A}_t\}$. Figure 4.3 shows the seasonally adjusted series $\{\hat{A}_t\}$ and its sample autocorrelations; formal diagnostics do not detect significant departures from white noise.

Table 4.1: Monthly Lewiston Regression Estimates

<table>
<thead>
<tr>
<th>Month</th>
<th>Trend ($^\circ$C/century)</th>
<th>Standard Error ($^\circ$C/century)</th>
<th>$z$-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>3.6561</td>
<td>0.9065</td>
<td>4.0334</td>
</tr>
<tr>
<td>February</td>
<td>5.4745</td>
<td>0.8848</td>
<td>6.1872</td>
</tr>
<tr>
<td>March</td>
<td>3.9715</td>
<td>0.7472</td>
<td>5.3152</td>
</tr>
<tr>
<td>April</td>
<td>2.9800</td>
<td>0.5920</td>
<td>5.0336</td>
</tr>
<tr>
<td>May</td>
<td>2.3805</td>
<td>0.5861</td>
<td>4.0615</td>
</tr>
<tr>
<td>June</td>
<td>2.4377</td>
<td>0.5701</td>
<td>4.2758</td>
</tr>
<tr>
<td>July</td>
<td>2.2714</td>
<td>0.5703</td>
<td>3.9831</td>
</tr>
<tr>
<td>August</td>
<td>2.8016</td>
<td>0.5912</td>
<td>4.7389</td>
</tr>
<tr>
<td>September</td>
<td>2.4143</td>
<td>0.5870</td>
<td>4.1127</td>
</tr>
<tr>
<td>October</td>
<td>2.4369</td>
<td>0.6128</td>
<td>3.9767</td>
</tr>
<tr>
<td>November</td>
<td>3.7531</td>
<td>0.6288</td>
<td>5.9683</td>
</tr>
<tr>
<td>December</td>
<td>4.1629</td>
<td>0.8922</td>
<td>4.6658</td>
</tr>
</tbody>
</table>
Figure 4.3: Lewiston ARMA(3,2) Residuals and Sample Autocorrelations
Regression-based trends for the four seasons and their standard errors are reported in Table 4.2 below. These figures suggest that winter minimum temperatures have warmed slightly more than summer minimums. The annual trend estimate is $\hat{\alpha}_{\text{yr}} = 3.2284 \pm 0.5304^\circ\text{C/century}$ (all error margins quoted represent one standard error).

Table 4.2: Seasonal Lewiston Regression Estimates

<table>
<thead>
<tr>
<th>Season</th>
<th>Trend ($^\circ\text{C/century}$)</th>
<th>Standard Error ($^\circ\text{C/century}$)</th>
<th>$z$-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>4.4312</td>
<td>0.7186</td>
<td>6.1666</td>
</tr>
<tr>
<td>Spring</td>
<td>3.1107</td>
<td>0.5696</td>
<td>5.4616</td>
</tr>
<tr>
<td>Summer</td>
<td>2.5036</td>
<td>0.5362</td>
<td>4.6695</td>
</tr>
<tr>
<td>Fall</td>
<td>2.8681</td>
<td>0.5543</td>
<td>5.1743</td>
</tr>
</tbody>
</table>

Turning to GEV modeling of the Lewiston minima as discussed in Section 4.3.2, we obtain the monthly trend estimates in Table 4.3.

Table 4.3: Monthly Lewiston GEV Estimates

<table>
<thead>
<tr>
<th>Month</th>
<th>Trend ($^\circ\text{C/century}$)</th>
<th>Standard Error ($^\circ\text{C/century}$)</th>
<th>$z$-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>2.9176</td>
<td>0.7203</td>
<td>4.0507</td>
</tr>
<tr>
<td>February</td>
<td>4.9865</td>
<td>0.6974</td>
<td>7.1501</td>
</tr>
<tr>
<td>March</td>
<td>3.3553</td>
<td>0.6461</td>
<td>5.1934</td>
</tr>
<tr>
<td>April</td>
<td>2.4413</td>
<td>0.5739</td>
<td>4.2535</td>
</tr>
<tr>
<td>May</td>
<td>1.8524</td>
<td>0.4861</td>
<td>3.8107</td>
</tr>
<tr>
<td>June</td>
<td>1.9125</td>
<td>0.4175</td>
<td>4.5811</td>
</tr>
<tr>
<td>July</td>
<td>1.2756</td>
<td>0.3814</td>
<td>3.3448</td>
</tr>
<tr>
<td>August</td>
<td>2.2098</td>
<td>0.4312</td>
<td>5.1248</td>
</tr>
<tr>
<td>September</td>
<td>1.9280</td>
<td>0.4921</td>
<td>3.9182</td>
</tr>
<tr>
<td>October</td>
<td>1.7657</td>
<td>0.5646</td>
<td>3.1275</td>
</tr>
<tr>
<td>November</td>
<td>3.2282</td>
<td>0.6777</td>
<td>4.7634</td>
</tr>
<tr>
<td>December</td>
<td>3.5864</td>
<td>0.7047</td>
<td>5.0892</td>
</tr>
</tbody>
</table>

GEV estimates of seasonal trends and standard errors are listed in Table 4.4.
Table 4.4: Seasonal Lewiston GEV Estimates

<table>
<thead>
<tr>
<th>Season</th>
<th>Trend (°C/century)</th>
<th>Standard Error (°C/century)</th>
<th>z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>3.8302</td>
<td>0.4805</td>
<td>7.9714</td>
</tr>
<tr>
<td>Spring</td>
<td>2.5497</td>
<td>0.4167</td>
<td>6.1188</td>
</tr>
<tr>
<td>Summer</td>
<td>1.7993</td>
<td>0.3487</td>
<td>5.1595</td>
</tr>
<tr>
<td>Fall</td>
<td>2.3073</td>
<td>0.4202</td>
<td>5.4907</td>
</tr>
</tbody>
</table>

The GEV annual trend is $2.6216 \pm 0.3412^\circ$C/century. Key parameters in the GEV model fit are $\hat{\xi} = -0.2791 \pm 0.0149^\circ$C, $\hat{\kappa}_0 = 1.6612 \pm 0.0338^\circ$C, $\hat{\kappa}_1 = 0.7046 \pm 0.0350^\circ$C, and $\hat{\eta} = 0.9545 \pm 0.1240$ month. All error margins represent one standard error and are computed by inverting the observed information matrix of the GEV likelihood.

Comparing regression and GEV results, the seasonal and annual trend estimates in Table 4.2 and Table 4.4 appear similar (roughly) with warming observed during all seasons. The GEV trends are somewhat smaller than the regression trends, but not markedly so. To further compare, Figure 4.4 plots the monthly trends from each method and shows good coherence.

Mathematically, we explain the agreement as follows. In the regression model, expected changes in mean are obtained from (4.4) as

$$E[X_{(n+1)T+\nu}] - E[X_{nT+\nu}] = T \alpha_\nu,$$

assuming no changepoints occur between times $(n+1)T + \nu$ and $nT + \nu$. Taking expectations in the GEV model via (4.17) yields (4.20) again. Such structure remains consistent if more elaborate periodic forms for $v_t$ were employed or if $\xi$ were allowed to be periodic.
In short, one should not expect radically different linear trend estimates from GEV and regression approaches as linear trends can be cast in terms of first moment properties (the reader is cautioned not to make similar conclusions about extreme quantiles and return levels where methods do matter). To confirm this agreement further in computational settings, monthly maximums at Fullerton, North Dakota, and Greenville, Ohio were also examined. Regression and GEV methods again gave very similar trends. Because of this, we use the more simplistic regression approach in what follows.
4.5 Results for all Stations

This section moves to the study of all trends in the 528 stations. Figure 4.5 shows boxplots of the regression trend estimates for each of the four seasons. The maximums are in the top graphic and the minimums in the bottom graphic. Observe that median trends exceed zero during all seasons and that winter trends are warming slightly more than the other three seasons. Variability is largest in the winter trends and smallest with summer trends. Minimum trends are slightly more variable than maximum trends. Figure 4.6 shows boxplots of ‘z-scores’ (trend divided by standard error) by season and were computed from the standard errors.

Figure 4.7—Figure 4.11 display the head-banging smoothed trend estimates in maximum and minimum temperatures for Winter, Spring, Summer, and Fall. The maximums are in the top graphic on each page, the minimums in the bottom. The contours are color coded with blue representing cooling and red warming. The brightest blues represent the strongest cooling and the darkest reds the most warming. All temperature trends are converted to degrees Celsius per century for a common basis of interpretation. Such a conversion entails multiplying seasonal trends and their standard errors by 400 and annual trends and their standard errors by 100. Of course, this scaling does not change z-scores.

The Winter trends in Figure 4.7 show fairly widespread warming in general areas, but most notably in the Plains, West Coast, Oregon, West Idaho, The Northeast, The Carolinas, and Northern Midwest. Cooling maximum temperatures are confined to Southern Louisiana, Southern Texas, and Western Iowa. Slight cooling in maximums takes place in Indiana and Eastern Illinois. For minimum trends, rapid warming is clustered in The Northern Midwest, The Dakotas, Nebraska, Wyoming, The North West, West Coast, Southern Arizona, and The Northeast. Minimum temperature cooling occurs in the Southeast, with more than 1°C per century of cooling.
Figure 4.5: Boxplot of Trend Estimates by Season
Figure 4.6: Boxplot of Trend z-Scores by Season
in Southern Georgia and Louisiana. A small decreasing trend also occurs in Southern Indiana.

The Spring trends in Figure 4.8 reveal warming maximums in many areas of the country, with significantly increasing trends in the Northeast, Central Plains, West and Southwest. The Ohio River Basin Region, extending out to Iowa, shows cooling Spring high temperatures. The Spring minimums are increasing in the Central Plains, the Northern Midwest, Northeastern Oregon, and Western Idaho. Noticeable cooling of Spring lows occurs in much of the Southeast, peaking in Southern Louisiana and Southern Georgia. Some cooling of Spring minimum is also evident in Montana.

The Summer trends in Figure 4.9 reveal increasing maximum trends over Colorado, Arizona, and New Mexico and are widespread over western Rocky Mountains during Summer seasons. A clustering of decreasing maximums is evident in Indiana, eastern Ohio, southern Illinois, western Iowa, and southern Texas. Summer minimum temperatures are warming in the Central Plains, the West Coast, and the Northeast, with the largest trends appearing in the northern Midwest. Northern Montana and Southern Louisiana have slight cooling trends with summer minimums.

The Fall trends in Figure 4.10 show warming maximums in the Northeast, the Carolinas, Oregon, Colorado, New Mexico, Arizona, and the West Coast. Cooling takes place in the Midwest, the Mideast, Texas, southern Louisiana, and Montana. Fall minimum temperatures are rising in central Plains, the West, Northern Minnesota, Wisconsin, Northern Michigan, and the Northeastern cities. The minimum temperatures show cooling over a large portion of the Southeast and in Montana.

Figure 4.11 displays the spatially smoothed trends of annual maximum and minimum temperatures. This graphic can be regarded as an aggregation of the last four figures and serves to summarize overall (yearly) trends in highs and lows. Figure 4.11 shows that The Northeast, The Southern Midwest, and Oregon have strongly increasing maximums. Significantly decreasing maximum annual trends
cluster in Western Iowa, South Texas, and Northern Georgia. Slight cooling of annual maximums is also evident in the Ohio River Basin. For annual minimum trends, significantly positive trends are present over the Western Rocky Mountains, Great Basin, and Central Plains. Northern Wisconsin shows the largest warming with an increase of more than 3°C per century. Only slight warming of annual minimums is evident in much of the country. Cooling minimums are evident in Western Texas, Louisiana, Southern Georgia, Southern Indiana, the Lower Mississippi River Basin.

4.6 Discussion

The temperature change maps in Figure 4.7—Figure 4.11 here, along with analogous maps for monthly averages in Lund et al. (2001), suggest that the Southeastern United States and the Tennessee and Ohio River Basins have cooled slightly over the period of record. We are not the first authors to note such cooling (cf. Karl et al. 1995). Though we have no firm basis to connect any physical cause with this cooling, we conjecture the cooling is rooted in reforestation. The type of tree in the reforestation may also be relevant. The Southern forest is becoming increasingly dominated by pine trees planted for lumber and paper pulp purposes. As pines are evergreen trees and not deciduous, they use more carbon dioxide on a yearly basis than the indigenous hardwood forests that were prevalent centuries ago. We note an edge demarcating cooling from warming in the annual minimum trend map very close to the southern extent of the Appalachian Mountains.

Figure 4.7—Figure 4.11 and the trend maps in Lund et al. (2001) suggest that New England, The West, and the Upper Midwest regions are warming. Whereas a North Dakota resident would consider this good news, a Utah skier may disagree. The warming in New England may be partly attributed to urban heat effects whereby concrete and asphalt impede nighttime radiational cooling. This has been conjec-
tured by many authors. It may be surprising to note that New England high temperatures have increased more than New England low temperatures (see Figure 4.11).

It is worth mentioning that the maximum and average of Gaussian series are asymptotically independent under very general conditions on a stationary sequence (see McCormick and Qi 2000). Using this as loose guidance (temperatures are not Gaussian), one should not necessarily expect monthly extremes and monthly averages to follow the same trend patterns. In particular, any complete study of temperature changes should examine both monthly averages and extremes. We reiterate that study of upper distributional quantiles, such as that in DeGaetano and Allen (2002), should be based on extreme valued methods.

It would be interesting to regress the estimated trends obtained here on factors such as forestation and pavement indices, local population counts, altitude of recording station, and perhaps even longitude and latitude of station. Such information would be very useful in physically explaining observed climate change. Some of these factors (altitude of recording station for example) are readily available, but others are not. Forestation indices, for example, are difficult to obtain during pre-satellite times. We leave such explorations to future work.

Acknowledgements. The authors acknowledge support from National Science Foundations grants DMS 0071383 and DMS 0304407. We are indebted to Mrs. QiQi Lu for making the color coded GIS contour maps.
Winter Maximum Trends

Winter Minimum Trends

Figure 4.7: Smoothed Winter Trends
Spring Maximum Trends

Spring Minimum Trends

Figure 4.8: Smoothed Spring Trends
Figure 4.9: Smoothed Summer Trends

Summer Maximum Trends

Summer Minimum Trends
Fall Maximum Trends

Fall Minimum Trends

Figure 4.10: Smoothed Fall Trends
Figure 4.11: Smoothed Annual Trends
4.7 References


This chapter proposes some avenues for future research.

5.1 Simple Linear Regression in a Periodic Setting

Consider the classical simple linear regression
\[ X_t = \mu + \alpha t + \epsilon_t, \quad (5.1) \]
where \( \{\epsilon_t\} \) is mean zero white noise; specifically, \( E[\epsilon_t] \equiv 0 \), \( \text{Cov}(\epsilon_t, \epsilon_s) = 0 \) if \( t \neq s \), and \( \text{Var}(\epsilon_t) \equiv \sigma^2 \). As little generality is gained by considering the regression \( X_i = \mu + \alpha t_i + \epsilon_i \) with general \( t_i \) (merely replace \( \sum_{t=1}^{n}(t - \bar{t})^2 \) in the discourse below with \( \sum_{i=1}^{n}(t_i - \bar{t})^2 \)), we focus on (5.1) and some of its variants in what follows.

Suppose that one has observations \( X_1, \ldots, X_N \). The ordinary least squares (OLS) estimates of the trend and location parameters in (5.1) are
\[ \hat{\alpha} = \frac{\sum_{t=1}^{N}(X_t - \bar{X})(t - \bar{t})}{\sum_{t=1}^{N}(t - \bar{t})^2} \text{ and } \hat{\mu} = \bar{X} - \hat{\alpha}\bar{t}, \quad (5.2) \]
where \( \bar{X} = N^{-1} \sum_{t=1}^{N} X_t \) and \( \bar{t} = (N + 1)/2 \) are the average series and observation times. The OLS estimates minimize the sum of squares \( \sum_{t=1}^{N}(X_t - \mu - \alpha t)^2 \) over \( \alpha \) and \( \mu \). The denominator in (5.2) can be explicitly written \( \sum_{t=1}^{N}(t - \bar{t})^2 = N(N + 1)(N - 1)/12 \).

If \( \{\epsilon_t\} \) is a stationary mean zero time series, OLS estimates are still very good (note that stationarity implies that \( \text{var}(\epsilon_t) \) is constant in \( t \)). Although the OLS
estimates do not account for autocorrelations in \( \{\epsilon_t\} \), Grenander (1954) shows that
they have the same asymptotic variance as best linear unbiased estimates (also called
BLUEs or weighted least squares estimates) as long as the spectral density of \( \{\epsilon_t\} \)
exists and is bounded away from zero (this is almost always the case). Hence, there
is little reason to diverge from the explicit and convenient expressions in (5.2) in a
stationary setting.

The analogous version of a simple linear regression in a periodic environment is

\[
X_{nT+\nu} = \mu_\nu + \alpha_\nu(nT + \nu) + \epsilon_{nT+\nu},
\]

where \( \{\epsilon_t\} \) is a mean zero periodic time series with known period \( T \). Note that
trend and location parameters are now allowed to depend on the season \( \nu \). The OLS
estimates of \( \alpha_\nu \) and \( \mu_\nu \) are

\[
\hat{\alpha}_\nu = \frac{\sum_{n=0}^{d-1}(nT + \nu - \bar{t}_\nu)(X_{nT+\nu} - \bar{X}_\nu)}{\sum_{n=0}^{d-1}(nT + \nu - \bar{t}_\nu)^2} \quad \text{and} \quad \hat{\mu}_\nu = \bar{X}_\nu - \hat{\alpha}_\nu \bar{t}_\nu,
\]

where \( \bar{X}_\nu = d^{-1}\sum_{n=0}^{d-1}X_{nT+\nu} \) and \( \bar{t}_\nu = d^{-1}\sum_{n=0}^{d-1}(nT + \nu) = \nu + (d - 1)/2 \) are the
series and observation time averages for season \( \nu \). The notation here uses \( d = N/T \)
as the number of observed data cycles and assumes that \( d \) is integral for simplicity
of exposition only.

There is reason to question the optimality of OLS estimates in a periodic setting.
Viewing the matter intuitively, observe from (5.4) that \( \hat{\alpha}_\nu \) is a linear combination of
series values from season \( \nu \) only. The reasoning is that the series values outside of
season \( \nu \) contain some information about \( \alpha_\nu \) and \( \mu_\nu \) when \( \{\epsilon_t\} \) is correlated.

As we have seen, the model in (5.3) has climate ramifications as it makes a
reasonable model for monthly temperatures. Whereas the true temperature response
of the Earth is undoubtedly non-linear in time (we would have fried or frozen long
ago, if linear), a linear trend is geologically justifiable over a short time interval,
easily interpretable, and statistically parsimonious. The seasonal dependence of \( \alpha_\nu \)
on $\nu$ allows us to quantify whether or not all months are warming/cooling at the same rate. In fact, some global warmists believe that winter temperatures should be warming more rapidly than other temperatures; however, temperature changes should be more easily detected during summer, when temperature variability is minimal, than during winter (see Callendar 1961 and Madden and Ramanathan 1980). We have put this issue in Section 4.4 with periodic analysis of variance.

Future research will examine time series regression models with periodic properties. Periodic time series arise naturally in climatology, meteorology, astronomy, hydrology, economics, etc. It has been demonstrated that seasonal autoregressive moving-average (SARMA) models are inappropriate for modeling such series. Rather, one should look at autoregressive moving-average (ARMA) models with coefficients that vary periodically. Our general goal is to extend the ARMA modeling paradigm to allow for periodicity in the series.

Future work will investigate optimality of OLS estimates in a periodic environment further. Such knowledge would extend Chapter 9 of Fuller (1996) to the periodic setting. Many non-statisticians, and even some statisticians, believe that OLS estimates are very inferior except when series errors are uncorrelated. However, as we showed in Chapter 3, OLS estimates are very good practically even when the responses are serially correlated. All that is needed for asymptotic efficiency is stationary errors with an spectral density bounded away from zero.

We are unsure if the same efficiency holds in a periodic setting. Whereas we do not believe that the OLS estimates in (5.4) are radically inefficient, we do not believe they are asymptotically most efficient either. The tradeoff needs to be understood and quantified — especially if maximum likelihood or weighted least squares techniques are demanded in periodic environments. In Chapter 3, we studied efficiency issues in the time-homogeneous simple linear regression in (5.1) with stationary errors, rehashing the hallowed ground of Grenander (1954) from a modern time-domain
stochastics viewpoint. An extension of this study to the periodic setting will prove insightful.

5.2 Application to Interdisciplinary Research

Beyond this dissertation, I seek to expand my research into three main directions. First, I seek to quantify trends in other scientific areas by investigating time series regression models in periodic settings. A spatial analysis of estimated trends or regression coefficients over varying stations could be conducted if needed. Second, I plan to explore spatio-temporal regression models in environmental and epidemiologic settings in which responses are observed over time in spatially dispersed locations. The area of spatio-temporal modeling is rapidly developing and gives physical insight into many natural phenomena. Third, I plan on developing statistical tools for analyzing extremes of non-stationary and periodically correlated sequences with application to real-world problems. In environmental settings, extreme values have important effects. Statistically, extremes are notoriously more difficult to model than averages. Analysis of non-stationary and periodically correlated extremes is untilled land, which is attractive to me.

Finally, I am also interested in future empirical research that develops statistical methods and models for problems involving longitudinal data, spatial mapping, missing data, and survival analysis. I am particularly looking forward to working in interdisciplinary settings with scientists in climatology, biology, epidemiology, biomedical science, ecology, geophysics, and econometrics.

5.3 References

