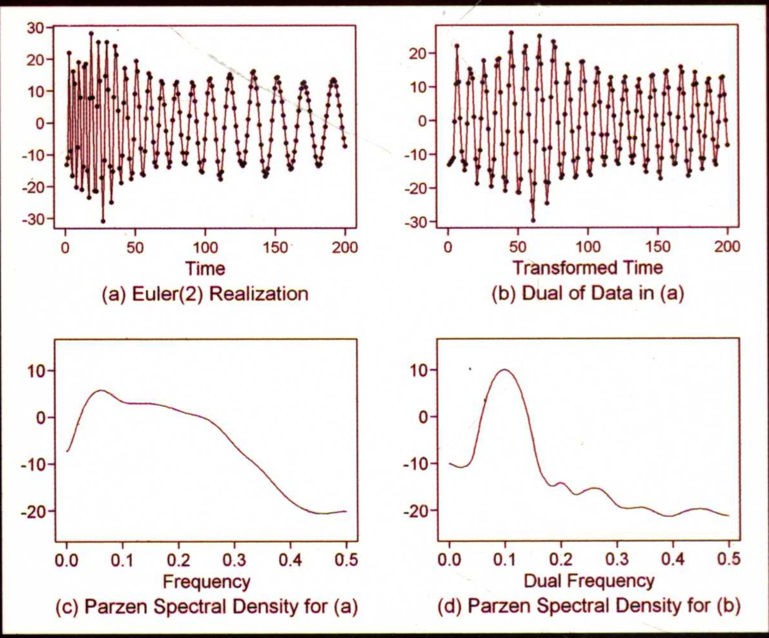
STATISTICS:
TEXTBOOKS and MONOGRAPHS

Applied Time Series Analysis



Wayne A. Woodward, Henry L. Gray, and Alan C. Elliott

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Applied Time Series Analysis

Wayne A. Woodward

Southern Methodist University

Dallas, Texas, USA

Henry L. Gray

Southern Methodist University

Dallas, Texas, USA

Alan C. Elliott

University of Texas Southwestern Medical Center at Dallas

Dallas, Texas, USA

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
Version Date: 20110629

International Standard Book Number: 978-1-4398-1837-4 (Hardback)

Library of Congress Cataloging-in-Publication Data

Woodward, Wayne A.
Applied time series analysis / Wayne A Woodward, Henry L. Gray, and Alan C. Elliott.
p. cm. -- (Statistics: textbooks and monographs)
Includes bibliographical references and index.
ISBN 978-1-4398-1837-4 (alk. paper)
1. Time-series analysis. I. Gray, Henry L. II. Elliott, Alan C. III. Title.

QA280.W66 201
519.5'5--dc23 2011025090

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Applied Time Series Analysis

STATISTICS: Textbooks and Monographs

D. B. Owen

Founding Editor, 1972–1991

STATISTICS: Textbooks and Monographs

Recent Titles

Multivariate Statistical Analysis, Second Edition, Revised and Expanded, *Narayan C. Giri*

Computational Methods in Statistics and Econometrics, *Hisashi Tanizaki*

Applied Sequential Methodologies: Real-World Examples with Data Analysis, *edited by Nitis Mukhopadhyay, Sujay Datta, and Saibal Chattopadhyay*

Handbook of Beta Distribution and Its Applications, *edited by Arjun K. Gupta and Saralees Nadarajah*

Item Response Theory: Parameter Estimation Techniques, Second Edition, *edited by Frank B. Baker and Seock-Ho Kim*

Statistical Methods in Computer Security, *edited by William W. S. Chen*

Elementary Statistical Quality Control, Second Edition, *John T. Burr*

Data Analysis of Asymmetric Structures, *Takayuki Saito and Hiroshi Yadohisa*

Mathematical Statistics with Applications, *Asha Seth Kapadia, Wenyaw Chan, and Lemuel Moyé*

Advances on Models, Characterizations and Applications, *N. Balakrishnan, I. G. Bairamov, and O. L. Gebizlioglu*

Survey Sampling: Theory and Methods, Second Edition, *Arijit Chaudhuri and Horst Stenger*

Statistical Design of Experiments with Engineering Applications, *Kamel Rekab and Muzaffar Shaikh*

Quality by Experimental Design, Third Edition, *Thomas B. Barker*

Handbook of Parallel Computing and Statistics, *Erricos John Kontoghiorghes*

Statistical Inference Based on Divergence Measures, *Leandro Pardo*

A Kalman Filter Primer, *Randy Eubank*

Introductory Statistical Inference, *Nitis Mukhopadhyay*

Handbook of Statistical Distributions with Applications, *K. Krishnamoorthy*

A Course on Queueing Models, *Joti Lal Jain, Sri Gopal Mohanty, and Walter Böhm*

Univariate and Multivariate General Linear Models: Theory and Applications with SAS, Second Edition, *Kevin Kim and Neil Timm*

Randomization Tests, Fourth Edition, *Eugene S. Edgington and Patrick Onghena*

Design and Analysis of Experiments: Classical and Regression Approaches with SAS, *Leonard C. Onyiah*

Analytical Methods for Risk Management: A Systems Engineering Perspective, *Paul R. Garvey*

Confidence Intervals in Generalized Regression Models, *Esa Uusipaikka*

Introduction to Spatial Econometrics, *James LeSage and R. Kelley Pace*

Acceptance Sampling in Quality Control, *Edward G. Schilling and Dean V. Neubauer*

Applied Statistical Inference with MINITAB®, *Sally A. Lesik*

Nonparametric Statistical Inference, Fifth Edition, *Jean Dickinson Gibbons and Subhabrata Chakraborti*

Bayesian Model Selection and Statistical Modeling, *Tomohiro Ando*

Handbook of Empirical Economics and Finance, *Aman Ullah and David E. A. Giles*

Randomized Response and Indirect Questioning Techniques in Surveys, *Arijit Chaudhuri*

Applied Time Series Analysis, *Wayne A. Woodward, Henry L. Gray, and Alan C. Elliott*

To Beverly, Becky, and E'Lynne

Preface

Why another time series book? That is a good question since there are so many excellent books available for a first and/or second course in time series analysis. *Applied Time Series Analysis* is a product of more than 25 years of teaching a course in introductory time series at Southern Methodist University (SMU). This course has always been substantially integrated with the use of an in-house time series package designed not only to do data analysis but also to educate users about properties of models. These time series analysis capabilities are now integrated into the software package `GW-WINKS` that can be downloaded from the website given later in the Preface.

Please keep in mind that this book is written to accomplish several distinct purposes. Among these are (1) provide a teachable book with innovative class-proven material that can be used in first and second semester time series courses, (2) cover areas such as long-memory models and data with time-varying frequencies/autocorrelations that are not typically dealt with in other such books, and (3) provide the student with the tools to solve *real* problems. Upon completion of this book, the student should have the understanding necessary to model, forecast, and identify the underlying components of a time series.

We believe that the writing style and the accompanying software will make this book a functional text for experts in the field of time series analysis along with being an invaluable resource for faculty members who have been "assigned" to teach the time series course but have no real experience in this area.

Features of the Book

Chapter 1 provides a treatment of stationarity and an analysis of stationary time series from both the time and frequency domains. We believe that it is important for students to become comfortable with the frequency domain from the get-go. The chapter contains a readable coverage of the frequency domain that does not assume a prior familiarity with it. Frequency domain concepts covered in this chapter include (1) frequency and period, (2) Nyquist frequency, (3) spectrum and spectral density, and (4) periodogram and window-based estimation of the spectral density.

Chapter 2 discusses basic filter material. In addition to the typical discussion of the general linear process and its relationship to stationarity, we provide a section containing practical filtering applications, which includes coverage of the frequency response function of a filter and how it determines the high pass, low pass, band pass, etc. properties of the filter. Unique to this statistical time series text is a brief discussion of the Butterworth filter, a widely used filter in science and engineering but one not usually covered in time series texts.

Chapter 3 gives an extensive coverage of ARMA(p,q) models. Learning the material in this chapter is greatly enhanced by using the `GW-WINKS` software, which provides unique capabilities for generating and plotting realizations from specified models along with their theoretical autocorrelations and spectral densities. Numerous simulation-based examples are given in this chapter to show properties of models. Although basic theorems are given and their proofs either included in the chapter's appendix or referenced, our coverage of these models focuses on applications over theory.

Factor tables: Unique to our treatment of ARMA processes is the use of factor tables that are based on a factoring of the pth and qth order characteristic polynomials of the autoregressive and moving average portions of an ARMA (p,q) model into their first and second order factors (see Woodward et al., 2009). We have used factor tables extensively in our introductory time series course at SMU for many years, and our experience is that these are indispensable tools for understanding ARMA models. The factor table and examples in the book are used to aid in understanding

- The frequency behavior of models
- How the proximity of a root (of the characteristic equation) affects the contribution of that root to the model characteristics
- The effect of near cancellation of factors in ARMA(p,q) models

Chapter 4 goes beyond the ARMA(p,q) models and discusses other important stationary models. We discuss the harmonic component model and its relationship to ARMA(p,q) models. We also give some basic material on ARCH/GARCH models that is designed to give readers an intuitive grasp of these models, which have proven to be useful in the analysis of economic time series data.

Chapter 5 contains a treatment of nonstationary models. These include signal-plus-noise models along with ARMA-type models with roots on the unit circle. Our treatment of nonstationary ARMA-type models has the following unique features:

- Use of results due to Findley (1978) and Quinn (1980) that illustrate the behavior of autocorrelations associated with roots on or close to the unit circle.
- An extension of the ARIMA(p,d,q) models to ARUMA(p,d,q) models that include the ARIMA models but also allow for complex conjugate roots on the unit circle. That is, these models allow for cyclic nonstationary behavior.
- A discussion of the seasonal model from the perspective of the associated factor table.

Chapter 6 covers forecasting, focusing on the Box–Jenkins approach. We discuss forecasting using both stationary and nonstationary (ARUMA) models. Several examples are given.

Chapter 7 provides basic material on parameter estimation. While the main focus is on maximum likelihood estimation, we discuss computationally efficient methods and discuss their advantages and disadvantages.

Chapter 8 discusses the topic of ARMA model identification. Our focus is on the use of AIC and its variations as the primary tool for ARMA model identification. Other model identification techniques such as the Box–Jenkins and array methods are discussed in an appendix. One unique feature is our use of "overfitting" to assist in model identification when some roots are on (i.e., nonstationary models) or close to the unit circle. We use factor tables to identify nonstationary ARIMA/ARUMA and seasonal time series models.

Chapter 9 is a relatively unique wrap-up of previous material that discusses issues related to selecting a final model for a set of data. Through examples (real and simulated), we give practical advice concerning checking the whiteness of residuals, deciding whether to use a stationary or nonstationary model, and whether to use signal-plus-noise or ARMA/ARUMA models. We also discuss checking the characteristics of realizations simulated from a fitted model to ascertain whether the characteristics of generated realizations are consistent with those of the actual data. Students have told us they like this chapter because it helps them put together everything they have learned to this point.

Chapter 10 covers multivariate time series and vector autoregressive models. We also provide an introduction to state-space models that will provide the reader with an overview of this important topic.

Chapter 11 discusses long-memory models that have received a considerable amount of recent attention in the literature. This chapter discusses not only the fractional (and FARMA) models typically covered in texts such as these, but also the Gegenbauer and GARMA models that allow for long-memory behavior associated with any frequency between 0 and 0.5 inclusive.

Chapter 12 covers wavelets and short-term Fourier transforms for analyzing data having behavior patterns that change with time and cannot successfully be modeled using ARMA/ARUMA models. We believe that this material will form the basis for a useful introduction to these ideas for the instructor who wants to spend a few weeks on wavelets.

Chapter 13 introduces innovative techniques to deal with data with time-varying frequencies (TVF) such as Doppler, bat echolocation, seismic, and chirp signals. The fundamental idea is that for many data sets with TVF behavior there is a transformation of the time axis on which the data are stationary. This is a topic that is based on recent developments in the literature and is totally unique to this book. GWS software, which runs on the `S-Plus` system, is available on the website.

Software Available with This Book

One of the unique features of this book is that we provide the (Windows®-based) software, GW-WINKS, which is easy to learn and use and is designed to help students understand time series models as well as analyze data. As mentioned, *we highly recommend using the GW-WINKS software with the book.* Sections entitled "Using GW-WINKS" are included throughout the book to discuss the use of GW-WINKS for performing analyses. GW-WINKS is a special purpose version of the WINKS SDA™ software package from TexaSoft (http://www.texasoft.com). A license for GW-WINKS is included with the book. While other programs are available for performing many of the procedures discussed in this book, GW-WINKS provides a "learning environment" that enhances the understanding of the material.

The website associated with this book is http://www.texasoft.com/atsa. This website will be referred to as the ATSA (*Applied Time Series Analysis*) website. This site can also be accessed by visiting http://faculty.smu.edu/waynew and selecting the link to Applied Time Series Analysis.

In addition to providing information on installing and running GW-WINKS, the website also contains R programs for performing computations related to some of the topics, primarily those covered in Chapters 11 and 12. Some of these programs can be run directly through GW-WINKS using data in the GW-WINKS data sheet. We also provide stand-alone R programs on the website. The R package can be downloaded free from the Comprehensive R Archive Network (CRAN) website at http://cran.r-project.org.

The GWS package for using the techniques discussed in Chapter 13 to analyze TVF data is also available on the ATSA website. GWS requires the S-Plus® package. Instructions for the use of the GWS software are provided on the ATSA website. Additionally, step-by-step directions for use of GWS are included in the Chapter 13 exercises. In the text, we also discuss the use of the Wavelets module in S-Plus®, along with MATLAB® (Chapter 2) and SAS® (Chapter 10) for performing certain analyses. Related code is provided on the ATSA website.

The ATSA website contains the following:

1. A *Getting Started Guide* for GW-WINKS that includes
   a. Instructions for downloading, installing, and activating the GW-WINKS program
   b. Step-by-step tutorials showing how to manage data in GW-WINKS and how to run a selection of GW-WINKS Time Series procedures
2. R programs for some procedures discussed in the book
3. Instructions on how to write, modify, and run WINKS-R programs (programs that require R software and are run from the GW-WINKS menus)

4. Instructions for downloading and running `GWS` software
5. Datasets used in examples and problems in the book
6. Updated information on software and a list of corrections for typos and other errors in the book

For MATLAB® and Simulink® product information, please contact:

The MathWorks, Inc.
3 Apple Hill Drive
Natick, MA, 01760-2098 USA
Tel: 508-647-7000
Fax: 508-647-7001
E-mail: info@mathworks.com
Web: www.mathworks.com

Acknowledgments

We would be remiss to not mention some of the people who have helped with this project, which actually started in the 1980s but got sidetracked for a variety of reasons. They include Nien Fan Zhang, who was involved in the early stages of the book, Tom Fomby, who has worked hard to keep us informed about econometric techniques, and countless graduate students who have taken the time series course and helped us to refine our pedagogical methods. The first author (Woodward) taught the graduate level time series course in the Spring 2010 semester using a preliminary set of notes for the first nine chapters. The following students did a great job of beta testing the software and copyediting the text, and they deserve to be thanked individually: Aymen Al-Rawashdeh, Fang Duan, Xiaowen Hu, Yalan Hu, Dong Lin, Runqi Lin, Long Luo, Katherine Sisk, Holly Stovall, Miao Zang, Peiqi Xhai, and Josh Zorsky. Special thanks go to Joel O'Hair who helped with the graphics and the wavelets material, Wenkai Bao who was a student in the time series class and also helped develop R code related to the long-memory models, and James Haney who played an integral role related to the material on the use of state-space methodology for computing additive components of an autoregressive model. We would also like to thank Giovanni Petris who has given the manuscript a careful reading and provided several very useful suggestions. Sheila Crain deserves a great deal of credit for her excellent job of inputting the manuscript and helping in all phases of the manuscript preparation. To all our colleagues with whom we have discussed these topics over the years, we thank you for your positive influence on this work.

Contents