

University of Nevada, Reno

**Seasonal Time Series Models and Their Application
to the Modeling of River Flows**

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

by

Yonas Gebeyehu Tesfaye

Dr. Mark M. Meerschaert/Dissertation Advisor

May, 2005

UNIVERSITY
OF NEVADA
RENO

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

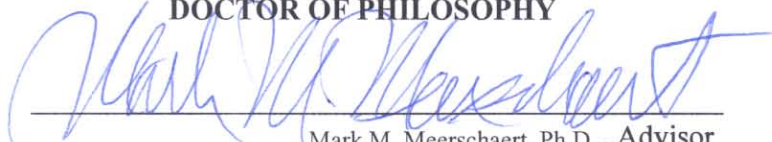
YONAS GEBEYEHU TESFAYE

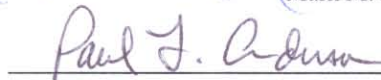
entitled

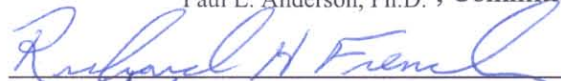
**Seasonal Time Series Models and Their Application to the Modeling of
River flows**

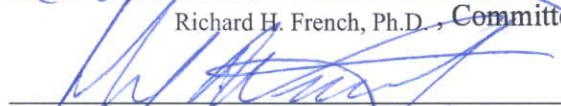
be accepted in partial fulfillment of the
requirements for the degree of

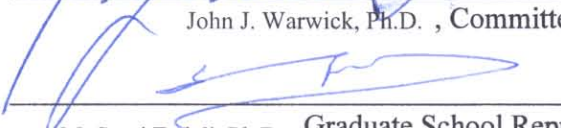
DOCTOR OF PHILOSOPHY

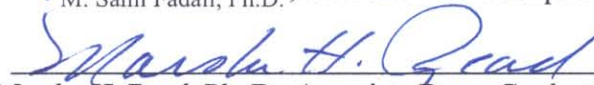

Mark M. Meerschaert, Ph.D. , Advisor


Paul L. Anderson, Ph.D. , Committee Member


Richard H. French, Ph.D. , Committee Member


John J. Warwick, Ph.D. , Committee Member


M. Sami Fadali, Ph.D. , Graduate School Representative


Marsha H. Read, Ph. D., Associate Dean, Graduate School

May, 2005

Abstract

This dissertation is concerned with the problem of modeling a periodic hydrological time series in general and river flows in particular. River flow series usually exhibit periodical stationarity; that is, their mean and covariance functions are periodic with respect to time. A class of models useful in such situations consists of periodic autoregressive moving average (PARMA) models, which are extensions of commonly used ARMA models that allow parameters to depend on season. The innovations algorithm can be used to obtain estimates of certain model parameters. Asymptotic distributions for the innovations estimates and model parameters provide us with a general technique for identifying PARMA models with the minimum number of parameters. A detailed simulation study is conducted to test the effectiveness of the innovation estimation procedures and asymptotics as a method for PARMA model identification. We use monthly river flow data for the Fraser River in British Columbia and the Salt River in Arizona so as to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows. In both cases, we are able to generate realistic synthetic river flows by developing a mixture probability model for the residuals.

Fitting the PARMA model to weekly or daily data, however, requires estimation of too many parameters which violates the principle of parsimony. In an effort to obtain a parsimonious model representing these periodically stationary series, we develop the asymptotic distribution of the discrete Fourier transform of the innovation estimates and PARMA model parameters, and then determine those statistically significant Fourier coefficients. We demonstrate the effectiveness of these techniques using simulated data from different PARMA models. An application of the techniques is demonstrated through the analysis of

a weekly river flow series for the Fraser River.

Hydrologic droughts are characterized by their duration and severity. We use PARMA models to model river flows and generate a large sequence of synthetic flows, and then drought properties are derived from the simulated samples based on the theory of runs. The applicability of these methods is demonstrated by using weekly river flow data for the Fraser River.

Dedication

To the memory of my father

To my mother, brothers and sisters

Acknowledgements

Completing a research project and writing a dissertation is obviously not possible without support of numerous people. First of all, I would like to express my gratitude to Prof. Mark Meerschaert for giving me the chance to work on a river modeling project and for his valuable consultation and support throughout the research work. I am also deeply grateful to Prof. Meerschaert for his never ending patience, enthusiasm and encouragement in guiding the preparation and production of this manuscript. Especially the extensive comments and the many discussions and the interactions with him had a direct impact on the final form and quality of the manuscript. He has got all the best qualities that any graduate student could expect from his/her research supervisor. I am very lucky to have him as my advisor and learned a lot from him. The support from the National Science Foundation grants GMS-0139927 and GMS-0417869 are gratefully acknowledged.

The members of my PhD committee Prof. Paul Anderson, Prof. Sami Fadali, Prof. Richard French and Prof. John Warwick have generously given their time and expertise to make my work better. I thank them for their contribution, support, careful reading the entire manuscript and providing valuable comments. In particular, I am grateful to Prof. Anderson for coauthoring and reviewing the papers produced from the dissertation.

I am greatly indebted to Dr. Laurel Saito who facilitated things for me to get here at the University of Nevada-Reno (UNR), supported me during my first year study, helped me in finding the project of my interest. She is open, kind, cooperative, and very helpful. Thank you for everything you have done for me. I wish also to thank Prof. Scott Tyler for his advice, encouragement, his willingness to help students and for all administrative support. I thank Sam Miller and Pamela Love for their help during the Ph.D. process. I am a recipient of the 2003 Douglas Paul Rennie Memorial Graduate Scholarship and it is

gratefully acknowledged. I also thank the Graduate Student Association for partial financial support during my travel to present my work to conferences. I must acknowledge as well the professors, students, librarians and others at UNR who assisted and advised me during my study period. In general, UNR is a wonderful atmosphere for intellectual growth and accomplishment.

I am thankful to Mrs. Lenette Grant and other staff at University of Otago in New Zealand for making a short period of research work in Duendin an enjoyable one. Many other people have contributed to my success directly or indirectly by their advice, moral and material support. In particular, I thank Dr. Shifferaw Taye, Dr-ing Girma Boled and others colleagues in the Addis Ababa University, Prof. Ånund Killingtveit (from Norway) and Prof. Dr.-Ing Manfred Ostrowski (from Germany) for their genuine support of my career development. I would also like to thank my mother, brothers and sisters for their unconditional love and continuous support of my academic career.

Finally, I owe whatever I have accomplished to God without Whom I could accomplish nothing.

Contents

1	Introduction	1
1.1	Characteristics of Hydrological Time Series	2
1.2	Modeling and Simulation of Hydrological Series	7
1.3	Literature Review	8
1.4	Objectives of the Research	10
2	PARMA Models and Their Parameter Estimation	13
2.1	General	13
2.2	Mathematical Formulation of PARMA Model	15
2.3	Parameter Estimation for PARMA Model	17
3	PARMA Model Identification: Finite Fourth Moment Case	22
3.1	General	22
3.2	Asymptotic Distributions for Moving Average Processes	23
3.3	Simulation Study	40
3.4	Application to Modeling of Natural River Flows	46
4	Modeling River Flows with Heavy Tails	58
4.1	General	58
4.2	Definition of Heavy Tails	59
4.3	Heavy Tailed Time Series Model	60
4.4	Checking Heavy Tails in Time Series Data	65
4.5	Application to Modeling of Natural River Flows	69
5	Estimation of Periodic Parameters by Fourier Series	80

5.1	Overview	80
5.2	Selection of Significant Harmonics and Fourier Coefficients	81
5.3	Estimation of Fourier Coefficients of Model Parameters	88
5.4	Simulation Study	95
5.5	Application to Modeling of River Flows	98
6	Application of PARMA Models: Hydrological Drought Analysis	111
6.1	General	111
6.2	Definitions	112
6.3	Methodology	113
6.4	Example: Hydrologic Drought Analysis for the Fraser River	116
6.4.1	Modeling of $\text{PARMA}_{52}(1, 1)$ Model Residuals	116
6.4.2	Drought Statistics Estimation	120
7	Summary and Conclusions	125
	References	130
	Appendices	144
A	Basic Concepts	144
A-1	Definitions, Theorems, etc.	144
A-2	Statistical Properties of Time Series	149
A-3	Spectral Representation of a Stationary Process	152
A-4	Simulation of Continuous Random Variates	154
A-5	Akaike Information Criterion	156
A-6	The Bonferroni Inequality and Simultaneous Confidence Intervals	157

B	Vector Difference Equation for Discrete Fourier Coefficients	159
C	Figures and Tables	162

List of Tables

1.1	Hill estimators for monthly river flow data.	6
3.1	Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 500$ years of simulated PARMA ₄ (0, 1) data.	41
3.2	Model parameters and estimates for simulated PARMA ₄ (0, 1) data.	42
3.3	Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 500$ years of simulated PARMA ₄ (1, 1) data.	45
3.4	Model parameters and estimates (with 95% confidence interval) for simulated PARMA ₄ (1, 1) data.	46
3.5	Sample mean, standard deviation and autocorrelation at lag 1 and 2 of average monthly flow series for the Fraser River at Hope, BC.	48
3.6	Moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, 2, \dots, 6$, and p -values, after $k = 20$ iterations of the innovations algorithm applied to average monthly flow series for the Fraser River at Hope, BC.	49
3.7	Parameter estimates (with 95% confidence interval) for PARMA model (3.63) of average monthly flow series for the Fraser River at Hope, BC.	51
4.1	Comparison of tail probabilities for standard normal, Cauchy and Lévy distributions.	60
4.2	Sample mean, standard deviation and autocorrelation at lag 1 and 2 of average monthly flow series for the Salt River near Roosevelt, AZ.	71

4.3	Moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, 2, \dots, 10$, after $k = 15$ iterations of the innovations algorithm applied to average monthly flow series for the Salt River near Roosevelt, AZ.	72
4.4	Parameter estimates for PARMA ₁₂ (1, 1) model of average monthly flow series for the Salt River near Roosevelt, AZ.	73
4.5	Parameter estimates for PARMA ₁₂ (1, 1) model residuals.	75
5.1	Model parameters and estimates for simulated PARMA ₁₂ (0, 1) data.	96
5.2	Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 500$ years of simulated PARMA ₁₂ (0, 1) data. Note that the value in (.) is the test statistic (5.21). . .	103
5.3	Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 500$ years of simulated PARMA ₁₂ (1, 1) data. Note that the value in (.) is the test statistic (5.21). . .	104
5.4	Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated PARMA ₁₂ (1, 1) data. Note that the value in (.) is the test statistic (5.34) and (5.38).	105
5.5	Significant discrete Fourier transform coefficients for simulated PARMA ₁₂ (1, 1) model parameters.	105
5.6	Discrete Fourier transform of moving average parameter estimate, $\hat{\theta}_t$ (with standard error, SE = 0.009 for $m = 26$ and SE = 0.013 for $m \neq 26$) for the simulated PARMA ₅₂ (1, 1) data. Note that the value in (.) is the test statistic (5.38).	106

5.7	Discrete Fourier transform of autoregressive parameter estimate, $\hat{\phi}_t$ (with standard error, SE = 0.011 for $m = 26$ and SE = 0.015 for $m \neq 26$) for the simulated PARMA ₅₂ (1,1) data. Note that the value in (.) is the test statistic (5.34).	107
5.8	Significant discrete Fourier transform coefficients for simulated PARMA ₅₂ (1,1) model parameters.	108
5.9	Discrete Fourier transform of model parameters estimates and standard errors (SE) of average monthly flow series for the Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.34) and (5.38).	108
5.10	Discrete Fourier transform of moving average parameter estimate, $\hat{\theta}_t$ (with standard error, SE = 0.017 for $m = 26$ and SE = 0.024 for $m \neq 26$) for PARMA ₅₂ (1,1) model of average weekly flow series for Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.38).	109
5.11	Discrete Fourier transform of autoregressive parameter estimate, $\hat{\phi}_t$ (with standard error, SE = 0.023 for $m = 26$ and SE = 0.033 for $m \neq 26$) for PARMA ₅₂ (1,1) model of average weekly flow series for the Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.34).	110
6.1	Other parameter estimates for PARMA ₅₂ (1,1) model of average weekly flow series for the Fraser River at Hope, BC.	123
6.2	Drought statistics estimation of the historical and simulated data for the Fraser River at Hope, BC. Note that the drought duration (L), severity (D) and return period (T) are in weeks, in cms and in years, respectively.	124

C-1	Moving average parameter estimates and p -values after $k = 20$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (0, 1) data.	163
C-2	Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (0, 1) data.	164
C-3	Moving average parameter estimates and p -values after $k = 10$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (0, 1) data.	165
C-4	Moving average parameter estimates and p -values after $k = 20$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (1, 1) data.	166
C-5	Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (1, 1) data.	167
C-6	Moving average parameter estimates and p -values after $k = 10$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA ₄ (1, 1) data.	168
C-7	Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 50$ years of simulated PARMA ₁₂ (0, 1) data. Note that the value in (\cdot) is the test statistic (5.21).	169

C-8	Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 100$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in (.) is the test statistic (5.21).	170
C-9	Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 300$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in (.) is the test statistic (5.21).	171
C-10	Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated $\text{PARMA}_{12}(1, 1)$ data ($N_y = 100$ years). Note that the value in (.) is the test statistic (5.34) and (5.38).	172
C-11	Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated $\text{PARMA}_{12}(1, 1)$ data ($N_y = 300$ years). Note that the value in (.) is the test statistic (5.34) and (5.38).	172
C-12	Some statistical distributions	173

List of Figures

1.1	Daily (top left), monthly (top right), annual (bottom left) river flow series for Truckee River at Farad, CA. Autocorrelation function for monthly data for the same river (bottom right).	5
1.2	Time series plot (left) and histogram (right) of monthly river flows (cfs) for the Salt River near Roosevelt, Arizona, from October 1914 to September 1986.	6
3.1	Time series plot of the residuals	43
3.2	Probability plot of the residuals	43
3.3	The sample ACF and PACF for the residuals, showing the 95% confidence bounds $\pm 1.96/\sqrt{N}$, indicating iid sequence.	44
3.4	Average monthly flows (cms) for the Fraser River at Hope, BC, indicate a seasonal pattern.	47
3.5	Statistical summary of the residuals.	50
3.6	ACF and PACF for model residuals, showing the bounds $\pm 1.96/\sqrt{N}$, indicate no serial dependence.	50
3.7	Lognormal probability plot (a) and histogram (b) for model residuals, Fraser river at Hope, BC.	52
3.8	Log-log plot of upper(left) and lower(right) residual tails and fitted Pareto and truncated Pareto distributions, Fraser river at Hope, BC.	54
3.9	Probability plot of simulated innovations using the mixed three parameter lognormal and truncated Pareto distributions. Compare Figure 3.7 (a).	56
3.10	Plot of (a) observed and (b) simulated monthly river flows for the Fraser River at Hope, BC, indicating similarity.	56

3.11	Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed monthly river flow data for the Fraser River at Hope, BC. . . .	57
4.1	Plots of standardized normal, Cauchy and Lévy densities.	61
4.2	(a) Hill's estimator, $\hat{\alpha}$ versus r , and (b) log-log plot for the upper tail of the monthly flows for the Salt River near Roosevelt, AZ.	67
4.3	Average monthly flows (cfs) for the Salt River near Roosevelt, AZ, indicate a seasonal pattern and heavy tails.	70
4.4	(a) ACF and (b) PACF for PARMA ₁₂ (1, 1) model residuals, showing the 95% confidence bounds $\pm 1.96/\sqrt{N}$	74
4.5	PARMA ₁₂ (1, 1) Model residuals of the Salt River near Roosevelt, AZ: Statistical summary (left) and three parameter lognormal probability plot (right).	74
4.6	Hill's estimator, $\hat{\alpha}$ versus r , of (a) upper tail and (b) lower tail for the PARMA ₁₂ (1, 1) model residuals.	75
4.7	Density, variance stabilized p - p plot and q - q plot for the PARMA ₁₂ (1, 1) model residuals.	76
4.8	Log-log plot of upper (left) and lower (right) residual tails and fitted Pareto and truncated Pareto distributions, Salt River near Roosevelt, AZ.	77
4.9	Probability plot of simulated innovations using the mixed stable and truncated Pareto distributions. Compare Figure 4.5 (right).	77
4.10	Plot of (a) observed and (b) simulated monthly river flows for the Salt River near Roosevelt, AZ, indicating statistical similarity.	78
4.11	Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed monthly river flow data for the Salt River near Roosevelt, AZ.	79

5.1	Plot of PARMA ₁₂ (1, 1) model parameters (with their Fourier fit) of average monthly flow series for the Fraser River at Hope, BC.	99
5.2	Statistical summary (a) and lognormal probability plot (b) for model residuals, Fourier model parameters, monthly Fraser river data. Compare Figure 3.7 (a).	99
5.3	Sample Statistics (a) mean (b) standard deviation and autocorrelations at lag 1 (c) and lag 2 (d) of weekly river flow data for the Fraser River at Hope, BC.	100
5.4	ACF and PACF for model residuals, showing the bounds $\pm 1.96/\sqrt{N}$, indicate no serial dependence.	101
5.5	Statistical summary (left) and lognormal probability plot (right) for model residuals, weekly Fraser river data. Compare Figure 3.7	101
5.6	Plot of PARMA ₅₂ (1, 1) model parameters (with their Fourier fit) of average weekly flow series for the Fraser River at Hope, BC.	102
6.1	Definition of hydrological drought properties for a periodic river flow series X_t (solid line) and a periodic water demand X_{0t} (broken line); and definition of variables for estimation of return period.	114
6.2	(a) ACF and (b) PACF for PARMA ₅₂ (1, 1) model residuals of the Fraser River at Hope, BC.	117
6.3	Statistical summary (left) and lognormal probability plot (right) for PARMA ₅₂ (1, 1) model residuals of the weekly Fraser River data, after applying significant Fourier coefficients. Compare Figure 5.5 (right).	118
6.4	Log-log plot of upper (left) and lower (right) residual tails and fitted Pareto and truncated Pareto distributions, Fraser river at Hope, BC.	119
6.5	Probability plot of simulated innovations using the mixed three parameter lognormal and truncated Pareto distributions. Compare Figure 6.3 (right). . .	120

6.6	Plot of (a) observed and (b) simulated weekly river flows for the Fraser River at Hope, BC, indicating similarity.	121
6.7	Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed weekly river flow data for the Fraser River at Hope, BC.	122
C-1	Normal probability plots of the residuals (for $k=15$) from $\text{PARMA}_4(0, 1)$, support the normality assumptions of residuals.	162
C-2	Normal probability plots of the residuals (for $k=15$) from $\text{PARMA}_4(1, 1)$, support the normality assumptions of residuals.	162

1 Introduction

Time series analysis and modelling is an important tool in hydrology and water resources. It is used for building mathematical models to generate synthetic hydrologic records, to determine the likelihood of extreme events, to forecast hydrologic events, to detect trends and shifts in hydrologic records, and to fill in missing data and extend records. Synthetic river flow series are useful for determining the dimensions of hydraulic works, for flood and drought studies, for optimal operation of reservoir systems, for determining the risk of failure of dependable capacities of hydroelectric systems, for planning capacity expansion of water supply systems, and for many other purposes (Salas *et. al.*, 1985; Salas, 1993). For example, hydrologic drought properties (severity and duration as defined in Chapter 6) of various return periods are needed to assess the degree to which a water supply system will be able to cope with future droughts and, accordingly, to plan alternative water supply strategies. They can be determined from the historical record alone by using non-parametric methods but, because the number of drought events that can be drawn from the historical sample is generally small, the “historical” drought properties have a large degree of uncertainty. Other alternatives for finding drought properties include using stochastic models (such as PARMA models) that can represent the underlying river flows, simulating long records of such river flows, and then deriving droughts properties from the simulated samples based on the theory of runs (see Chapter 6 for a detailed discussion).

The main purposes of this dissertation are to fit a PARMA model to represent a given river flow data, estimate parameters, check for goodness of fit to the data, model the residuals, and to use the fitted model for generating synthetic river flows and apply them to extreme analysis such as droughts. The innovations algorithm can be used to obtain estimates of the PARMA model parameters. Asymptotic distributions for the innovations estimates

and model parameters (which are developed here) provide us with a general technique for identifying PARMA models with the minimum number of parameters. A detailed simulation study is conducted to test the effectiveness of the innovation estimation procedures and PARMA model identification. We use monthly river flow data for the Fraser River in British Columbia in order to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows. We apply the same methods to monthly flow data for the Salt River in Arizona, which has heavy tails. In both cases, we are able to generate realistic synthetic river flows by developing a mixture probability model for the residuals. In an effort to obtain a parsimonious model representing periodically stationary series, we develop the asymptotic distribution of the discrete Fourier transform of the innovation estimates and other model parameters, and then determine those statistically significant Fourier coefficients. We demonstrate the effectiveness of these techniques using simulated data from different PARMA models. An application of the techniques is demonstrated through the analysis of a weekly river flow series for the Fraser River. Finally, we demonstrate the application of synthetically generated weekly river flows using PARMA model for hydrologic drought analysis of the Fraser River.

1.1 Characteristics of Hydrological Time Series

The structure of hydrological time series consists of mainly one or more of these four basic structural properties and components (Salas, 1993):

- Over year trends and other deterministic changes (such as shifts in the parameters). In general, natural and human induced factors may produce gradual and instantaneous trends and shift in hydrological time series. See, for example, Salas (1980,1993) for a detailed discussion of trends and shifts in hydrological data and their removal.

- Intermittency in the processes, mainly consisting of the hydrology of intermittent sequences of zero and non-zero values.
- Seasonal or periodic changes of days, weeks or months within the annual cycle. Periodicity means that the statistical characteristic changes periodically within the year. For example, in hydrologic data concerning river flows, we expect high runoff periods in the spring and low flow periods in the summer. Thus the river flow correlations between spring months may be different from the correlations between summer months.
- Stochasticity or random variations.

All the rivers used as case studies in this research are perennial rivers. None of them revealed any significant trend or shift. Therefore, we consider only periodicity and stochasticity in the modeling of the river flow series.

Although basic hydrologic processes evolve on a continuous time scale, most analysis and modeling of such processes are made by defining them on a discrete time scale. In most cases, a discrete time processes is derived by aggregating the continuous time process within a given time interval while in other cases it is derived by sampling the continuous process at a discrete point in time. For example, a weekly river flow series can be derived by aggregating the continuous flow hydrograph on a weekly basis while a daily river flow series can be the result of simply sampling the flows of a river once daily or by integrating the continuous flow hydrograph on a daily basis. In any case, most time series of hydrological processes are defined at hourly, daily, weekly, monthly, bimonthly, quarterly, and annual time intervals. It must be emphasized at this point that the term period (also used interchangeably with season) applies for any time resolution smaller than one year.

The plot of a hydrological series gives a good indication of its main characteristics. For example, the daily, monthly, annual flows of the Truckee River at Farad (gauge 10346000), CA are given in Figure 1.1. The daily series (for this case, beginning October 1, 1912) shows a typical behavior in which during some days of the year the flows are low and in other days the flows are high. During the days of low flows the variability is small, on the other hand, during high flows the variability is large. This characteristic behavior of daily flows indicates that the daily means vary throughout the year. This is further substantiated if one observes the plot of monthly flows (ten years of data beginning October, 1912). In general, most hydrological time series exhibit periodic variations in mean, variance, covariance and skewness. The periodicity in mean, as can be seen from Figure 1.1, may be easily observed in the plot of the underlying hydrologic time series. However, the periodicity in higher order moments is not so easy to observe and usually requires further mathematical analysis (see, for example, Chapters 4 and 5 for a detailed discussion).

Autocorrelation analysis can also be used to identify cycles or periodic components of hydrological time series. Figure 1.1 (bottom right) shows the autocorrelation of monthly river flows for the Truckee River at Farad, CA. The non-decaying nature of the correlation function can be observed, indicating a correlation between flows at 12 months' lag. Spectral analysis has also been used to detect cycles in hydrological data (Rosner and Yevjevich, 1966; Quimpo, 1967; Salas, 1993). In addition, periodicity in autocorrelation can be detected using asymptotic theory (see, for example, Anderson and Vecchia, 1993; Anderson and Meerschaert, 1997,1998).

Many river flow time series also exhibit occasional sharp spikes (also called heavy tail characteristics; that is, some observations are several times greater than the average), and as a result they are skewed. For example, Figure 1.2 shows a time series plot and histogram

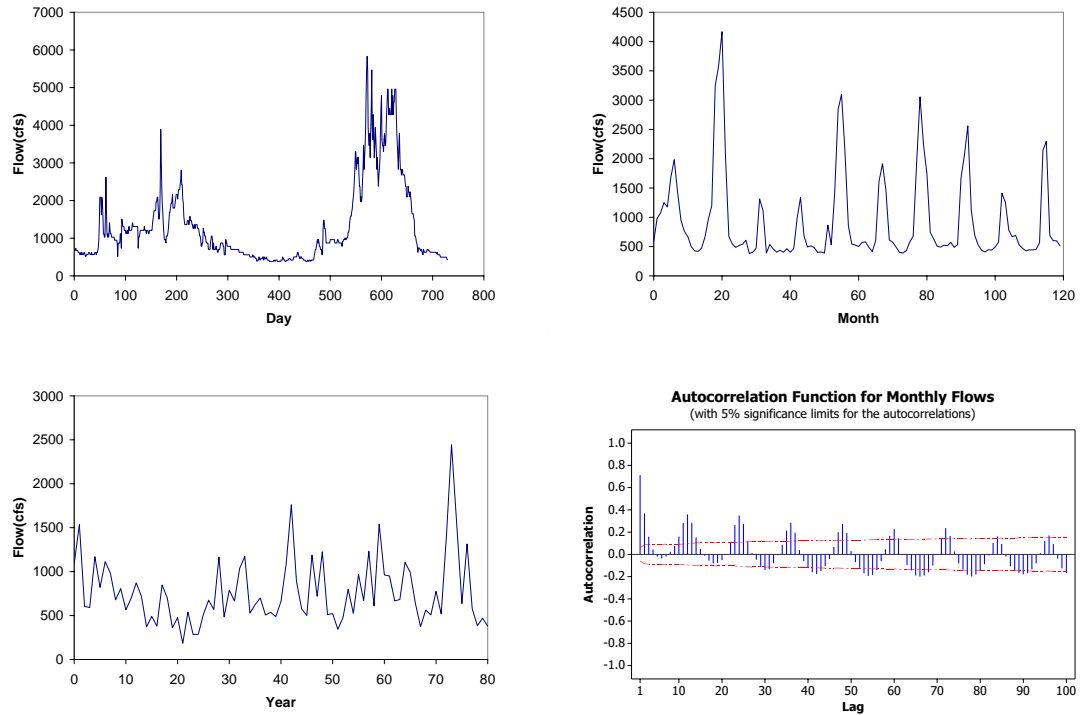


Figure 1.1: Daily (top left), monthly (top right), annual (bottom left) river flow series for Truckee River at Farad, CA. Autocorrelation function for monthly data for the same river (bottom right).

for the 72 years (October 1914 to September 1986) of average monthly flows for the Salt River near Roosevelt, AZ. There are occasional sharp spikes in the time series plot of the data (for example, 14 observations are greater than $6 \times \text{Average}$, where $\text{Average} = 905.50$ cfs), which indicates the heavy tail characteristics of the data. Since the data set exhibits heavy tails and skewness, it can be well described by a heavy tail distribution (for example, a stable distribution¹). The tail behavior of a non-normal stable distribution is asymptotically equivalent to a Pareto distribution (see Chapter 4). A Pareto random variable X , is defined by $P(X > x) = Cx^{-\alpha}$ for some $C > 0$, $\alpha > 0$. The parameter α is called the "tail index" and can be used as a measure of tail-heaviness. For data analysis applications, the actual

¹See the definition in Chapter 4

distribution is often unknown, so robust estimators (for instance, Hill's estimator) are needed to fit the tail parameter α of an unknown heavy tail distribution (see Chapter 4 for a detailed discussion of tail index estimation). Table 1.1 summarizes the Hill's estimators for the average monthly flow data of Fraser, Salt, Truckee rivers. It must be noted that r is the number of order statistics.

Table 1.1: Hill estimators for monthly river flow data.

$r + 1$	$\hat{\alpha}$			\hat{C}		
	Truckee	Salt	Fraser	Truckee	Salt	Fraser
5	2.548	4.825	10.600	3.163E+06	5.773E+16	6.198E+54
10	3.557	3.182	6.889	4.127E+10	1.371E+10	1.541E+34
15	3.408	2.925	8.355	9.828E+09	1.396E+09	1.753E+42
20	3.784	3.023	9.556	3.135E+11	3.318E+09	6.986E+48
25	4.097	2.698	9.090	3.326E+12	1.924E+08	1.938E+46
30	4.314	2.621	9.031	3.909E+13	9.870E+07	9.196E+45
:	:	:	:	:	:	:

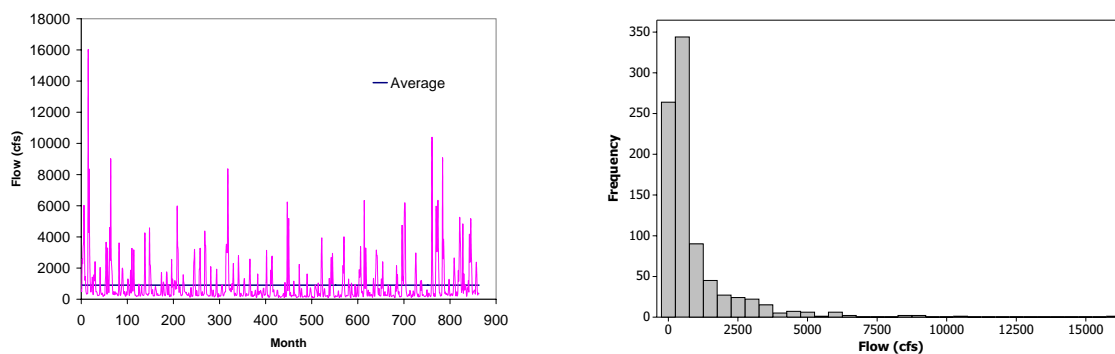


Figure 1.2: Time series plot (left) and histogram (right) of monthly river flows (cfs) for the Salt River near Roosevelt, Arizona, from October 1914 to September 1986.

In time series modeling, the tail parameter α also dictates the modeling approach for data with a Pareto tail. If $\alpha > 4$, then the time series has finite fourth moments, and the classical approach based on normal asymptotics is appropriate. If $2 < \alpha < 4$, then the time series has finite variance but infinite fourth moment and the asymptotics of the sample autocorrelation functions (ACF) are governed by stable laws. Finally, if $\alpha < 2$, the time series has infinite second moments, and the asymptotics are governed by stable laws. Preliminary tail estimations from Table 1.1 (for small $r = 15$) indicates that for the Fraser River $\alpha > 4$, so the time series has light tails and normal asymptotics apply whereas $2 < \alpha < 4$ for the Salt and Truckee rivers, so that these time series have heavy tails with a finite variance but infinite fourth moment, and stable asymptotics apply.

1.2 Modeling and Simulation of Hydrological Series

A number of approaches have been suggested for modeling hydrological time series defined at time intervals less than a year (Salas, 1993; Hipel and McLeod, 1994). The common procedure in modeling such periodic river flow series is first to standardize or filter the series and then fit an appropriate stationary stochastic model to the reduced series (Salas, *et al.*, 1980; Thompstone *et al.*, 1985; Vecchia, 1985; Salas, 1993; Chen, 2002). However, standardizing or filtering most river flow series may not yield stationary residuals due to periodic autocorrelations. In these cases, the resulting model is misspecified (Tiao and Grupe, 1980). Periodic models can, therefore, be employed to remove the periodic correlation structure. An important class of periodic models useful in such situations consists of periodic autoregressive moving average (PARMA) models, which are extensions of commonly used ARMA models that allow periodic parameters. The PARMA modeling procedure involves iterative steps of model identification, parameter estimation, model diagnosis and fitting the

residuals (noise) with a probability distribution function (pdf). The opposite process to step-by-step modeling is the use of models to generate (simulate) new samples or a long sample of the process. One starts with the random noise and its pdf by generating its sample(s). Then generate the corresponding data samples by using the fitted PARMA model.

1.3 Literature Review

Early studies by Hazen (1914) and Sudler (1927) showed the feasibility of using statistics and probability theory in analyzing river flow sequences. Hurst (1951) in investigating the Nile River for the Aswan Dam project, reported studies of long records of river flows and other geophysical series, which years later tremendously impacted the theoretical and practical aspect of time series analysis of hydrologic and geophysical phenomena. Barnes (1954) extended the early empirical studies of Hazen and Sudler and introduced the idea of synthetic generation of stream flow by using a table of random numbers. However, it was not until the beginning of the 1960's that the formal development of stochastic modeling started with introduction and application of autoregressive models for annual and seasonal stream flows (Thomas and Fiering, 1962; Yevjevich, 1963; Roesner and Yevjevich; 1966).

Since then, extensive research efforts have been made towards improving the earlier concepts and models, providing physical justification of some models, introducing alternative models, developing and applying improved estimation procedures as well as fitness tests for such models, and studying their impact in water resources systems planning and management. The concept of periodically correlated processes was introduced by Gladyshev (1961). He gave a formal definition of periodic stationarity for a general periodic process and showed that a necessary and sufficient condition for such stationarity is the stationarity of the so-called "lumped" vector process. The first application of periodic time series mod-

els seems to have been by hydrologists Thomas and Fiering (1962). They used the lag-one autoregressive (AR) model for modeling monthly stream flow. Since then there have been many discussions and summaries about periodic time series models: Jones and Breilsford (1967), Moss and Bryson (1974), Delleur *et al.* (1976), Pagano (1978), Troutman (1979), Tiao and Grupe (1980), Tjostheim and Paulsen (1982), Noakes *et al.* (1985), Salas *et al.* (1981,1982,1985), Thompstone *et al.* (1985b), Vecchia (1985a,1985b), Vecchia *et al.* (1983), Li and Hui (1988), Vecchia and Ballerini (1991), Anderson and Vecchia (1993), Bentarzi *et al.* (1993), Ula (1990,1993), Ula and Smadi (1997,2003), Adams and Goodwin (1995) and Anderson and Meerschaert (1997,1998), Lund and Basawa (1999,2000), and Shao and Lund (2004).

For example, Noakes *et al.* (1985) demonstrate the superiority of periodic autoregressive models among several other competitors in forecasting thirty monthly river flow time series. Troutman (1979) studies some properties of the periodic autoregressive models using a related multivariate autoregressive representation. Tiao and Grupe (1980) show how periodic autoregressive moving average models may be misspecified as homogeneous models. Because of the seasonal parameters, the estimation of periodic models is more difficult than that of the homogeneous models. Pangano (1978) dealt with moment estimation of parameters in periodic autoregressive (PAR) models. He showed that estimates of the seasonal parameters obtained by using the seasonal Yule-Walker equations possess many desirable properties, including maximum asymptotic efficiency under normality. Salas *et al.* (1982) suggest estimating the parameters of PARMA models using the seasonal Yule-Walker equations. Vecchia (1985b) proposed an algorithm for the maximum likelihood estimation for periodic ARMA models. Li (1988) developed an algorithm for exact likelihood of periodic moving average models. Anderson, Meerschaert and Vecchia (1999) developed the innovations algorithm for

estimation of PARMA model parameters. Lund and Basawa (2000) examine the recursive prediction and likelihood evaluation techniques for PARMA models, which are extensions of Ansley (1979) for ARMA series. Shao and Lund (2004) study correlation and partial auto-correlation properties PARMA time series models. Additional review of literature is given at the beginning of the remaining chapters.

1.4 Objectives of the Research

Hydrologic river flows can exhibit both heavy tails and/or nonstationarity. Therefore, the main goal of the research presented here is to fit a PARMA model to represent a given river flow data, estimate parameters, check for goodness of fit to the data, and possibly to use the fitted model for different purposes such as extreme analysis (flood and drought).

Specific objectives:

1. Anderson, Meerschaert and Vecchia (1999) developed a parameter estimation technique (known as innovations algorithm) for PARMA models with finite fourth moment as well as infinite fourth moment but finite variance. In addition, Anderson and Meerschaert (2003) provided a method for model identification when the time series have finite fourth moment. Therefore, in this part of the research, asymptotic distributions for model parameters are developed. These results provide us with a general technique for identifying PARMA models with minimum number of parameters. A detailed simulation study is conducted to test the effectiveness of the innovation estimation procedures and those asymptotics as method for PARMA model identification. We use monthly river flow data for the Fraser River at Hope in British Columbia so as to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows.

2. We describe the PARMA modeling procedure for time series with heavy tails. As an application, we use monthly flow data for the Salt River near Roosevelt in Arizona so as to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows in the heavy tail case.
3. For analysis and design of water resources systems, it is sometimes required to generate river flow data with high resolution (that is, weekly or daily values). Fitting the PARMA model to historical weekly or daily data, however, requires estimation of too many parameters which violates the principle of parsimony (model with minimum number of parameters). In an effort to obtain a parsimonious model representing these periodically stationary series, we develop the asymptotic distribution of the discrete Fourier transform of the innovation estimates and then determine those statistically significant Fourier coefficients. We also extend these results to other periodic model parameters. We demonstrate the effectiveness of the technique using simulated data from different PARMA models. An application of the technique is demonstrated through the analysis of a weekly river flow series for the Fraser River.
4. Hydrologic drought properties (such as duration and severity) of various return periods, for example, are needed to assess the degree to which a water supply system will be able to cope with future droughts and, accordingly, to plan alternative water supply strategies. They can be determined from the historical record alone by using non-parametric methods but, because the number of drought events that can be drawn from the historical sample is generally small, the “historical” drought properties have a large degree of uncertainty. In such cases, drought properties can be derived by

synthetically generating river flows at key points in the water supply system under consideration. In this part of the research, PARMA models are used to model river flows and to generate a large sequence of synthetic flows, and then drought properties are derived from the simulated samples based on the theory of runs. The applicability of these methods is demonstrated by using weekly river flow data for the Fraser River.

The remainder of the dissertation is organized as follows. Chapter 2 describes the general properties of PARMA models and the innovations algorithm for parameter estimation of PARMA models. Asymptotic distributions for the innovations estimates and model parameters are developed in Chapter 3, which also deals with PARMA model identification using those asymptotic distributions, and demonstrates the PARMA modeling techniques with a simulation study and with an example. Chapter 4 concerns modeling of river flows with heavy tails. In an effort to obtain a parsimonious model for periodically stationary series, asymptotic distributions of the discrete Fourier transform of the innovations estimates and PARMA model parameters are developed in Chapter 5 and used to determine statistically significant Fourier coefficients. In Chapter 5, we also included a simulation study and an example to demonstrate the effectiveness of the technique. Applications of PARMA models to analysis of droughts are presented in Chapter 6. Finally, conclusions are drawn and recommendations for further research are made in Chapter 7.

2 PARMA Models and Their Parameter Estimation

2.1 General

A time series is a set of observations $\{X_t\}$, each one being recorded at a specific time t . For example, $\{X_t\}$ can be river flow (daily, monthly, etc.) measurements at time t . In general, a collection of a random variables, indexed by t is referred to as a stochastic process. The observed values of a stochastic process are referred to as a realization of the stochastic process. A mathematical model representing a stochastic process is called a stochastic model or time series model. The model consists of a certain mathematical form or structure and a set of parameters. Such models are built to “reproduce” or to “resemble” the main statistical characteristics (mean, standard deviation, skewness, autocorrelation, range and run) of the time series.

Several stochastic models have been used for modeling hydrological time series in general and stream flow time series in particular (Salas *et al.*, 1981). Unfortunately, the exact mathematical model of a hydrological time series is never known. The exact model parameters are also never known, they must be estimated from limited data. Identification of the models and estimation of their parameters from the available data are often referred to in the literature as time series modeling or stochastic modeling of hydrologic time series. A systematic approach to hydrologic time series modeling may be composed of six main phases (Salas *et al.*, 1981):

1. Identification of model composition. The modeler has to decide whether the model(s) will be a univariate, a multivariate, or a combination of a univariate and a disaggregation, etc.
2. Identification of model type. Select a type of models among the various types available

in literature.

3. Identification of the model form. Determine, for example, the orders p and q of a PARMA model.
4. Estimation of model parameters. Use of proper parameter estimation method.
5. Testing the goodness of fit of the model. The model estimated in phase (4) needs to be checked to see whether it complies with certain assumptions about the model and to verify how well it represents the historic hydrologic time series.
6. Evaluation of uncertainties. It may be model uncertainty or parameter uncertainty.

The overall time series modeling is actually an iterative process with the feed back and interaction between each of the above-referred stages.

The selection or identification of the type of hydrologic stochastic model generally depends on (Salas et al, 1981):

- The modelers' input such as judgment, experience, and personal preference;
- the physical basis of the process under study (physical or stochastic generating mechanisms of the data); and
- the statistical characteristics of a given time series.

Based on the conceptual physical representation of a natural watershed, river flow processes are often modeled by the general class of ARMA processes (Salas et al, 1981). As discussed in Chapter 1, observed river flow series are characterized by seasonal variations in both the mean and covariance structure. Periodic ARMA models are, therefore, an appropriate class of models to represent the river flow series.

In this chapter and the following three chapters, we deal with the 3rd, 4th and 5th phases of hydrologic time series modeling. Parameter estimation for PARMA models is more complex than for normal ARMA models because of the higher number of parameters to be estimated. Pagano (1978) has used seasonal Yule-Walker equations (see equation (2.6)) to estimate the parameters of PAR models. Vecchia (1985) has used ML^2 for parameter estimation of PARMA models, but this only works well for lower order models. With higher-order PARMA models it becomes difficult to obtain estimates using the ML algorithm. Adams and Goodwin (1995) have employed pseudo-linear regression algorithms for parameter estimation of PARMA models with finite fourth moment. Anderson, Meerschaert and Vecchia (1999) developed the innovations algorithm for parameter estimation of an infinite moving average representation of PARMA models, which is applicable for PARMA models with finite fourth moment as well as infinite fourth moment but finite variance. The innovations algorithm parameter estimation technique is the focus of this chapter.

2.2 Mathematical Formulation of PARMA Model

A stochastic process X_t is periodically stationary (in the wide sense) if $\mu_t = EX_t$ and $\gamma_t(h) = \text{Cov}(X_t, X_{t+h}) = EX_t X_{t+h}$ for $h = 0, \pm 1, \pm 2, \dots$ are all periodic functions of time t with the same period $\nu \geq 1$. That is, for some integer ν , for $i = 0, 1, \dots, \nu - 1$, and for all integers k and h , $\mu_i = \mu_{i+k\nu}$ and $\gamma_i(h) = \gamma_{i+k\nu}(h)$. If $\nu = 1$ then the process is stationary³.

The periodic ARMA process $\{\tilde{X}_t\}$ with period ν (denoted by $\text{PARMA}_\nu(p, q)$) has representation

²Maximum likelihood(ML) estimation, also called the maximum likelihood method, is the procedure of finding the value of one or more parameters for a given statistic which makes the known likelihood distribution a maximum.

³See Definition A.2, Appendix A.

$$X_t - \sum_{j=1}^p \phi_t(j) X_{t-j} = \varepsilon_t - \sum_{j=1}^q \theta_t(j) \varepsilon_{t-j} \quad (2.1)$$

where $X_t = \tilde{X}_t - \mu_t$ and $\{\varepsilon_t\}$ is a sequence of random variables with mean zero and scale σ_t such that $\{\delta_t = \sigma_t^{-1} \varepsilon_t\}$ is independent and identically distributed (iid). The autoregressive parameters $\phi_t(j)$, the moving average parameters $\theta_t(j)$, and the residual standard deviations σ_t are all periodic functions of t with the same period $\nu \geq 1$. The residual standard deviations parameters σ_t , assumed to be strictly positive. We also assume:

(i) Finite variance: $E\varepsilon_t^2 < \infty$.

(ii) We will consider models where $E\varepsilon_t^4 < \infty$ (Finite Fourth Moment Case) and also models in which $E\varepsilon_t^4 = \infty$ (Infinite Fourth Moment Case). In the latter case, we say that the iid sequence $\{\delta_t = \sigma_t^{-1} \varepsilon_t\}$ is $RV(\alpha)$ if $P[|\delta_t| > x]$ varies regularly⁴ with the index $-\alpha$ and $P[\delta_t > x]/P[|\delta_t| > x] \rightarrow p$ for some $p \in [0, 1]$. In the case where the noise sequence has infinite fourth moment, it is assumed that the sequence is $RV(\alpha)$ with $\alpha > 2$. This assumption implies that $E|\delta_t|^\rho < \infty$ if $0 < \rho \leq \alpha$, in particular the variance of ε_t exists. With this technical condition, Anderson and Meerschaert (1997) showed that the sample autocovariance is a consistent estimator of the autocovariance, and asymptotically stable⁵ with tail index $\alpha/2$.

(iii) The model admits a causal representation

$$X_t = \sum_{j=0}^{\infty} \psi_t(j) \varepsilon_{t-j} \quad (2.2)$$

where $\psi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\psi_t(j)| < \infty$ for all t . The absolute summability of the ψ -weights ensures that (2.2) converges almost surely for all t , and in the mean-square to the same limit. The causality restriction imposes conditions on the seasonal autoregressive parame-

⁴See Definition 4.1, Section 4.3.

⁵See Chapter 4 for the definition.

ters in (2.1). These conditions may be stated by expressing the PARMA model given by (2.1) in terms of an equivalent multivariate stationary ARMA process (Tiao and Grupe, 1980; Vecchia, 1985a). For example, if $p = 1$, the causality condition for the autoregressive parameters in (2.1) is given by $|\phi_0(1)\phi_1(1)\dots\phi_{\nu-1}(1)| < 1$. It should be noted that $\psi_t(j) = \psi_{t+k\nu}(j)$ for all j .

(iv) The model also satisfies an invertibility condition

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_t(j) X_{t-j} \quad (2.3)$$

where $\pi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\pi_t(j)| < \infty$ for all t . The invertibility condition places constraints on the moving average parameters in the same way that (2.2) places constraints on the autoregressive parameters. Again, $\pi_t(j) = \pi_{t+k\nu}(j)$ for all j .

2.3 Parameter Estimation for PARMA Model

The innovations algorithm, developed for periodic time series models by Anderson, Meerschaert and Vecchia (1999), is used to obtain parameters estimates for PARMA models. It can be used to estimate parameters for the PARMA models with finite fourth moments ($E\varepsilon_t^4 < \infty$) as well as infinite fourth moment but finite variance ($E\varepsilon_t^4 = \infty$). When $E\varepsilon_t^2 = \infty$, the autocovariance and autocorrelation functions are not defined.

Let $\hat{X}_{i+k}^{(i)} = P_{\mathcal{H}_{k,i}} X_{i+k}$ denote the one-step predictors, where $\mathcal{H}_{k,i} = \overline{\text{sp}}\{X_i, \dots, X_{i+k-1}\}$, $k \geq 1$, and $P_{\mathcal{H}_{k,i}}$ is the orthogonal projection onto this space, which minimizes the mean squared error

$$v_{k,i} = \left\| X_{i+k} - \hat{X}_{i+k}^{(i)} \right\|^2 = E \left(X_{i+k} - \hat{X}_{i+k}^{(i)} \right)^2 \quad (2.4)$$

Then

$$\hat{X}_{i+k}^{(i)} = \phi_{k,1}^{(i)} X_{i+k-1} + \dots + \phi_{k,k}^{(i)} X_i, \quad k \geq 1, \quad (2.5)$$

where the vector of coefficients $\phi_k^{(i)} = \left(\phi_{k,1}^{(i)}, \dots, \phi_{k,k}^{(i)}\right)'$ solves the prediction equations

$$\Gamma_{k,i} \phi_k^{(i)} = \gamma_k^{(i)} \quad (2.6)$$

with $\gamma_k^{(i)} = (\gamma_{i+k-1}(1), \gamma_{i+k-2}(2), \dots, \gamma_i(k))'$ and

$$\Gamma_{k,i} = \left[\gamma_{i+k-\ell}(\ell - m) \right]_{\ell, m=1, \dots, k} \quad (2.7)$$

is the covariance matrix of (X_{i+k-1}, \dots, X_i) for each $i = 0, \dots, \nu - 1$. Because the process is nonstationary, the Durbin-Levinson algorithm (Brockwell and Davis, 1991) for computing $\hat{\phi}_{k,j}^{(i)}$ does not apply. However, the innovations algorithm still applies to a nonstationary process. Writing

$$\hat{X}_{i+k}^{(i)} = \sum_{j=1}^k \theta_{k,j}^{(i)} \left(X_{i+k-j} - \hat{X}_{i+k-j}^{(i)} \right) \quad (2.8)$$

yields the one step predictors in terms of the innovations $X_{i+k-j} - \hat{X}_{i+k-j}^{(i)}$. If $\sigma_i^2 > 0$ for $i = 0, \dots, \nu - 1$, then for a causal $\text{PARMA}_\nu(p, q)$ process the covariance matrix $\Gamma_{k,i}$ is non-singular for every $k \geq 1$ and each i (see Lund and Basawa, 2000). Anderson, Meerschaert, and Vecchia (1999) showed that if $EX_t = 0$ and $\Gamma_{k,i}$ is non-singular for each $k \geq 1$, then the one-step predictors \hat{X}_{i+k} , $k \geq 0$, and their mean-square errors $v_{k,i}$, $k \geq 1$, are given by

$$v_{0,i} = \gamma_i(0)$$

$$\theta_{k,k-\ell}^{(i)} = (v_{\ell,i})^{-1} \left[\gamma_{i+\ell}(k - \ell) - \sum_{j=0}^{\ell-1} \theta_{\ell,\ell-j}^{(i)} \theta_{k,k-j}^{(i)} v_{j,i} \right] \quad (2.9)$$

$$v_{k,i} = \gamma_{i+k}(0) - \sum_{j=0}^{k-1} \left(\theta_{k,k-j}^{(i)} \right)^2 v_{j,i}$$

where (2.9) is solved in the order $v_{0,i}$, $\theta_{1,1}^{(i)}$, $v_{1,i}$, $\theta_{2,2}^{(i)}$, $\theta_{2,1}^{(i)}$, $v_{2,i}$, $\theta_{3,3}^{(i)}$, $\theta_{3,2}^{(i)}$, $\theta_{3,1}^{(i)}$, $v_{3,i}$, \dots . The

results in Anderson, Meerschaert and Vecchia (1999) show that

$$\begin{aligned}\theta_{k,j}^{((i-k))} &\rightarrow \psi_i(j) \\ v_{k,\langle i-k \rangle} &\rightarrow \sigma_i^2\end{aligned}\tag{2.10}$$

$$\phi_{k,j}^{((i-k))} \rightarrow -\pi_i(j)$$

as $k \rightarrow \infty$ for all i, j , where

$$\langle j \rangle = \begin{cases} j - \nu [j/\nu] & \text{if } j = 0, 1, \dots \\ \nu + j - \nu [j/\nu + 1] & \text{if } j = -1, -2, \dots \end{cases}$$

and $[\cdot]$ is the greatest integer function. For example, for monthly data X_t , $\langle t \rangle$ is the month of the year = 0, 1, ..., 11. Given N_y years of data, $N = N_y\nu$, define the sample mean

$$\hat{\mu}_i = N_y^{-1} \sum_{k=0}^{N_y-1} \tilde{X}_{k\nu+i}\tag{2.11}$$

the sample autocovariance

$$\hat{\gamma}_i(\ell) = (N_y - m)^{-1} \sum_{j=0}^{N_y-1-m} X_{j\nu+i} X_{j\nu+i+\ell}\tag{2.12}$$

and the sample autocorrelation

$$\hat{\rho}_i(j) = \frac{\hat{\gamma}_i(j)}{\sqrt{\hat{\gamma}_i(0)\hat{\gamma}_{i+j-m\nu}(0)}}\tag{2.13}$$

where $X_t = \tilde{X}_t - \hat{\mu}_t$, $m = [(i + \ell)/\nu]$ and $[\cdot]$ is the greatest integer function. If we replace the autocovariances in the prediction equation (2.6) with their corresponding sample autocovariances, we obtain the estimator $\hat{\phi}_{k,j}^{(i)}$ of $\phi_{k,j}^{(i)}$. If we also replace the autocovariances in (2.9) with corresponding sample autocovariances (2.12), we obtain the innovations estimates $\hat{\theta}_{k,l}^{(i)}$. The consistency of these estimators was also established in the following result due to Anderson, Meerschaert and Vecchia (1999).

Theorem 2.1 (*Finite Fourth Moment Case*) Suppose that $\{X_t\}$ be the mean zero PARMA process with period ν given by (2.1) and that $E\varepsilon_t^4 < \infty$. Assume that the spectral density matrix⁶, $f(\lambda)$, of its equivalent vector ARMA process is such that $mzz' \leq z'f(\lambda)z \leq Mzz'$, $-\pi \leq \lambda \leq \pi$, for some m and M such that $0 < m \leq M < \infty$ and for all z in R^ν . If k is chosen as a function of the sample size N_y so that $k^2/N_y \rightarrow 0$ as $N_y \rightarrow \infty$ and $k \rightarrow \infty$, then

$$\begin{aligned} \hat{\theta}_{k,j}^{((i-k))} &\xrightarrow{P} \psi_i(j) \\ \hat{\nu}_{k,(i-k)} &\xrightarrow{P} \sigma_i^2 \end{aligned} \tag{2.14}$$

$$\hat{\phi}_{k,j}^{((i-k))} \xrightarrow{P} -\pi_i(j)$$

for all i, j where “ \xrightarrow{P} ” denotes convergence in probability⁷.

PROOF. See Anderson, Meerschaert and Vecchia (1999) for a proof.

This yields a practical method for estimating the model parameters, in the classical case of finite fourth moments. The asymptotic results in Chapter 3 can then be used to determine which of these model parameters are statistically significantly different from zero.

Define

$$a_{N_y} = \inf\{x : P(|\delta_t| > x) < 1/N_y\},$$

where

$$a_{N_y}^{-2} = \sum_{t=0}^{N_y-1} \delta_{t\nu+i}^2 \Rightarrow S^{(i)},$$

$S^{(i)}$ is $\alpha/2$ -stable law, and “ \Rightarrow ” denotes convergence in distribution⁸.

⁶See Section A-3, Appendix A.

⁷See Definition A.10, Appendix A.

⁸See Definition A.11, Appendix A.

Theorem 2.2 (*Infinite Fourth Moment Case*) Suppose that $\{X_t\}$ be the mean zero PARMA process with period ν given by (2.1) and $\{\delta_t\}$ is $RV(\alpha)$ with $2 < \alpha \leq 4$. Assume that the spectral density matrix, $f(\lambda)$, of its equivalent vector ARMA process is such that $mzz' \leq z'f(\lambda)z \leq Mzz'$, $-\pi \leq \lambda \leq \pi$, for some m and M such that $0 < m \leq M$ and for all z in R^ν . If k is chosen as a function of the sample size N_y so that $k^{5/2}a_{N_y}^2/N_y \rightarrow 0$ as $N_y \rightarrow \infty$ and $k \rightarrow \infty$, then

$$\begin{aligned} \hat{\theta}_{k,j}^{((i-k))} &\xrightarrow{P} \psi_i(j) \\ \hat{\nu}_{k,\langle i-k \rangle} &\xrightarrow{P} \sigma_i^2 \end{aligned} \tag{2.15}$$

$$\hat{\phi}_{k,j}^{((i-k))} \xrightarrow{P} -\pi_i(j)$$

for all j and every $i = 0, 1, \dots, \nu - 1$, where “ \xrightarrow{P} ” denotes convergence in probability.

PROOF. See Anderson, Meerschaert and Vecchia (1999) for a proof.

Theorem 2.2 yields a practical method for estimating the model parameters, in the case of infinite fourth moments.

3 PARMA Model Identification: Finite Fourth Moment Case

3.1 General

The PARMA modeling procedure usually involves three main steps: model identification, parameter estimation and diagnostic checking. The goal of PARMA model identification is to identify the order of the PARMA model. Ula (2003) has used the cut-off property of the periodic autocorrelation function for the identification of the orders of periodic MA (PMA) models, which is an extension of Box-Jenkins (Box and Jenkins, 1976) identification techniques for ordinary MA and AR models. Anderson and Vecchia (1993) computed the asymptotic distributions of the moment estimates of the periodic autocovariance and autocorrelation functions of a PARMA process and used them for detecting periodicity in autocorrelations and model parameters. Also, they derived the asymptotic distribution of the Fourier-transformed sample periodic autocorrelation function and applied those asymptotics for identification of a parsimonious PARMA model.

Anderson and Meerschaert (2003) provided an asymptotic distribution for innovation estimates of PARMA models in the case where the innovations have finite fourth moment. These results are used for the identification of the order of PARMA models. However, such techniques often do not adequately identify the order of the model when the model is mixed in the sense that both the autoregressive (AR) and moving average (MA) components are present. If the innovations algorithm is supplemented by modeler experience or the Akaike's information criterion⁹ (AIC), then an approximate order of the mixed PARMA model can be obtained. This chapter, which is an extension of previous work by Anderson and Meerschaert,

⁹See Section A-5, Appendix A.

has three objectives. The first is to develop asymptotic distributions for PARMA model parameters for the finite fourth moment case. These asymptotic results provide us with a general technique for computing asymptotic confidence regions for the model parameters from the innovation algorithm estimates, and for identifying PARMA models with minimum number of parameters. The second is to demonstrate the effectiveness of the innovation estimation procedures and those asymptotics as a method for PARMA model identification by using simulated data from different PARMA models for the finite fourth moment case. Finally we use monthly river flow data for the Fraser River at Hope in British Columbia so as to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows.

3.2 Asymptotic Distributions for Moving Average Processes

Any mean-centered finite variance periodic ARMA process (2.1) can be expressed in terms of a periodic moving average

$$X_t = \sum_{j=-\infty}^{\infty} \psi_t(j) \varepsilon_{t-j} \quad (3.1)$$

where $\delta_t = \sigma_t^{-1} \varepsilon_t$ is a sequence of iid random variables with mean zero and variance one.

We also assume that the moving average parameters $\psi_t(j)$ are periodic in t with the same period ν , and that $\sum_{j=-\infty}^{\infty} |\psi_t(j)| < \infty$ for all t . Also, we consider a finite fourth moment case: $E\varepsilon_t^4 < \infty$.

For any periodically stationary time series, we can construct an equivalent (stationary) vector moving average process in the following way: Let $Z_t = (\delta_{t\nu}, \dots, \delta_{(t+1)\nu-1})'$ and $Y_t =$

$(X_{t\nu}, \dots, X_{(t+1)\nu-1})'$, so that

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j} \quad (3.2)$$

where Ψ_j is the $\nu \times \nu$ matrix with ij entry $\sigma_i \psi_i(t\nu + i - j)$, and we number the rows and columns $0, 1, \dots, \nu - 1$ for ease of notation.

We define the sample autocovariance matrix by $\hat{\Gamma}(h) = N_y^{-1} \sum_{t=0}^{N_y-1} (Y_t - \hat{\mu})(Y_{t+h} - \hat{\mu})'$, where $\hat{\mu} = N_y^{-1} \sum_{t=0}^{N_y-1} Y_t$ is the sample mean and the autocovariance matrix by $\Gamma(h) = E(Y_t - \mu)(Y_{t+h} - \mu)'$, where $\mu = EY_t$. Note that the ij entry of $\Gamma(h)$ is $\gamma_i(h\nu + j - i)$ and likewise for $\hat{\Gamma}(h)$. The autocorrelation matrix $R(h)$ has ij entry equal to $\rho_i(h\nu + j - i)$ and likewise for the sample autocorrelation matrix $\hat{R}(h)$. The ij term of $R(h)$ is also called the cross correlation of the i and j components of the vector process at lag h . In this application it represents the correlation between season i at year t and season j at year $t + h$.

Theorem 3.1 *Let $X_t = \tilde{X}_t - \mu_t$, where $\{X_t\}$ is the periodic moving average process (3.1) and μ_t is a periodic mean function with period ν . Then, for $\sigma_t^2 = E\varepsilon_t^2 < \infty$, $\hat{\mu}_t$ is a consistent estimator of μ_t and*

$$N_y^{1/2} \{\hat{\mu} - \mu\} \Rightarrow \mathcal{N}(0, C) \quad (3.3)$$

where $N = N_y \nu$, $\mathcal{N}(0, C)$ is a Gaussian random vector with mean zero and covariance

$$C = (\sum_j \Psi_j)(\sum_j \Psi_j)',$$

$$\begin{aligned} \hat{\mu} &= [\hat{\mu}_0, \hat{\mu}_1, \dots, \hat{\mu}_{\nu-1}]^T \\ \mu &= [\mu_0, \mu_1, \dots, \mu_{\nu-1}]^T \end{aligned} \quad (3.4)$$

PROOF. See Anderson and Meerschaert (1998).

This theorem can be used for constructing confidence regions for μ . We adapt the procedure of Brockwell and Davis (1991, p. 407) to obtain the simultaneous $(1 - \alpha)100\%$ confidence region for μ_t . By Theorem 3.1, we have $N_y^{1/2} \{\hat{\mu}_i - \mu_i\} \Rightarrow \mathcal{N}(0, v_i)$, where v_i is a

variance

$$v_i = 2\pi f_i(0) = \sum_{j=-\infty}^{\infty} \gamma_i(j) \quad (3.5)$$

and $f_i(\omega)$ is the spectral density for the i th seasonal component. The lag window estimator

$$2\pi f_i(0) = \sum_{|h| \leq r} \left(1 - \frac{|h|}{r}\right) \hat{\gamma}_i(h) \quad (3.6)$$

for v_i is consistent when $r = r_n \rightarrow \infty$ in such a way that $r_n/n \rightarrow 0$. Then, with probability $1 - \alpha$,

$$|\hat{\mu}_i - \mu_i| \leq \Phi_{1-\alpha/(2\nu)} \sqrt{2\pi \hat{f}_i(0)/N_y} \quad (3.7)$$

is approximately true for large r and n . Here Φ_p is the p percentile of a standard normal distribution.

For finite fourth moment case, Anderson and Vecchia (1993) compute the asymptotic distributions of the moment estimates of the periodic autocovariance function $\gamma_t(\ell)$ and of the periodic autocorrelation function $\rho_t(\ell)$, where $\rho_t(\ell) = \gamma_t(\ell)/\{\gamma_t(0)\gamma_{t+\ell}(0)\}^{1/2}$. The joint asymptotic distribution of the estimated periodic autocovariances are given in Theorem 3.2.

Theorem 3.2 *Let $X_t = \tilde{X}_t - \mu_t$, where X_t is the periodic moving average process (3.1) and μ_t is a periodic mean function with period ν . Then, for any nonnegative integers j and h with $j \neq h$,*

$$N^{1/2} \begin{pmatrix} \hat{\gamma}(j) - \gamma(j) \\ \hat{\gamma}(h) - \gamma(h) \end{pmatrix} \Rightarrow \mathcal{N} \left(0, \begin{pmatrix} O_{jj} & O_{jh} \\ O_{hj} & O_{hh} \end{pmatrix} \right) \quad (3.8)$$

where $N = N_y \nu$, $\mathcal{N}(\mu, O)$ denotes the multivariate normal distribution with mean μ and covariance matrix O ,

$$\begin{aligned} \hat{\gamma}(\ell) &= [\hat{\gamma}_0(\ell), \hat{\gamma}_1(\ell), \dots, \hat{\gamma}_{\nu-1}(\ell)]^T \\ \gamma(\ell) &= [\gamma_0(\ell), \gamma_1(\ell), \dots, \gamma_{\nu-1}(\ell)]^T \end{aligned} \quad (3.9)$$

and, for $i, \ell = 0, 1, \dots, \nu - 1$,

$$(O_{jh})_{il} = \sum_{n=-\infty}^{\infty} \{ \gamma_i(n\nu + \ell - i) \gamma_{i+j}(n\nu + \ell - i - j + h) + \gamma_i(n\nu + \ell - i + h) \gamma_{i+j}(n\nu + \ell - i - j) \} \quad (3.10)$$

$$O_{jh} = \sum_{n=-\infty}^{\infty} \{ G_n \Pi^j G_{n+h-j} \Pi^{n-j} + G_{n+h} \Pi^j G_{n-j} \Pi^{n-j} \}$$

where

$$G_n = \text{diag}\{\gamma_0(n), \gamma_1(n), \dots, \gamma_{\nu-1}(n)\} \quad (3.11)$$

and Π an orthogonal $\nu \times \nu$ cyclic permutation matrix

$$\Pi = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \quad (3.12)$$

Note that

$$\Pi^j \text{diag}\{a_0, a_1, \dots, a_{\nu-1}\} \Pi^{-j} = \text{diag}\{a_j, a_{j+1}, \dots, a_{j+\nu-1}\} \quad (3.13)$$

where $\{a_t\}$ is any sequence of constants satisfying $a_{t+k\nu} = a_t$ for all t .

PROOF. See Anderson and Vecchia (1993).

Theorem 3.3 *Under the assumption of Theorem 3.2,*

$$N_y^{1/2} \begin{pmatrix} \hat{\rho}(j) - \rho(j) \\ \hat{\rho}(h) - \rho(h) \end{pmatrix} \Rightarrow \mathcal{N} \left(0, \begin{pmatrix} W_{jj} & W_{jh} \\ W_{hj} & W_{hh} \end{pmatrix} \right) \quad (3.14)$$

where $N = N_y \nu$, j and h are positive integers with $j \neq h$,

$$\begin{aligned} \hat{\rho}(\ell) &= [\hat{\rho}_0(\ell), \hat{\rho}_1(\ell), \dots, \hat{\rho}_{\nu-1}(\ell)]^T \\ \rho(\ell) &= [\rho_0(\ell), \rho_1(\ell), \dots, \rho_{\nu-1}(\ell)]^T \end{aligned} \quad (3.15)$$

and, for $i, \ell = 0, 1, \dots, \nu - 1$,

$$W_{jh} = \sum_{n=-\infty}^{\infty} \{L_n \Pi^j L_{n+h-j} \Pi^{n-j} + L_{n+h} \Pi^j L_{n-j} \Pi^{n-j} - L_j (I + \Pi^j) L_n L_{n+h} \Pi^n - L_n \Pi^j L_{n-j} \Pi^{n-j} (I + \Pi^{-h}) L_h + \frac{1}{2} L_j (I + \Pi^j) L_n^2 \Pi^n (I + \Pi^{-h}) L_h\} \quad (3.16)$$

where

$$L_n = \text{diag}\{\rho_0(n), \rho_1(n), \dots, \rho_{\nu-1}(n)\} \quad (3.17)$$

and Π an orthogonal $\nu \times \nu$ cyclic permutation matrix.

PROOF. See Anderson and Vecchia (1993) for a proof.

For $\nu = 1$, the expression for W_{jh} in (3.16) reduces to

$$W_{jh} = \sum_{n=-\infty}^{\infty} \{\rho(n)\rho(n+h-j) + \rho(n+h)\rho(n-j) - 2\rho(j)\rho(n)\rho(n+h) - 2\rho(n)\rho(n-j)\rho(h) + 2\rho(j)\rho(h)\rho^2(n)\} \quad (3.18)$$

See also Theorem A.15, Appendix A.

Anderson and Meerschaert (2003) also provide the asymptotic distribution for the innovations estimates of the parameters in a periodically stationary series (2.2) with period $\nu \geq 1$ under the classical assumption that the innovations have finite fourth moment.

Theorem 3.4 *Suppose that a periodically stationary moving average (2.2) is causal, invertible, $E\varepsilon_t^4 < \infty$, and that for some $0 < g \leq G < \infty$ we have $gz'z \leq z'f(\lambda)z \leq Gz'z$, for all $-\pi \leq \lambda \leq \pi$, and for all z in R^ν , where $f(\lambda)$ is the spectral density matrix of equivalent vector moving average process (3.2). If $k = k(N_y) \rightarrow \infty$ as $N_y \rightarrow \infty$ with $k^3/N_y \rightarrow 0$ and*

$$N_y^{1/2} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \rightarrow 0 \quad \text{for } \ell = 0, 1, \dots, \nu - 1 \quad (3.19)$$

then for any fixed positive integer D

$$N_y^{1/2} (\hat{\theta}_{k,u}^{(i-k)} - \psi_i(u) : u = 1, \dots, D, i = 0, \dots, \nu - 1) \Rightarrow \mathcal{N}(0, V) \quad (3.20)$$

where

$$V = A \operatorname{diag} (\sigma_0^2 D^{(0)}, \dots, \sigma_{\nu-1}^2 D^{(\nu-1)}) A' \quad (3.21)$$

$$A = \sum_{n=0}^{D-1} E_n \Pi^{[D\nu-n(D+1)]} \quad (3.22)$$

$$D^{(i)} = \operatorname{diag} (\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \dots, \sigma_{i-D}^{-2}) \quad (3.23)$$

$$E_n = \operatorname{diag} \left\{ \underbrace{0, \dots, 0}_0, \underbrace{\psi_0(n), \dots, \psi_0(n)}_{D-n}, \underbrace{0, \dots, 0}_n, \underbrace{\psi_1(n), \dots, \psi_1(n)}_{D-n} \right. \\ \left. , \dots, \underbrace{0, \dots, 0}_n, \underbrace{\psi_{\nu-1}(n), \dots, \psi_{\nu-1}(n)}_{D-n} \right\} \quad (3.24)$$

and Π an orthogonal $D\nu \times D\nu$ cyclic permutation matrix.

PROOF. See Anderson and Meerschaert (2003) for a proof.

Note that Π^0 is $D\nu \times D\nu$ identity matrix and $\Pi^{-\ell} \equiv (\Pi')^\ell$. Matrix multiplication yields the following corollary.

Corollary 3.5 *Regarding Theorem 3.4, in particular, we have that*

$$N_y^{1/2} (\hat{\theta}_{k,u}^{(i-k)} - \psi_i(u)) \Rightarrow \mathcal{N} \left(0, \sum_{n=0}^{u-1} \frac{\sigma_{i-n}^2}{\sigma_{i-u}^2} \psi_i^2(n) \right) \quad (3.25)$$

PROOF. See Anderson and Meerschaert (2003) for a proof.

Remark 3.6 *Corollary 3.5 also holds the asymptotic result for second-order stationary process where the period is just $\nu = 1$. In this case $\sigma_i^2 = \sigma^2$ so (3.25) becomes*

$$N_y^{1/2} (\hat{\theta}_{k,u} - \psi(u)) \Rightarrow \mathcal{N} \left(0, \sum_{n=0}^{u-1} \psi^2(n) \right) \quad (3.26)$$

The α -level test statistic rejects the null hypothesis ($H_o : \psi_i(u) = 0$) in favor of the alternative ($H_a : \psi_i(u) \neq 0$) if $|Z| > z_{\alpha/2}$. The p -value of this test is

$$p = P(|Z| > |z|) \quad \text{where } Z \sim \mathcal{N}(0, 1) \quad (3.27)$$

and

$$Z = \frac{N_y^{1/2} \hat{\theta}_{k,u}^{(i-k)}}{W}, \quad W^2 = \frac{\sum_{n=0}^{u-1} \hat{v}_{k, \langle i-k-n \rangle} \left(\hat{\theta}_{k,n}^{(i-k)} \right)^2}{\hat{v}_{k, \langle i-k-u \rangle}} \quad (3.28)$$

The joint asymptotic distribution of the innovation estimates are given in Theorem 3.7.

Theorem 3.7 *average process(2) and μ_t is a Let $\hat{\theta}_{k,\ell}^{(i-k)} = \hat{\psi}_i(\ell)$, under the assumption of Theorem 3.4, for any nonnegative integers j and h with $j \neq h$,*

$$N_y^{1/2} \begin{pmatrix} \hat{\psi}(j) - \psi(j) \\ \hat{\psi}(h) - \psi(h) \end{pmatrix} \Rightarrow \mathcal{N} \left(0, \begin{pmatrix} V_{jj} & V_{jh} \\ V_{hj} & V_{hh} \end{pmatrix} \right) \quad (3.29)$$

where $\mathcal{N}(\mu, V)$ denotes the multivariate normal distribution with mean μ and covariance matrix V ,

$$\begin{aligned} \hat{\psi}(\ell) &= [\hat{\psi}_0(\ell), \hat{\psi}_1(\ell), \dots, \hat{\psi}_{\nu-1}(\ell)]^T \\ \psi(\ell) &= [\psi_0(\ell), \psi_1(\ell), \dots, \psi_{\nu-1}(\ell)]^T \end{aligned} \quad (3.30)$$

and for $x = \text{Min}(h, j)$

$$V_{jh} = \sum_{n=0}^{x-1} \left\{ F_{j-1-n} \Pi^{(\nu-1)(j-1-n)} B_{n+1} \left(F_{h-1-n} \Pi^{(\nu-1)(h-1-n)} \right)^T \right\} \quad (3.31)$$

where

$$\begin{aligned} F_n &= \text{diag}\{\psi_0(n), \psi_1(n), \dots, \psi_{\nu-1}(n)\} \\ B_n &= \text{diag}\{\sigma_0^2 \sigma_{0-n}^{-2}, \sigma_1^2 \sigma_{1-n}^{-2}, \dots, \sigma_{\nu-1}^2 \sigma_{\nu-1-n}^{-2}\} \end{aligned} \quad (3.32)$$

and Π an orthogonal $\nu \times \nu$ cyclic permutation matrix.

PROOF. This is simply a rearrangement of rows and columns of the matrix V in (3.21).

Vector Difference Equation for the ψ -weights of the PARMA process

For a $\text{PARMA}_\nu(p, q)$ model given by (2.1), we develop a vector difference equation for the ψ -weights of the PARMA process so that we can determine feasible values of p and q . For fixed p and q ($p + q = m$; assuming m statistically significant values of $\psi_t(j)$), to determine

the parameters $\phi_t(\ell)$ and $\theta_t(\ell)$, (2.2) is substituted in to (2.1) to obtain

$$\sum_{j=0}^{\infty} \psi_t(j) \varepsilon_{t-j} - \sum_{\ell=1}^p \phi_t(\ell) \sum_{j=0}^{\infty} \psi_{t-\ell}(j) \varepsilon_{t-\ell-j} = \varepsilon_t - \sum_{\ell=1}^q \theta_t(\ell) \varepsilon_{t-\ell} \quad (3.33)$$

and then the coefficients on both sides are equated in order to calculate the $\phi_t(\ell)$ and $\theta_t(\ell)$.

$$\begin{aligned} \psi_t(0) &= 1 \\ \psi_t(1) - \phi_t(1)\psi_{t-1}(0) &= -\theta_t(1) \\ \psi_t(2) - \phi_t(1)\psi_{t-1}(1) - \phi_t(2)\psi_{t-2}(0) &= -\theta_t(2) \\ \psi_t(3) - \phi_t(1)\psi_{t-1}(2) - \phi_t(2)\psi_{t-2}(1) - \phi_t(3)\psi_{t-3}(0) &= -\theta_t(3) \\ &\vdots = \vdots \end{aligned} \quad (3.34)$$

where we take $\phi_t(\ell) = 0$ for $\ell > p$, and $\theta_t(\ell) = 0$ for $\ell > q$. Hence the ψ -weights satisfy the homogeneous difference equation given by

$$\begin{cases} \psi_t(j) - \sum_{k=1}^p \phi_t(k) \psi_{t-k}(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi_t(j) - \sum_{k=1}^j \phi_t(k) \psi_{t-k}(j-k) = -\theta_t(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \quad (3.35)$$

Defining

$$\begin{aligned} A_\ell &= \text{diag}\{\phi_0(\ell), \phi_1(\ell), \dots, \phi_{\nu-1}(\ell)\} \\ \psi(j) &= (\psi_0(j), \psi_1(j), \dots, \psi_{\nu-1}(j))^T \\ \theta(j) &= (\theta_0(j), \theta_1(j), \dots, \theta_{\nu-1}(j))^T \\ \psi_{jk}(j-k) &= (\psi_{-k}(j-k), \psi_{-k+1}(j-k), \dots, \psi_{-k+\nu-1}(j-k))^T \end{aligned} \quad (3.36)$$

then (3.35) becomes the vector difference equation

$$\begin{cases} \psi(j) - \sum_{k=1}^p A_k \psi_{jk}(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi(j) - \sum_{k=1}^j A_k \psi_{jk}(j-k) = -\theta(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \quad (3.37)$$

Note that $\psi_t(0) = 1$

We can obtain a difference equation for $\psi_{j0}(j-0) = \psi(j)$ by introducing the orthogonal $\nu \times \nu$ cyclic permutation matrix, and observing that $\psi_{jk}(j-k) = \Pi^{-k}\psi(j-k)$. So, (3.37)

becomes

$$\begin{cases} \psi(j) - \sum_{k=1}^p A_k \Pi^{-k} \psi(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi(j) - \sum_{k=1}^j A_k \Pi^{-k} \psi(j-k) = -\theta(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \quad (3.38)$$

The vector difference equation (3.38) can be helpful for the analysis of higher order PARMA models using matrix algebra.

(a) Periodic Moving Average Processes

The vector difference equation for the $\text{PMA}_\nu(q)$ process given by (3.38) is

$$\begin{cases} \psi(j) = -\theta(j) & 0 \leq j \leq q \\ \psi(j) = 0 & j > q \end{cases} \quad (3.39)$$

We see that the innovation estimates of a periodic moving average process of order q has a “cut-off” at lag q . Therefore, Corollary 3.5 can be directly applied to identify the order of the PMA process.

$$N_y^{1/2}(-\hat{\theta}(j) + \theta(j)) \Rightarrow \mathcal{N}(0, V_{jj}) \quad (3.40)$$

where V_{jj} is obtained from (3.31).

(b) Periodic Autoregressive Processes

The vector difference equation for the $\text{PAR}_\nu(p)$ process given by (3.38) is

$$\psi(j) = \sum_{k=1}^p A_k \Pi^{-k} \psi(j-k) \quad j \geq p \quad (3.41)$$

For $\text{PAR}_\nu(1)$, $\psi(1) = A_1 \Pi^{-1} \psi(0) = \phi$ where $\phi = \{\phi_0, \phi_1, \dots, \phi_{\nu-1}\}^T$.

(c) $\text{PARMA}_\nu(1, 1)$ Process

For higher order PAR or PARMA models, it is difficult to obtain the explicit solution for $\phi(\ell)$ and $\theta(\ell)$, and hence the model identification is a complicated problem. However, we develop herein an asymptotic distribution for $\text{PARMA}_\nu(1, 1)$ model parameters.

Consider a PARMA $_{\nu}(1, 1)$ model

$$X_t = \phi_t X_{t-1} + \varepsilon_t - \theta_t \varepsilon_{t-1} \quad (3.42)$$

where $\{\sigma_i^{-1} \varepsilon_{k\nu+i}\}$ form an iid sequences of normal random variables with mean zero and standard deviation one. Then from (3.34) given the innovations coefficients, we can solve for ϕ_t and θ_t using the following equations.

$$\psi_t(0) = 1 \quad (3.43)$$

$$\theta_t = \phi_t - \psi_t(1) \quad (3.44)$$

$$\psi_t(2) = \phi_t \psi_{t-1}(1) \quad (3.45)$$

We now use Theorem 3.7 to derive the asymptotic distribution of the estimates of the autoregressive as well as the moving average parameters (ϕ_t, θ_t) .

Proposition 3.8 *Suppose that \mathbf{X}_n is AN $(\mu, c_n^2 \Sigma)$ where Σ is a symmetric non-negative definite matrix and $c_n \rightarrow 0$ as $n \rightarrow \infty$. If $\mathbf{g}(\mathbf{X}) = (g_1(\mathbf{X}), \dots, g_m(\mathbf{X}))'$ is a mapping from \mathbb{R}^k into \mathbb{R}^m such that each $g_i(\cdot)$ is continuously differentiable in a neighborhood of μ , and if $D\Sigma D'$ has all of its diagonal elements non-zero, where D is the $m \times k$ matrix $[(\partial g_i / \partial x_j)(\mu)]$, then $\mathbf{g}(\mathbf{X}_n)$ is AN $(g(\mu), c_n^2 D\Sigma D')$*

PROOF. See Proposition 6.4.3 of Brockwell and Davis (1991) for a proof.

Theorem 3.9 *Under the assumption of Theorem 3.7, we have*

$$N_y^{1/2} (\hat{\phi} - \phi) \Rightarrow \mathcal{N}(0, Q) \quad (3.46)$$

where

$$\begin{aligned} \hat{\phi} &= [\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_{\nu-1}]^T \\ \phi &= [\phi_0, \phi_1, \dots, \phi_{\nu-1}]^T \end{aligned} \quad (3.47)$$

and the $\nu \times \nu$ matrix Q is defined by

$$Q = \sum_{k,\ell=1}^2 H_\ell V_{\ell k} H_k^T \quad (3.48)$$

where the matrix H is a $\nu \times 2\nu$ matrices of partial derivatives

$$H = (H_1, H_2) = \left(\frac{\partial \phi_\ell}{\partial \psi_m(1)}, \frac{\partial \phi_\ell}{\partial \psi_m(2)} \right)_{\ell, m = 0, 1, \dots, \nu-1} \quad (3.49)$$

H can be expressed using $\nu \times \nu$ permutation matrix as

$$H = (-F_2 \Pi^{-1} F_1^{-2}, \Pi^{-1} F_1^{-1} \Pi) \quad (3.50)$$

and for $\ell, k = 1, 2$, $V_{\ell k}$ is given by

$$\begin{aligned} V_{11} &= B_1 \\ V_{12} &= B_1 \Pi^{-(\nu-1)} F_1 = B_1 \Pi F_1 \\ V_{21} &= F_1 \Pi^{\nu-1} B_1 = F_1 \Pi^{-1} B_1 \\ V_{22} &= B_2 + F_1 \Pi^{\nu-1} B_1 \Pi^{-(\nu-1)} F_1 = B_2 + F_1 \Pi^{-1} B_1 \Pi F_1 \end{aligned} \quad (3.51)$$

in which B_n and F_n are obtained from 3.32.

PROOF. Using Proposition 3.8, together with the relation $\phi_t = \frac{\psi_t(2)}{\psi_{t-1}(1)}$ and Theorem 3.7 for $j = 1$ and $h = 2$ we have that (3.46) holds with

$$Q = H V H^T$$

where

$$\begin{aligned} \Sigma = V &= \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \\ \mathbf{g}(\mu) = \phi &\text{ where } \mu = \begin{pmatrix} \psi(1) \\ \psi(2) \end{pmatrix} \end{aligned}$$

$$g \begin{pmatrix} \psi_0(1) \\ \vdots \\ \psi_{\nu-1}(1) \\ \psi_0(2) \\ \vdots \\ \psi_{\nu-1}(2) \end{pmatrix} = \begin{pmatrix} \phi_0 \\ \vdots \\ \phi_{\nu-1} \end{pmatrix} = \begin{pmatrix} \frac{\psi_0(2)}{\psi_{\nu-1}(1)} \\ \vdots \\ \frac{\psi_{\nu-1}(2)}{\psi_{\nu-2}(1)} \end{pmatrix} = \begin{pmatrix} g_0 \\ \vdots \\ g_{\nu-1} \end{pmatrix}$$

and $\mathbf{g}(\mathbf{X}_n) = \hat{\phi}$ where $\mathbf{X}_n = \begin{pmatrix} \hat{\psi}(1) \\ \hat{\psi}(2) \end{pmatrix}$

Then,

$$\begin{aligned} D = H &= \begin{pmatrix} \frac{\partial \phi_0}{\partial \psi_0(1)} & \cdots & \frac{\partial \phi_0}{\partial \psi_{\nu-1}(1)} & \frac{\partial \phi_0}{\partial \psi_0(2)} & \cdots & \frac{\partial \phi_0}{\partial \psi_{\nu-1}(2)} \\ \vdots & \vdots & \vdots & \vdots & & \\ \frac{\partial \phi_{\nu-1}}{\partial \psi_{\nu-1}(1)} & \cdots & \frac{\partial \phi_{\nu-1}}{\partial \psi_{\nu-1}(1)} & \frac{\partial \phi_{\nu-1}}{\partial \psi_0(2)} & \cdots & \frac{\partial \phi_{\nu-1}}{\partial \psi_{\nu-1}(2)} \end{pmatrix} \\ &= \left(\left[\frac{\partial \phi_\ell}{\partial \psi_m(1)} \right]_{\ell,m=0,\dots,\nu-1} \quad \left[\frac{\partial \phi_\ell}{\partial \psi_m(2)} \right]_{\ell,m=0,\dots,\nu-1} \right) \\ &= (H_1 \quad H_2) \quad \text{same as (3.49)} \end{aligned}$$

Therefore,

$$\begin{aligned} Q &= HVH^T \\ &= (H_1 \quad H_2) \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} H_1^T \\ H_2^T \end{pmatrix} \\ &= (H_1 V_{11} + H_2 V_{21} \quad H_1 V_{12} + H_2 V_{22}) \begin{pmatrix} H_1^T \\ H_2^T \end{pmatrix} \\ &= H_1 V_{11} H_1^T + H_2 V_{21} H_1^T + H_1 V_{12} H_2^T + H_2 V_{22} H_2^T \\ &= \sum_{k,\ell=1}^2 H_\ell V_{\ell k} H_k^T \quad \text{same as (3.48)} \end{aligned}$$

where $V_{\ell k}$ is from (3.31).

For $\phi_i = \frac{\psi_i(2)}{\psi_{i-1}(1)}$, we obtain

$$H_1 = \begin{pmatrix} 0 & 0 & \dots & 0 & \frac{-\psi_0(2)}{\psi_{\nu-1}^2(1)} \\ \frac{-\psi_1(2)}{\psi_0^2(1)} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \frac{\psi_{\nu-1}(2)}{\psi_{\nu-2}^2(1)} & 0 \end{pmatrix} = -F_1 \Pi^{-1} F_1^{-2} \quad \text{same as (3.50)}$$

and

$$H_2 = \begin{pmatrix} \frac{1}{\psi_{\nu-1}(1)} & 0 & \dots & 0 \\ 0 & \frac{1}{\psi_0(1)} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{1}{\psi_{\nu-2}(1)} \end{pmatrix} = \Pi^{-1} F_1^{-1} \Pi \quad \text{same as (3.50)}$$

From (3.32)

$$B_1 = \begin{pmatrix} \frac{\sigma_0^2}{\sigma_{\nu-1}^2} & 0 & \dots & 0 \\ 0 & \frac{\sigma_1^2}{\sigma_0^2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} \end{pmatrix} \quad F_1 = \begin{pmatrix} \psi_0(1) & 0 & \dots & 0 \\ 0 & \psi_1(1) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \psi_{\nu-1}(1) \end{pmatrix}$$

and

$$B_2 = \begin{pmatrix} \frac{\sigma_0^2}{\sigma_{\nu-2}^2} & 0 & \dots & 0 \\ 0 & \frac{\sigma_1^2}{\sigma_{\nu-1}^2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-3}^2} \end{pmatrix}$$

Using (3.31), we obtain (3.51) and then we can compute

$$V_{11} = B_1 = \begin{pmatrix} \frac{\sigma_0^2}{\sigma_{\nu-1}^2} & 0 & \dots & 0 \\ 0 & \frac{\sigma_1^2}{\sigma_0^2} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} \end{pmatrix}$$

$$V_{12} = B_1 \Pi F_1 = \begin{pmatrix} 0 & \frac{\psi_1(1)\sigma_0^2}{\sigma_{\nu-1}^2} & 0 & \dots & 0 \\ 0 & 0 & \frac{\psi_2(1)\sigma_1^2}{\sigma_0^2} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\psi_0(1)\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$V_{21} = F_1 \Pi^{-1} B_1 = \begin{pmatrix} 0 & 0 & 0 & \dots & \frac{\psi_0(1)\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} \\ \frac{\psi_1(1)\sigma_0^2}{\sigma_{\nu-1}^2} & 0 & 0 & \dots & 0 \\ 0 & \frac{\psi_2(1)\sigma_1^2}{\sigma_0^2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$$V_{22} = B_2 + F_1 \Pi^{-1} B_1 \Pi F_1$$

$$= \begin{pmatrix} \psi_0^2(1) \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} + \frac{\sigma_0^2}{\sigma_{\nu-2}^2} & 0 & \dots & 0 \\ 0 & \psi_1^2(1) \frac{\sigma_0^2}{\sigma_{\nu-1}^2} + \frac{\sigma_1^2}{\sigma_{\nu-1}^2} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Corollary 3.10 *Regarding Theorem 3.9, in particular, we have that*

$$N_y^{1/2} \left(\hat{\phi}_i - \phi_i \right) \Rightarrow \mathcal{N} \left(0, w_{\phi_i}^2 \right) \quad (3.52)$$

where

$$w_{\phi_i}^2 = \psi_{i-1}^{-4}(1) \left\{ \psi_i^2(2) \sigma_{i-2}^{-2} \sigma_{i-1}^2 \left(1 - \frac{2\psi_i(1)\psi_{i-1}(1)}{\psi_i(2)} \right) + \psi_{i-1}^2(1) \sigma_{i-2}^{-2} \sum_{n=0}^1 \sigma_{i-n}^2 \psi_i^2(n) \right\} \quad (3.53)$$

PROOF. Using the formulas from the proof of Theorem 3.9, we obtain

$$H_1 V_{11} H_1^T = \begin{pmatrix} \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} \frac{\psi_0^2(2)}{\psi_{\nu-1}^4(1)} & 0 & \dots & 0 \\ 0 & \frac{\sigma_0^2}{\sigma_{\nu-1}^2} \frac{\psi_1^2(2)}{\psi_0^4(1)} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad i^{th} \text{ diagonal element is } \frac{\sigma_{i-1}^2}{\sigma_{i-2}^2} \frac{\psi_i^2(2)}{\psi_{i-1}^4(1)}$$

$$H_1 V_{12} H_2^T = \begin{pmatrix} \frac{-\psi_0(2)\psi_0(1)}{\psi_{\nu-1}^3(1)} \frac{\sigma_{\nu-1}^2}{\sigma_{\nu-2}^2} & 0 & \dots & 0 \\ 0 & \frac{-\psi_1(2)\psi_1(1)}{\psi_0^3(1)} \frac{\sigma_0^2}{\sigma_{\nu-1}^2} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

\rightarrow i^{th} diagonal element is $\frac{-\psi_i(2)\psi_i(1)}{\psi_{i-1}^3(1)} \frac{\sigma_{i-1}^2}{\sigma_{i-2}^2}$.

$$H_2 V_{21} H_1^T = H_2 V_{12}^T H_1^T = (H_1 V_{12} H_2^T)^T$$

so i^{th} diagonal element of $H_2 V_{21} H_1^T$ is the same that of $H_1 V_{12} H_2^T$, so i^{th} diagonal element of

$$H_1 V_{12} H_2^T + H_2 V_{21} H_1^T \text{ is } \frac{-2\psi_i(2)\psi_i(1)\sigma_{i-1}^2}{\psi_{i-1}^3(1)\sigma_{i-2}^2}.$$

$$H_2 V_{22} H_2^T = \begin{pmatrix} \frac{\psi_0^2(1)\sigma_{\nu-1}^2}{\psi_{\nu-1}^2(1)\sigma_{\nu-2}^2} + \frac{\sigma_0^2}{\psi_{\nu-1}^2(1)\sigma_{\nu-2}^2} & 0 & \dots & 0 \\ 0 & \frac{\psi_1^2(1)\sigma_0^2}{\psi_0^2(1)\sigma_{\nu-1}^2} + \frac{\sigma_1^2}{\psi_0^2(1)\sigma_{\nu-1}^2} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

\rightarrow i^{th} diagonal element is $\frac{\psi_i^2(1)\sigma_{i-1}^2}{\psi_{i-1}^2(1)\sigma_{i-2}^2} + \frac{\sigma_i^2}{\psi_{i-1}^2(1)\sigma_{i-2}^2}$.

and hence the i^{th} diagonal element of Q is

$$\begin{aligned} &= \frac{\sigma_{i-1}^2 \psi_i^2(2)}{\sigma_{i-2}^2 \psi_{i-1}^4(1)} - \frac{2\psi_i(2)\psi_i(1)}{\psi_{i-1}^3(1)} \frac{\sigma_{i-1}^2}{\sigma_{i-2}^2} + \frac{\psi_i^2(1)\sigma_{i-1}^2}{\psi_{i-1}^2(1)\sigma_{i-2}^2} + \frac{\sigma_i^2}{\psi_{i-1}^2(1)\sigma_{i-2}^2} \\ &= \psi_{i-1}^{-4}(1) \left\{ \frac{\sigma_{i-1}^2 \psi_i^2(2)}{\sigma_{i-2}^2} - \frac{2\psi_i(2)\psi_i(1)\psi_{i-1}(1)\sigma_{i-1}^2}{\sigma_{i-2}^2} + \frac{\psi_i^2(1)\psi_{i-1}^2(1)\sigma_{i-1}^2}{\sigma_{i-2}^2} + \frac{\psi_{i-1}^2(1)\sigma_i^2}{\sigma_{i-2}^2} \right\} \end{aligned}$$

Corollary 3.10 is useful for finding approximate large-sample confidence interval for ϕ_i .

The $(1 - \alpha)100\%$ confidence interval for ϕ_i is $(\hat{\phi}_i - z_{\alpha/2} N_y^{-1/2} w_{\phi_i}, \hat{\phi}_i + z_{\alpha/2} N_y^{-1/2} w_{\phi_i})$, where $Z \sim \mathcal{N}(0, 1)$. Note that non-overlapping confidence intervals give statistically significant conclusions.

Theorem 3.11 *Under the assumption of Theorem 3.7, we have*

$$N_y^{1/2} (\hat{\theta} - \theta) \Rightarrow \mathcal{N}(0, S) \quad (3.54)$$

where

$$\begin{aligned}\hat{\theta} &= [\hat{\theta}_0, \hat{\theta}_1, \dots, \hat{\theta}_{\nu-1}]^T \\ \theta &= [\theta_0, \theta_1, \dots, \theta_{\nu-1}]^T\end{aligned}$$

and the $\nu \times \nu$ matrix S is defined by

$$S = \sum_{k,\ell=1}^2 M_\ell V_{\ell k} M_k^T \quad (3.55)$$

where $V_{\ell k}$ is given in (3.51) and M is a $\nu \times 2\nu$ matrix of partial derivatives

$$M = (M_1, M_2) = \left(\frac{\partial \theta_\ell}{\partial \psi_m(1)}, \frac{\partial \theta_\ell}{\partial \psi_m(2)} \right)_{\ell, m = 0, 1, \dots, \nu-1} \quad (3.56)$$

M can be expressed using $\nu \times \nu$ permutation matrix (22) as

$$M = (-I - F_2 \Pi^{-1} F_1^{-2}, \Pi^{-1} F_1^{-1} \Pi). \quad (3.57)$$

PROOF. Using proposition 3.8, the relation $\theta_t = \frac{\psi_t(2)}{\psi_{t-1}(1)} - \psi_t(1)$, and Theorem 3.7 for $j = 1$ and $h = 2$, we have that (3.55) holds with

$$S = MVM^T$$

where

$$\begin{aligned}\mathbf{g}(\mu) &= \theta \quad \text{where} \quad \mu = \begin{pmatrix} \psi(1) \\ \psi(2) \end{pmatrix} \\ g \begin{pmatrix} \psi_0(1) \\ \vdots \\ \psi_{\nu-1}(1) \\ \psi_0(2) \\ \vdots \\ \psi_{\nu-1}(2) \end{pmatrix} &= \begin{pmatrix} \theta_0 \\ \vdots \\ \theta_{\nu-1} \end{pmatrix} = \begin{pmatrix} \frac{\psi_0(2)}{\psi_{\nu-1}(1)} - \psi_0(1) \\ \vdots \\ \frac{\psi_{\nu-1}(2)}{\psi_{\nu-2}(1)} - \psi_{\nu-1}(1) \end{pmatrix} = \begin{pmatrix} g_0 \\ \vdots \\ g_{\nu-1} \end{pmatrix} \\ \text{and } \mathbf{g}(\mathbf{X}_n) &= \hat{\theta} \quad \text{where} \quad \mathbf{X}_n = \begin{pmatrix} \hat{\psi}(1) \\ \hat{\psi}(2) \end{pmatrix}\end{aligned}$$

Following the same procedure as in the proof of Theorem 3.9, we obtain

$$\begin{aligned} M &= \left(\left[\frac{\partial \theta_\ell}{\partial \psi_m(1)} \right]_{\ell,m=0,\dots,\nu-1} \quad \left[\frac{\partial \theta_\ell}{\partial \psi_m(2)} \right]_{\ell,m=0,\dots,\nu-1} \right) \\ &= (M_1 \quad M_2) \quad \text{same as (3.56)} \end{aligned}$$

Therefore,

$$\begin{aligned} S &= MVM^T \\ &= (M_1 \quad M_2) \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} M_1^T \\ M_2^T \end{pmatrix} \\ &= (M_1V_{11} + M_2V_{21} \quad M_1V_{12} + M_2V_{22}) \begin{pmatrix} M_1^T \\ M_2^T \end{pmatrix} \\ &= M_1V_{11}M_1^T + M_2V_{21}M_1^T + M_1V_{12}M_2^T + M_2V_{22}M_2^T \\ &= \sum_{k,\ell=1}^2 M_\ell V_{\ell k} M_k^T \quad \text{same as (3.55)} \end{aligned}$$

where $V_{\ell k}$ is from (3.31).

A comparison to the proof of Theorem 3.9 shows that $M_1 = -I + H_1$ and $M_2 = H_2$ and then (3.57) follows.

Corollary 3.12 *Regarding Theorem 3.11, in particular, we have that*

$$N_y^{1/2} \left(\hat{\theta}_i - \theta_i \right) \Rightarrow \mathcal{N}(0, w_{\theta_i}^2) \quad (3.58)$$

where

$$w_{\theta_i}^2 = \psi_{i-1}^{-4}(1) \left\{ \psi_i^2(2) \sigma_{i-2}^{-2} \sigma_{i-1}^2 \left(1 - \frac{2\psi_i(1)\psi_{i-1}(1)}{\psi_i(2)} \right) + \sum_{j=1}^2 \psi_{i-1}^{4/j}(1) \sigma_{i-j}^{-2} \sum_{n=0}^{j-1} \sigma_{i-n}^2 \psi_i^2(n) \right\}$$

PROOF. A comparison to the proof of Theorem 3.9 shows that $S = Q + \bar{S}$ where

$$\bar{S} = V_{11} - H_1V_1 - V_{11}H_1^T - H_2V_{21} - V_{12}H_2^T$$

and the i^{th} diagonal element of \bar{S} is therefore $\frac{\sigma_i^2}{\sigma_{i-1}^2} = \psi_{i-1}^{-4}(1) \left\{ \frac{\psi_{i-1}^4(1)\sigma_i^2}{\sigma_{i-1}^2} \right\}$

and hence the i^{th} diagonal element of S is

$$= \psi_{i-1}^{-4}(1) \left\{ \frac{\sigma_{i-1}^2 \psi_i^2(2)}{\sigma_{i-2}^2} - \frac{2\psi_i(2)\psi_i(1)\psi_{i-1}(1)\sigma_{i-1}^2}{\sigma_{i-2}^2} + \frac{\psi_i^2(1)\psi_{i-1}^2(1)\sigma_{i-1}^2}{\sigma_{i-2}^2} + \frac{\psi_{i-1}^2(1)\sigma_i^2}{\sigma_{i-2}^2} + \frac{\psi_{i-1}^4(1)\sigma_i^2}{\sigma_{i-1}^2} \right\}.$$

Corollary 3.12 is useful for finding approximate large-sample confidence interval for θ . An α -level confidence interval for θ is $(\hat{\theta}_i - z_{\alpha/2} N_y^{-1/2} w_{\theta_i}, \hat{\theta}_i + z_{\alpha/2} N_y^{-1/2} w_{\theta_i})$, where $Z \sim \mathcal{N}(0, 1)$.

Note that non-overlapping confidence intervals give statistically significant conclusions.

3.3 Simulation Study

A detailed simulation study was conducted to investigate the practical utility of the innovations algorithm for model identification in the presence of seasonally correlated data. Data for several different $\text{PARMA}_\nu(p, q)$ models with finite fourth moment were simulated. For each model, individual realizations of $N_y = 50, 100, 300,$ and 500 years of data were simulated and the innovations algorithm was used to obtain parameter estimates for each realization. In each case, estimates were obtained for $k = 10, k = 15$ and $k = 20$ iterations of the innovations algorithm in order to examine the convergence, and p -values were computed using (3.27) to identify those estimates that were statistically significant ($p < 0.05$). A FORTRAN program was used to simulate the PARMA samples as well as to make all the necessary calculations. Some general conclusions can be drawn from this study. We found that 10 to 20 iterations of the innovations algorithm are usually sufficient to obtain reasonable estimates of the model parameters. We also found that $N_y = 50$ years of monthly or quarterly data give only rough estimates of the model parameters, while $N_y = 100$ years generally is enough to give good estimates. For the data between 50 and 100 years, the estimates are less accurate but generally adequate for practical applications.

In order to illustrate the general quality of those results, we present the results from two

of the models. The first model is a Gaussian PARMA₄(0, 1) model:

$$X_{k\nu+i} = \varepsilon_{k\nu+i} + \theta_i \varepsilon_{k\nu+i-1} \quad (3.59)$$

where $\{\delta_{k\nu+i} = \sigma_i^{-1} \varepsilon_{k\nu+i}\}$ is an iid sequence of normal random variables with mean zero and standard deviation one. The periodic notation $X_{k\nu+i}$ refers to the (mean zero) simulated data for season i of year k . From the above model, a single realization with $N_y = 500$ years of quarterly data (sample size of $N = N_y \nu = 500 \cdot 4 = 2000$) was generated. Table 3.1 shows the results after $k = 15$ iterations of the innovations algorithm.

Table 3.1: Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 500$ years of simulated PARMA₄(0, 1) data.

Lag ℓ	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.265	.00	0.640	.00	0.918	.00	0.348	.00
2	-0.104	.24	-0.026	.70	-0.058	.50	-0.065	.03
3	-0.015	.52	0.132	.46	-0.066	.29	-0.090	.15
4	0.003	.95	0.063	.18	0.149	.36	-0.071	.12
5	0.010	.76	0.049	.61	0.056	.19	0.118	.31
6	-0.015	.87	-0.113	.10	0.037	.67	-0.003	.92
7	0.013	.57	-0.202	.26	-0.115	.07	-0.091	.15
8	-0.009	.84	0.042	.37	-0.131	.42	-0.054	.23
9	-0.033	.34	0.071	.47	0.053	.21	0.304	.01
:	:	:	:	:	:	:	:	:

The innovation estimates in Table 3.1 cut-off at lag 1 is significantly different from zero, implying that a PARMA₄(0, 1) model is adequate to fit the simulated data, which agrees

with the model used to simulate the data. The parameter estimates for this model, obtained from the relationship (3.39), are summarized in Table 3.2.

Table 3.2: Model parameters and estimates for simulated PARMA₄(0, 1) data.

Parameter	Season 0	Season 1	Season 2	Season 3
θ	0.25	0.65	0.90	0.35
$\hat{\theta}$	0.265	0.640	0.918	0.348
σ	0.90	1.90	0.50	1.20
$\hat{\sigma}$	0.854	1.753	0.458	1.182

The model selected above needs to be checked in order to verify whether it complies with certain assumptions about the model, and to verify how well it represents the generated time series. The model assumptions to be checked are usually the independence and normality of residuals of the model. For PARMA₄(0, 1), the standardized residuals can be calculated using the following relationship:

$$\hat{\delta}_t = \frac{X_t + \sum_{j=1}^{\infty} (-1)^j \hat{\theta}_t \hat{\theta}_{t-1} \dots \hat{\theta}_{t-j+1} X_{t-j}}{\sqrt{\hat{\sigma}_t^2}} \quad (3.60)$$

Inspection of the time series plot of the residuals (see Figure 3.1) indicates no obvious pattern or trend. A normal probability plot of the residuals (see Figure 3.2) also supports the approximate normality of residuals.

The Anderson-Darling (AD) test rejects H_0 (the data follows a normal distribution) if the test statistic is greater than the critical value (or $p < 0.05$). Since the critical value 0.752 is greater than the test statistic 0.627 (and the p -value 0.102 is larger than the critical value 0.05), we conclude that there is no reason to reject the hypothesis of normality.

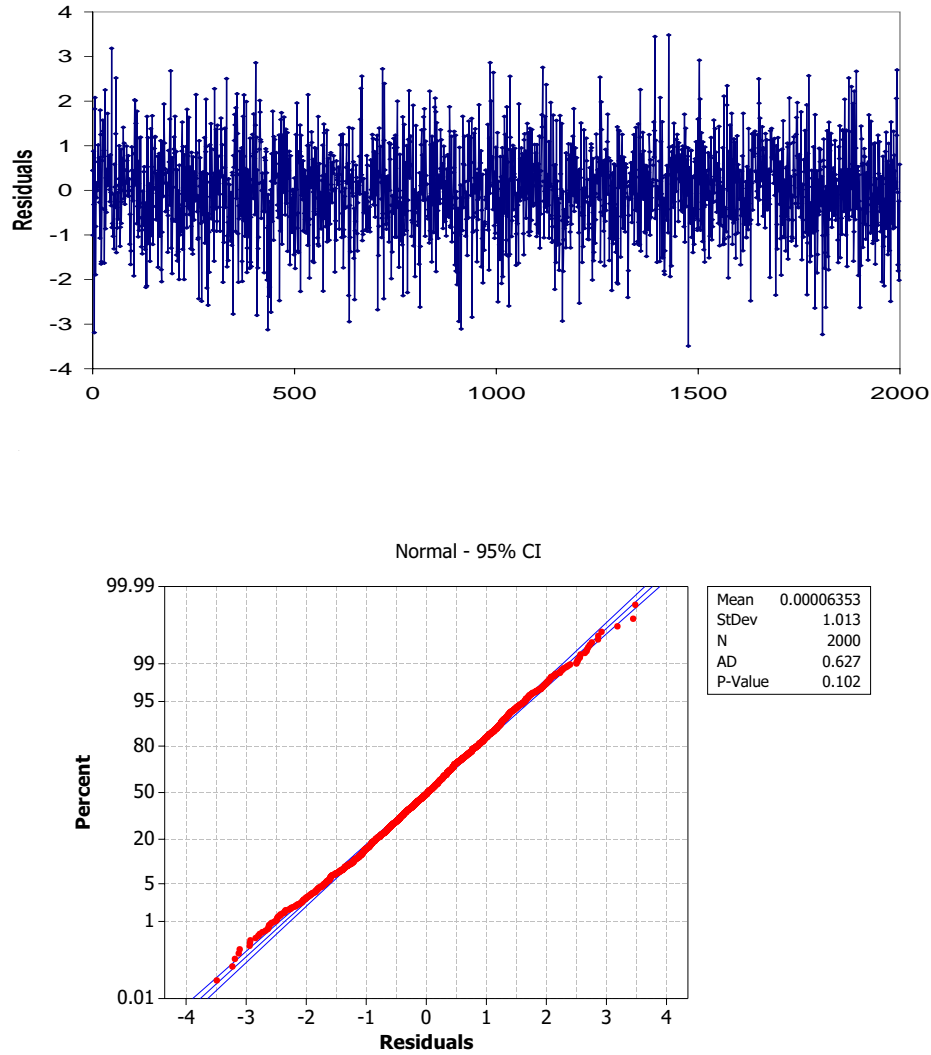


Figure 3.2: Probability plot of the residuals

The plots of the sample autocorrelation functions (ACF)¹⁰ and partial autocorrelation functions (PACF)¹¹ for the residuals are shown in Figure 3.3. All values of ACF and PACF fall inside or close to the 95% confidence bounds ($\pm 1.96/\sqrt{N}$) indicating iid sequence.

In summary, these plots show that the residuals look like iid normal sequence with mean zero and variance close to one, which agrees with the assumption made at the beginning of this section (see also Figure C-1, Appendix C for $N_y = 50, 100, 300$ years).

¹⁰See Section A-1, Appendix A.

¹¹See Section A-1, Appendix A.

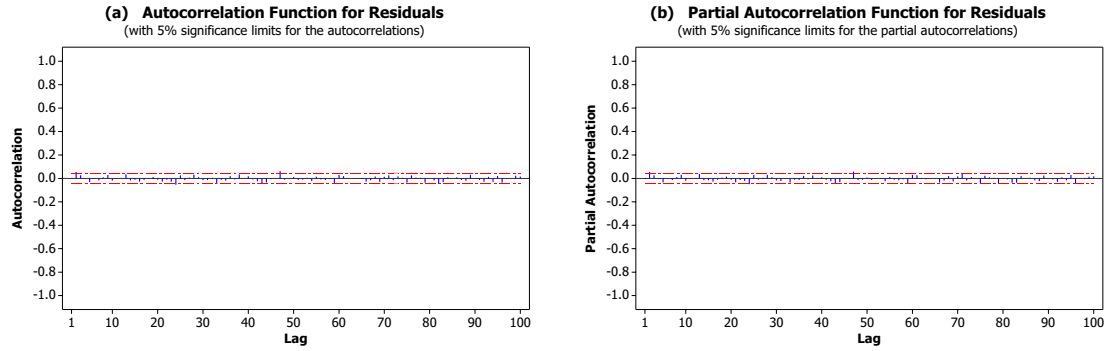


Figure 3.3: The sample ACF and PACF for the residuals, showing the 95% confidence bounds $\pm 1.96/\sqrt{N}$, indicating iid sequence.

The second model for which simulation result will be presented is a $\text{PARMA}_4(1, 1)$ model

$$X_{k\nu+i} = \phi_i X_{k\nu+i-1} + \varepsilon_{k\nu+i} + \theta_i \varepsilon_{k\nu+i-1} \quad (3.61)$$

where $\{\delta_{k\nu+i} = \sigma_i^{-1} \varepsilon_{k\nu+i}\}$ is an iid sequence of normal random variables with mean zero and standard deviation one. The periodic notation $X_{k\nu+i}$ refers to the (mean zero) simulated data for season i of year k . From the above model, a single realization with $N_y = 500$ years of quarterly data (sample size of $N = N_y \nu = 500 \cdot 4 = 2000$) was generated.

Table 3.3 shows the results after $k = 15$ iterations of the innovations algorithm. For season 0 the first four lags are statistically significant, for season 1 and 3 the first three lags are significant, while for season 2 only the first two are significant. Since parameter estimates do not generally cut-off to (statistically) zero at a certain lag, it is advantageous to seek a parsimonious mixed moving average/autoregressive model. The $\text{PARMA}_\nu(1, 1)$ model is the simplest mixed model, and thus is preferred so long as diagnostic plots of the residual autocorrelation (ACF) and/or partial autocorrelation (PACF) indicate no significant serial dependence. The parameter estimates for this model, obtained using equations (3.44) and

Table 3.3: Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 500$ years of simulated $\text{PARMA}_4(1, 1)$ data.

Lag ℓ	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.606	.00	1.231	.00	1.710	.00	0.617	.00
2	-0.620	.00	-0.360	.00	1.047	.00	0.387	.00
3	-0.329	.00	-0.250	.20	-0.346	.00	0.329	.00
4	-0.322	.00	-0.037	.48	-0.120	.69	-0.009	.87
5	0.041	.40	-0.055	.62	0.058	.47	0.087	.52
6	-0.004	.98	0.136	.08	0.128	.46	0.001	.98
7	0.017	.61	0.148	.45	0.187	.11	0.086	.27
8	-0.024	.74	0.042	.42	0.199	.50	-0.003	.95
9	-0.017	.72	-0.006	.96	0.049	.55	0.034	.80
:	:	:	:	:	:	:	:	:

(3.45), are summarized in Table 3.4. It must be noted that θ_t in (3.44) must be replaced with $-\theta_t$ so as to be consistent with the representation in (2.1).

Residuals for this $\text{PARMA}_4(1, 1)$ model can be computed using the equation

$$\hat{\delta}_t = \hat{\sigma}_t^{-1} \left\{ X_t - \left(\hat{\phi}_t + \hat{\theta}_t \right) X_{t-1} + \sum_{j=2}^{\infty} (-1)^j \left(\hat{\phi}_{t-j+1} + \hat{\theta}_{t-j+1} \right) \hat{\theta}_t \hat{\theta}_{t-1} \dots \hat{\theta}_{t-j+2} X_{t-j} \right\} \quad (3.62)$$

which was obtained by solving (2.1) for the innovations and substituting the estimated model parameters for their true values. The analysis of residuals also indicates that they are iid normal sequence with mean zero and variance close to one, which agrees with the assumption in (3.61). For more complex mixed PARMA models, Akaike's information criterion (AIC) could also be a useful guide for the selection process (see Vecchia, 1985b; Brockwell and

Table 3.4: Model parameters and estimates (with 95% confidence interval) for simulated $\text{PARMA}_4(1, 1)$ data.

Parameter	Season 0	Season 1	Season 2	Season 3
θ	0.25	0.65	0.90	0.35
$\hat{\theta}$	0.400 ± 0.265	0.636 ± 0.264	0.859 ± 0.127	0.391 ± 0.204
ϕ	-0.90	0.50	0.80	0.25
$\hat{\phi}$	-1.005 ± 0.259	0.595 ± 0.201	0.850 ± 0.125	0.226 ± 0.032
σ	0.90	1.90	0.50	1.20
$\hat{\sigma}$	0.832	1.771	0.482	1.215

Davis, 1991).

3.4 Application to Modeling of Natural River Flows

Next we model a monthly river flow time series from the Fraser River at Hope, British Columbia. The Fraser River is the longest river in BC, travelling almost 1400 km and sustained by a drainage area covering 220,000 km². It rises in the Rocky Mts., at Yellowhead Pass, near the British Columbia-Alta. line and flows northwest through the Rocky Mt. Trench to Prince George, thence south and west to the Strait of Georgia at Vancouver. Its main tributaries are the Nechako, Quesnel, Chilcotin, and Thompson rivers.

The data are obtained from daily discharge measurements, in cubic meter per second, averaged over each of the respective months to obtain the monthly series. The series contains 72 years of data from October 1912 to September 1984. In the following analysis, $\nu = 0$ corresponds to October and $\nu = 11$ corresponds to September. A partial plot of the original data, given Figure 3.4, shows the cyclic behavior of the monthly flows. The sample

mean, standard deviation and autocorrelations at lag 1 and lag 2 are given in Table 3.5 (see also Figure 3.11). The nonstationarity of the series is apparent since the mean, standard deviation and correlation functions vary significantly from month to month. For example, the confidence intervals for the means are non-overlapping, indicating statically significant difference. Removing the periodicity in mean and variance will not yield a stationary series. Therefore, a periodically stationary time series model is appropriate. After $k = 20$ iterations, the innovations algorithm yields the innovations estimates and associated p -values found in Table 3.6.

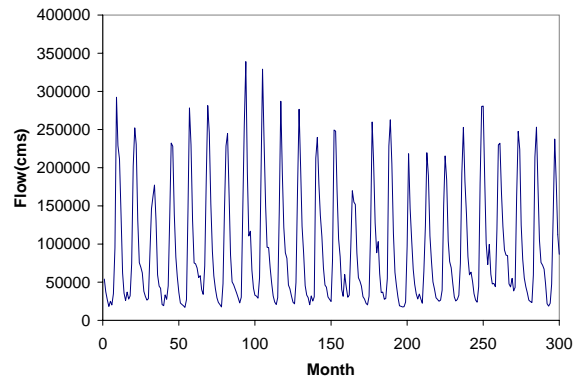


Figure 3.4: Average monthly flows (cms) for the Fraser River at Hope, BC, indicate a seasonal pattern.

Since the $\hat{\psi}_i$ weights do not generally cut-off to (statistically) zero at a certain lag, we choose a parsimonious mixed model that captures the periodic behavior as well as the exponential decay evidenced in the autocorrelation function. We find that a $\text{PARMA}_{12}(1,1)$ model

$$X_{k\nu+i} - \phi_i X_{k\nu+i-1} = \varepsilon_{k\nu+i} + \theta_i \varepsilon_{k\nu+i-1} \quad (3.63)$$

is sufficient in adequately capturing the series autocorrelation structure. The Akaike's information criterion (AIC) or physical basis of the river flow process could also be helpful

Table 3.5: Sample mean, standard deviation and autocorrelation at lag 1 and 2 of average monthly flow series for the Fraser River at Hope, BC.

month	Parameter			
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}(1)$	$\hat{\rho}(2)$
OCT	69763 \pm 11542	19997	0.688	0.517
NOV	56000 \pm 11497	17698	0.731	0.581
DEC	40352 \pm 7987	12817	0.715	0.531
JAN	33135 \pm 7420	9252	0.787	0.691
FEB	30861 \pm 7271	8845	0.779	0.385
MAR	29709 \pm 6492	8834	0.510	0.224
APR	59293 \pm 5766	20268	0.302	-0.294
MAY	171907 \pm 15616	40200	0.272	-0.047
JUN	248728 \pm 22226	45120	0.568	0.496
JUL	199118 \pm 20410	42543	0.779	0.462
AUG	127157 \pm 14149	28070	0.718	0.320
SEP	86552 \pm 11354	20052	0.635	0.454

in choosing the appropriate model (Vecchia, 1985b; Salas *et al.*, 1981,1992; Brockwell and Davis, 1991). The parameter estimates for this model, obtained using equations (3.44) and (3.45), are summarized in Table 3.7. The confidence intervals are narrower and non-overlapping, indicating strong statistical significance, which can also be verified by small p -values (not shown). Model residuals were estimated using equation (3.62). It must be noted that θ_t in (3.44) must be replaced with $-\theta_t$ so as to be consistent with the representation in (2.1). Although the model is periodically stationary, the standard residuals should be

Table 3.6: Moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, 2, \dots, 6$, and p -values, after $k = 20$ iterations of the innovations algorithm applied to average monthly flow series for the Fraser River at Hope, BC.

i	$\hat{\psi}_i(1)$	p	$\hat{\psi}_i(2)$	p	$\hat{\psi}_i(3)$	p	$\hat{\psi}_i(4)$	p	$\hat{\psi}_i(5)$	p	$\hat{\psi}_i(6)$	p	...
0	0.885	.00	0.134	.28	0.105	.10	0.163	.01	0.006	.93	0.038	.78	...
1	0.625	.00	0.625	.00	0.085	.46	0.140	.02	0.077	.17	-0.004	.94	...
2	0.508	.00	0.350	.00	0.419	.00	0.032	.72	0.097	.03	0.019	.65	...
3	0.515	.00	0.287	.00	0.140	.07	0.239	.00	0.034	.60	0.030	.37	...
4	0.791	.00	0.165	.10	0.295	.00	0.112	.12	0.160	.03	0.045	.43	...
5	0.567	.00	0.757	.00	0.057	.61	0.250	.00	0.062	.40	0.139	.06	...
6	1.076	.01	0.711	.11	0.856	.01	0.415	.13	0.241	.17	0.112	.52	...
7	0.522	.03	0.684	.41	0.988	.28	1.095	.09	0.350	.51	0.198	.56	...
8	0.451	.00	-1.014	.00	-0.062	.66	-0.745	.50	0.128	.87	-0.635	.31	...
9	0.618	.00	-0.041	.77	-0.746	.01	-1.083	.26	-0.047	.97	0.514	.50	...
10	0.448	.00	0.409	.00	0.026	.78	-0.241	.20	-1.125	.08	0.799	.26	...
11	0.677	.00	0.159	.01	0.194	.00	0.050	.46	-0.190	.17	-0.402	.38	...

stationary, so the standard $1.96/\sqrt{N}$ 95% confidence limits still apply. Figure 3.6 shows the ACF and PACF of the model residuals. Although a few values lie slightly outside of the 95% confidence bands, there is no apparent pattern, providing some evidence that the $\text{PARMA}_{12}(1, 1)$ model is adequate.

One reason for carefully modeling the river flow time series is to develop the ability to generate synthetic river flows for further analysis. This requires a realistic distributional model for the residuals that can be used to simulate the innovations. After exploring a num-

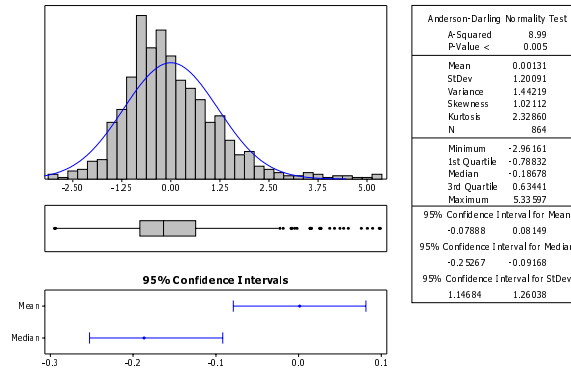
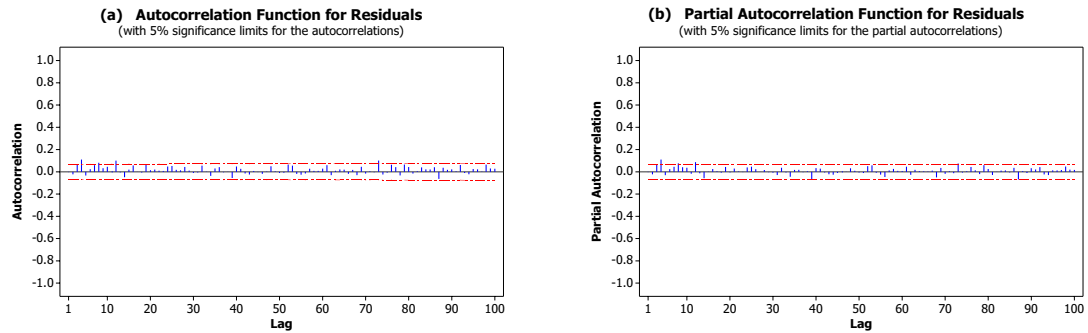


Figure 3.5: Statistical summary of the residuals.

Figure 3.6: ACF and PACF for model residuals, showing the bounds $\pm 1.96/\sqrt{N}$, indicate no serial dependence.

ber of possible distributions, we found that a three-parameter lognormal fits the residuals fairly well. A histogram of the residuals showing the best fitting lognormal density curve (scale = 0.217, location = 1.656 and threshold = -5.363), as well as the corresponding probability plot, are shown in Figure 3.7. On the probability plot, points along the diagonal line (model percentiles equal data percentiles) indicate a good fit. According to this lognormal model, residuals follow the distribution of a random variable $R = -5.363 + e^{(1.656+0.217Z)}$ where $Z \sim \mathcal{N}(0, 1)$.

The histogram in Figure 3.7 (b) shows that the three parameter lognormal gives an

Table 3.7: Parameter estimates (with 95% confidence interval) for PARMA model (3.63) of average monthly flow series for the Fraser River at Hope, BC.

month	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}$
OCT	0.198 ± 0.319	0.687 ± 0.392	11875.479
NOV	0.568 ± 0.251	0.056 ± 0.337	11598.254
DEC	0.560 ± 0.228	-0.052 ± 0.271	7311.452
JAN	0.565 ± 0.233	-0.050 ± 0.299	5940.845
FEB	0.321 ± 0.307	0.470 ± 0.347	4160.214
MAR	0.956 ± 0.240	-0.389 ± 0.351	4610.209
APR	1.254 ± 1.494	-0.178 ± 1.677	15232.867
MAY	0.636 ± 1.451	-0.114 ± 1.526	31114.514
JUN	-1.942 ± 2.362	2.393 ± 2.374	32824.370
JUL	-0.092 ± 0.621	0.710 ± 0.655	29712.190
AUG	0.662 ± 0.191	-0.213 ± 0.226	15511.187
SEP	0.355 ± 0.227	0.322 ± 0.289	12077.991

acceptable overall fit, but the probability plot in Figure 3.7 (a) reveals a lack of fit at both tails. This is important for practical applications, since tail behavior of the residuals (or the innovations) determines the extreme values of the times series, which govern both droughts and floods. None of the standard probability plots we tried (normal, lognormal, Weibull, gamma, etc.) gave an adequate fit at the tails. To check for a power law probability tail we constructed a Mandelbrot plot of each tail (Figure 3.8) as described in Mandelbrot (1963) and Anderson and Meerschaert (1998). Suppose that X_1, \dots, X_n are iid Pareto with distribution function $F(x) = Cx^{-\alpha}$. Then $F(x) = P[X > x] = Cx^{-\alpha}$ and so $\ln F(x) = \ln C - \alpha \ln x$.

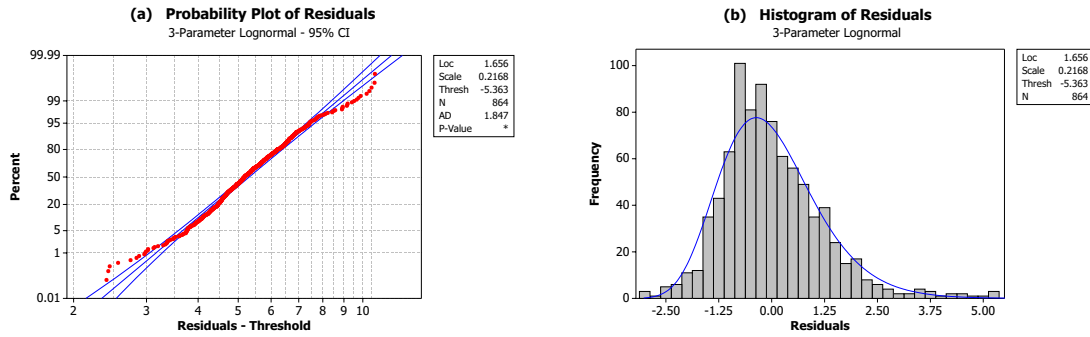


Figure 3.7: Lognormal probability plot (a) and histogram (b) for model residuals, Fraser river at Hope, BC.

Sorting the data in decreasing order so that $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ (order statistics) we should have approximately that $x = X_{(r)}$ when $F(x) = r/n$. Then a plot of $\ln X_{(r)}$ versus $\ln(r/n)$ should be approximately linear with slope $-\alpha$. In Figure 3.8, the downward curve indicates that a simple power law model is not appropriate. However, the shape of the plots is consistent with many examples of truncated Pareto distributions found in the geophysics literature (see for example Aban, Meerschaert and Panorska, 2004; Burroughs and Tebbens, 2001a, 2001b, 2002). This distribution is appropriate when a power law model is affected by an upper bound on the observations. In hydrology it is commonly believed that there is an upper bound on precipitation and therefore river flows (see, for example, Maidment, 1993).

A truncated Pareto random variable X has distribution function

$$F_X(x) = P(X \leq x) = \frac{1 - (\gamma/x)^\alpha}{1 - (\gamma/\beta)^\alpha} \quad (3.64)$$

and density

$$f_X(x) = \frac{\alpha \gamma^\alpha x^{-\alpha-1}}{1 - (\gamma/\beta)^\alpha} \quad (3.65)$$

with $0 < \gamma \leq x \leq \beta < \infty$ and $\gamma < \beta$. Aban, Meerschaert and Panorska (2004) develop maximum likelihood estimators (MLE) for the parameters of the truncated Pareto distribu-

tion. When a truncated Pareto is fit to the tail of the data, the parameters are estimated by obtaining the conditional maximum likelihood estimate based on the largest order statistics, representing only the portion of the tail where the truncated Pareto model holds.

Theorem 3.13 *When $X_{(r)} > X_{(r+1)}$, the conditional maximum likelihood estimator for the parameters of the upper-truncated Pareto in (3.64) based on the $r + 1$ largest order statistics is given by*

$$\hat{\beta} = X_{(1)} \quad (3.66)$$

$$\hat{\gamma} = r^{1/\hat{\alpha}} X_{(r+1)} \left[n - (n - r)(X_{(r+1)}/X_{(1)})^{\hat{\alpha}} \right]^{-1/\hat{\alpha}} \quad (3.67)$$

and $\hat{\alpha}$ solves the equation

$$\frac{r}{\hat{\alpha}} + \frac{r (X_{(r+1)}/X_{(1)})^{\hat{\alpha}} \ln(X_{(r+1)}/X_{(1)})}{1 - (X_{(r+1)}/X_{(1)})^{\hat{\alpha}}} - \sum_{i=1}^r [\ln X_{(i)} - \ln X_{(r+1)}] = 0 \quad (3.68)$$

The truncated Pareto distribution given by (3.64) with $0 < \gamma < \infty$ reduces to the Pareto distribution in the limit $\beta = \infty$. An asymptotic q -level test (see Aban et al., 2004) is used to distinguish whether a given data set is more appropriately modelled by a Pareto or truncated Pareto distribution. The asymptotic level q test ($0 < q < 1$) rejects the null hypothesis $H_0 : \beta = \infty$ (Pareto) in favor of the alternative $H_a : \beta < \infty$ (truncated Pareto) if $X_{(1)} < x_q$ where $q = \text{Exp}(-nC x_q^{-\alpha})$. This test rejects H_0 if and only if

$$X_{(1)} < \left(\frac{nC}{-\ln q} \right)^{1/\alpha} \quad (3.69)$$

and the p -value of this test is given by

$$p = \text{Exp}(-nC X_{(1)}^{-\alpha}) \quad (3.70)$$

Hill's estimator¹² (4.12) gives $\hat{\alpha} = 2.450$ and $\hat{C} = 0.282$ for the upper tail ($r = 39$ residuals), and $\hat{\alpha} = 4.184$ and $\hat{C} = 0.348$ for the lower tail ($r = 34$ residuals). Using (3.70),

¹²See Section 4.4 for a detailed discussion.

the p -values are 0.018 and 0.041, respectively. The small p -values as well as the visual evidence of the probability plot in Figure 3.8 suggests that the Pareto model is not a good fit for the tails of the estimated residuals. Therefore, we decided to fit a truncated Pareto to roughly the upper 5% and the lower 5% of the residuals. Then we computed $-1.697 =$ the 5th percentile and $2.122 =$ the 95th percentile of the three parameter lognormal distribution we fit to the body of the residuals. Next we determined that $r = 39$ residuals exceed the 95th percentile, and $r = 34$ residuals fall below the 5th percentile. Then the MLE from Theorem 3.13 was used to estimate the parameters ($\hat{\beta} = 5.336$, $\hat{\gamma} = 0.072$, $\hat{\alpha} = 0.722$) of the best fitting truncated Pareto distribution, and the theoretical distribution tail $P(R > r)$ was plotted over the 39 largest positive residuals in Figure 3.8(left). In Figure 3.8 (right), we used the same method to fit a truncated Pareto ($\hat{\beta} = 2.961$, $\hat{\gamma} = 0.291$, $\hat{\alpha} = 1.560$) to the 34 largest negative residuals, after a change of sign. Both of the plots in Figures 3.8 indicate an adequate fit.

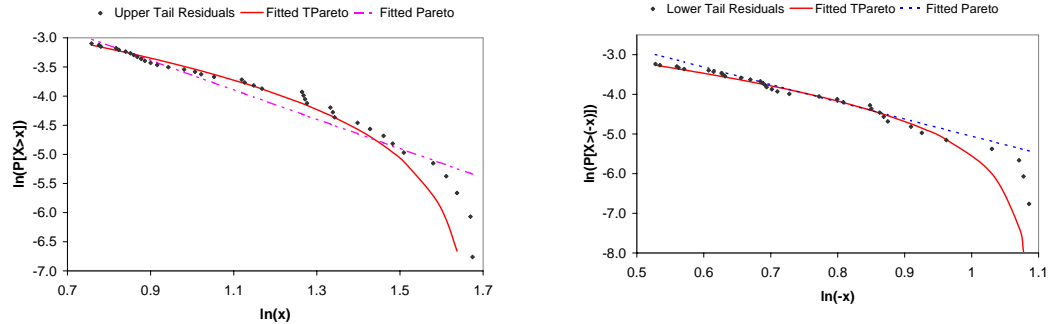


Figure 3.8: Log-log plot of upper(left) and lower(right) residual tails and fitted Pareto and truncated Pareto distributions, Fraser river at Hope, BC.

A mixture distribution with lognormal body and truncated Pareto tails was used to

simulate the innovations. The mixture has cumulative distribution function (cdf)

$$P(\delta \leq r) = \begin{cases} F_-(r) & \text{if } r < -1.697 \\ F_0(r) & \text{if } -1.697 \leq r \leq 2.122 \\ F_+(r) & \text{if } r > 2.122 \end{cases} \quad (3.71)$$

where F_0 is the cdf of the lognormal, and F_+ , F_- are truncated Pareto cdfs of the positive and negative tails, respectively. The truncated Pareto distributions were slightly shifted (by $s = 0.172$ on the positive tail and $s = 0.174$ on the negative tail) to make the mixture cdf continuous. Now innovations could be simulated by the inverse cumulative distribution function method $\delta = F^{-1}(U)$ where U is a pseudorandom number uniformly distributed on the unit interval $(0, 1)$. However, this is impractical in the present case since the lognormal cdf is not analytically invertible. Instead, we used the Box-Müller method¹³ to generate standard normal random variates Z (see Gentle, 2003). Then lognormal random variates were calculated using $\delta = -5.363 + \exp(1.656 + 0.217Z)$. If $R > 2.122$, the 95th percentile of the lognormal, we generated another uniform $(0, 1)$ random variate U and substituted $\delta = F_+^{-1}(0.95 + 0.05U)$. If $R < -1.697$, the 5th percentile of the lognormal, we substituted $\delta = F_-^{-1}(0.05U)$. This gives simulated innovations δ with the mixture distribution (3.71).

Figure 3.9 shows a probability plot for $N = N_y\nu$ simulated innovations (for $\nu = 12$ months and $N_y = 100$ years) from the mixture distribution (3.71). Comparison with Figure 3.7 (a) shows that the simulated innovations are statistically identical to the computed model residuals in terms of distribution. Substituting the simulated innovations into the model (3.63) generates N_y years of simulated river flow. It is advantageous to simulate several extra years of river flows and throw out the initial years (100 years in this case), since we did not simulate X_t for $t < 0$. This ensures that the simulated series is periodically stationary

¹³See Section A-4, Appendix A.

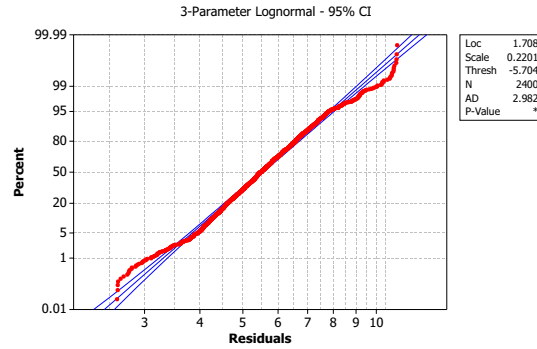


Figure 3.9: Probability plot of simulated innovations using the mixed three parameter lognormal and truncated Pareto distributions. Compare Figure 3.7 (a).

(see Figure 3.10). Figure 3.11 shows the main statistical characteristics (mean, standard deviation and autocorrelations) of a typical synthetic river flow time series obtained by this method, as well as the same statistical measures for the observed time series. It is apparent that this procedure closely reproduces the main statistical characteristics, indicating that the modeling procedure is trustworthy for generating synthetic river flows. Such synthetic river flows are useful for design of hydraulic structures, for optimal operation of reservoir systems, for calculating the risk of failure of water supply systems, etc.

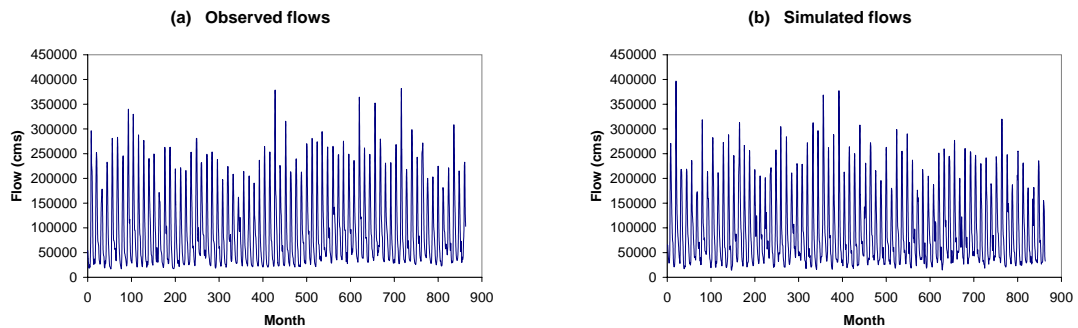


Figure 3.10: Plot of (a) observed and (b) simulated monthly river flows for the Fraser River at Hope, BC, indicating similarity.

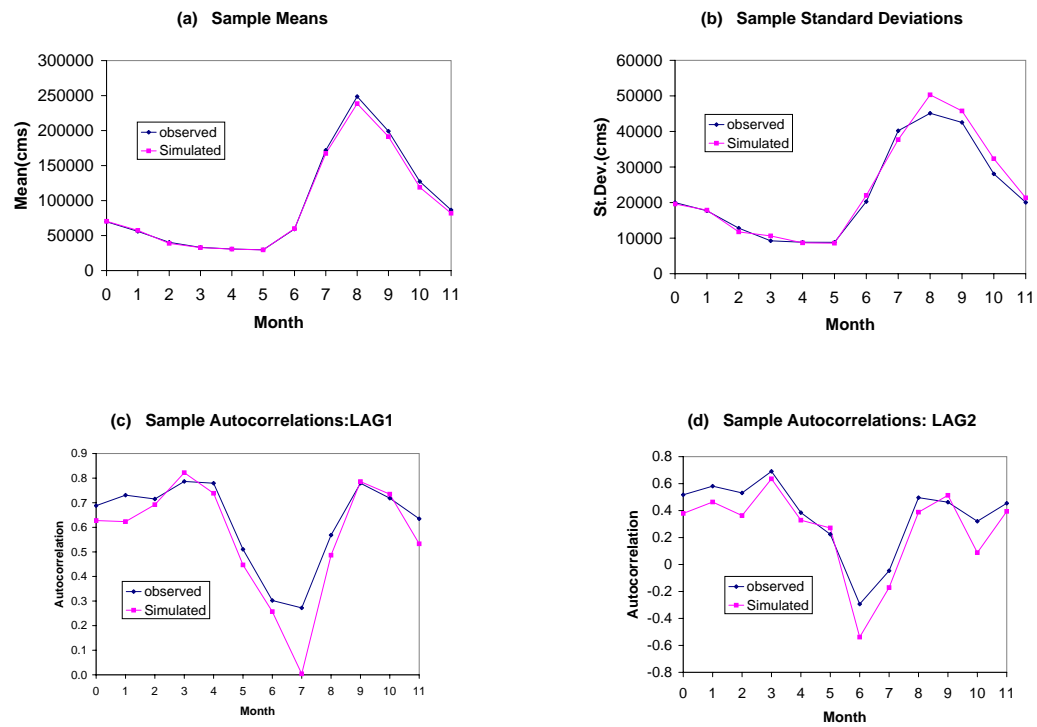


Figure 3.11: Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed monthly river flow data for the Fraser River at Hope, BC.

4 Modeling River Flows with Heavy Tails

4.1 General

Many river flow time series exhibit occasional sharp spikes (heavy tail characteristics). A model that captures these heavy tail characteristics is, therefore, important in adequately describing the series. Typically, a time series with heavy tails is appropriately transformed so that the normal asymptotics apply. The normal distribution is widely applicable because of the central limit theorem, which states that the sum of a large number of independent, identically distributed variables from a finite-variance distribution will tend to be normally distributed. An alternative model (based on stable asymptotics) allows a more faithful representation of the river flow without preliminary transformation. In this chapter, we fit a periodic ARMA model to river flow data without moment assumptions (Anderson and Meerschaert, 1997,1998), and model the innovations process using stable distributions or a mixture of stable with appropriate tail distributions. Although there are other heavy tailed alternatives to a normal law, stable distributions have been proposed herein mainly because of the generalized central limit theorem, which states that stable laws are the only possible limit distributions for properly normalized and centered sums of independent, identically distributed random variables.

Heavy-tailed distributions (also known as power-law distributions) are important in applications to electrical engineering, finance, geology, hydrology, and physics. Books and articles dealing with heavy tails in these fields include Mandelbrot (1963), Fama (1965), Feller (1971), Rybin (1978), Davis and Resnick (1985a,b,1986), Brockwell and Devis (1991), Jansen and de Vries (1991), Bhansali (1993), Janicki and Weron (1994a,b), Samorodnitsky and Taqqu (1994), Nikias and Shao (1995), Mittnik and Rachev (1998), Kokoszka (1996), Kokoszka

and Taqqu (1994,1996), Lotetan and Philips (1994), Mikosch, Gadrach, Klüppenberg and Adler (1995), Resnick and Stărică (1995), McCulloch (1997), Anderson and Meerschaert (1997,1998). Various probability distributions are used in these fields, such as Stable, Type II Extreme Value and the generalized Pareto.

In this chapter we will discuss the application of heavy tails models to river flow data. As an application, we consider the average monthly flow of the Salt River near Roosevelt, Arizona.

4.2 Definition of Heavy Tails

In the literature, different definitions of heavy tailed distributions exist (see, for example, Werner and Upper, 2002). We say that a probability distribution function has heavy tails if the tails of the distribution diminish ($P(|X| > x) \rightarrow 0$) at an algebraic rate (like some power of x) rather than at an exponential rate. In this case some of the moments of this probability distribution will fail to exist. The k th moment of a probability distribution function $F(x)$ with density function $f(x)$ is defined by $\mu_k = \int x^k f(x) dx$. The mean μ_1 and the variance σ^2 are related to the first two moments by familiar equations $\mu = \mu_1$ and the variance $\sigma^2 = \mu_2 - \mu_1^2$. Examples of probability distributions with heavy tails are stable (for instance, Cauchy and Lévy distribution), Student's t , and Pareto distributions.

In the Gaussian case, $X \sim \mathcal{N}(0, 1)$, there is a well known formula to approximate the tail probability: as $x \rightarrow \infty$,

$$P(X > x) \sim \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}} \quad (4.1)$$

(see Feller, 1971).

If X , for example, is standard Cauchy, then

$$F(x) = P[X \leq x] = \frac{1}{2} + \frac{1}{\pi} \arctan(x) \quad \text{and} \quad f(x) = \frac{dF(x)}{dx} = \frac{1}{\pi(1+x^2)} \quad (4.2)$$

Although the bell-shaped of the density of the Cauchy distribution appears to be similar to that of normal law, the tails are heavier, see Table 4.1. While the density of the normal law diminishes at an exponential rate, for the Cauchy we have $f(x) \sim \pi^{-1}x^{-2}$ as $|x| \rightarrow \infty$. This causes the integral defining the k th moment to diverge when $k \geq 1$, and hence the mean and standard deviations of the Cauchy are undefined. Figure 4.1 shows a plot of Normal, Cauchy and Lévy densities.

Table 4.1: Comparison of tail probabilities for standard normal, Cauchy and Lévy distributions.

c	$P(X > c)$		
	Normal	Cauchy	Lévy
0	0.5000	0.5000	1.0000
1	0.1587	0.2500	0.6827
2	0.0228	0.1476	0.5205
3	0.001347	0.1024	0.4363
4	0.00003167	0.0780	0.3829
5	0.0000002866	0.0628	0.3453

4.3 Heavy Tailed Time Series Model

Suppose that $\{\delta_t\}$ are independent, identically distributed random variables with common distribution function $F(x) = P[\delta_t \leq x]$. Regular variation is a technical condition that is required for the generalized central limit theorem to apply.

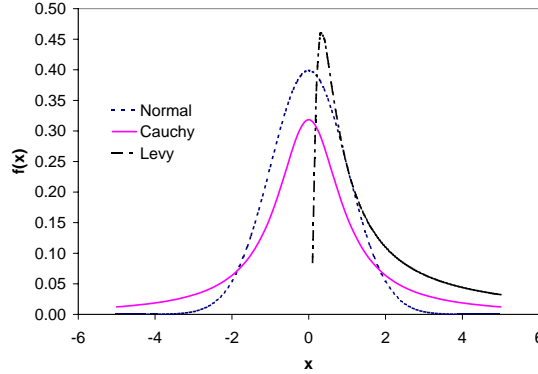


Figure 4.1: Plots of standardized normal, Cauchy and Lévy densities.

Definition 4.1 A Borel measurable function $R : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ said to be regularly varying at infinity with index $\rho \in \mathbb{R}$ if

$$\lim_{x \rightarrow \infty} \frac{R(\lambda x)}{R(x)} = \lambda^\rho \quad \text{for all } \lambda > 0. \quad (4.3)$$

R is slowly varying iff $\rho = 0$ in (4.3).

For example, the following are some regularly varying functions on \mathbb{R}^+ .

- Constant functions are slowly varying, and therefore regularly varying.
- If $f \sim g$ (meaning that $f(x)/g(x) \rightarrow 1$ as $x \rightarrow \infty$) and f is regularly varying, then g is regularly varying with the same index.
- $f(x) = ax^b$ is regularly varying with index b .

We say that the distribution $F(x)$ belongs to the domain of attraction of some nondegenerate random variable Y with distribution $G(x)$ if there exist real constants $a_n > 0$ and b_n such that

$$a_n^{-1}(\delta_1 + \dots + \delta_n - nb_n) \Rightarrow Y \quad (4.4)$$

where \Rightarrow indicates convergence in distribution. The generalized central limit theorem states that (4.4) holds with Y normal if and only if the truncated second moment function

$$\mu(x) = \int_{|u| \leq x} u^2 dF(u) \quad (4.5)$$

is slowly varying; (4.4) holds with Y nonnormal if and only if the tail function $T(x) = P[|\delta_t| > x] = F(-x) + 1 - F(x)$ varies regularly with index $-\alpha$ for some $0 < \alpha < 2$ and the tails satisfy the balancing condition

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{T(x)} = p \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{F(-x)}{T(x)} = q \quad (4.6)$$

for some $0 \leq p, q \leq 1$ with $p + q = 1$; see, for example, Feller (1971). The norming constants can be chosen to satisfy $nP[|\delta_t| > a_n x] \rightarrow C$ and are always of the form $a_n = n^{1/\alpha} \ell_n$, where ℓ_n is slowly varying. The domain of normal attraction, in which we assume that $a_n = n^{1/\alpha}$ is strictly smaller class of distributions. If $\alpha > 1$, then the mean $E\delta_t$ exists, and we can take $b_n = E\delta_t$. If $\alpha < 1$, the mean fails to exist, the norming constant $a_n \rightarrow \infty$ faster than $n \rightarrow \infty$, and we can let $b_n = 0$. The Pareto distribution $F(x) = 1 - Cx^{-\alpha}$ ($x \geq C^{1/\alpha}$) as well as the type II extreme value distribution $F(x) = \exp(-Cx^{-\alpha})$ ($x \geq 0$) belong to the stable domain of attraction when $0 < \alpha < 2$. Sums of iid random variables with these distributions are approximately stable. Stable distributions are a rich class of distributions that include the Gaussian and Cauchy distributions in a family that allow skewness and heavy tails. The class was characterized by Paul Lévy (1925) in his study of normalized sums of iid terms. The general stable distribution is described by four parameters: an index of stability $\alpha \in (0, 2]$, a skewness parameter β , a scale parameter γ and a location parameter η (see Nolan, 1999). We will use γ for the scale parameter and η for the location parameter to avoid confusion with the symbols σ and μ , which are used exclusively for the standard deviation

and mean. The parametrization most often used now (see Samorodnitsky and Taqqu, 1994) is the following : $Y \sim S1(\alpha, \beta, \gamma, \eta)$ if the characteristic function of Y is given by

$$E\{\exp(itY)\} = \begin{cases} \exp\{-\gamma^\alpha |t|^\alpha (1 - i\beta(\text{sign}t)\tan\frac{\pi\alpha}{2}) + i\eta t\} & \alpha \neq 1 \\ \exp\{-\gamma |t| (1 + i\beta\frac{2}{\pi}(\text{sign}t)\ln|t|) + i\eta t\} & \alpha = 1 \end{cases} \quad (4.7)$$

where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma > 0$, and $\eta \in R$. For a more detailed discussion of stable distributions and their properties, the reader is referred to Nolan (1999). If we choose the norming constant a_n in (4.4) to satisfy $nP[|\delta_t| > a_n x] \rightarrow C$, then the stable limit Y will have a scale factor satisfying $\gamma^\alpha = C\Gamma(1 - \alpha)\cos(\pi\alpha/2)$ when $\alpha \neq 1$ and $\gamma = C\pi/2$ when $\alpha = 1$ (Theorem 7.3.5 on p.265 of Meerschaert and Scheffler, 2001). The parameter C is called the dispersion and is sometimes used as an alternative scale parameter for stable laws. The parameters in the balance equations (4.6) are related to the skewness of the limit Y by $p = (1 + \beta)/2$ and $q = (1 - \beta)/2$. If $\alpha > 1$ and $b_n = E\delta_t$ or $\alpha < 1$ and $b_n = 0$, then the location parameter of the limit Y in (4.4) is $\eta = 0$. If $\alpha \leq 1$, then EY does not exist, but if $\alpha > 1$, then $EY = \eta$, and, in particular, we obtain $EY = 0$ when we center to zero expectation. If Y_n are iid stable with the same distribution as Y , then $\sum k_n Y_n$ is also stable with the same stable index and skewness parameters. The location parameter of $\sum k_n Y_n$ is $\gamma(\sum |k_n|^\alpha)^{1/\alpha}$, and so, in particular, the scale parameter of kY is $k\gamma$. The dispersion of $\sum k_n Y_n$ is $C \sum |k_n|^\alpha$.

Suppose that the iid sequence $\{\delta_t\}$ represents the innovations process of a time series (2.1). If δ_t has finite fourth moment, then normal asymptotics apply (see Chapter 3), but if δ_t has an infinite fourth moment, the asymptotics are governed by stable laws. Infinite fourth moments occur when the tail function $T(x) = P[|\delta_t| > x]$ varies regularly with index $-\alpha$ for some $0 < \alpha < 4$. Then $T(x) \rightarrow 0$ about as fast as $x^{-\alpha} \rightarrow 0$ as $x \rightarrow \infty$. See Anderson

and Meerschaert (1998) for a review of literature in heavy tail modeling in hydrology.

Suppose that $\{\delta_t\}$ are iid and that their common distribution $F(x)$ has a regularly varying tails with index $-\alpha$ for some $2 < \alpha < 4$. Then $\sigma_t^2 = E\varepsilon_t^2 < \infty$, and the autocovariance and autocorrelation of the infinite moving average (3.1) at season i and lag ℓ are

$$\begin{aligned}\gamma_i(\ell) &= \sum_j \sigma_t \psi_t(j) \sigma_{t+\ell} \psi_{t+\ell}(j + \ell) \\ \rho_i(\ell) &= \frac{\gamma_i(\ell)}{\sqrt{\gamma_i(0)\gamma_{i+\ell}(0)}} = \frac{\sum_j \psi_t(j) \psi_{t+\ell}(j + \ell)}{\sqrt{\sum_j \psi_t(j)^2 \sum_j \psi_{t+\ell}(j)^2}}\end{aligned}\tag{4.8}$$

Given a sample size of $N = N_y \nu$, the sample statistics are obtained using equations (2.11), (2.12) and (2.13).

We will say that the iid sequence $\{\delta_t\}$ is $RV(\alpha)$ if $P[|\delta_t| > x]$ varies regularly with index $-\alpha$ and $P[\delta_t > x]/P[|\delta_t| > x] \rightarrow p$ for some $p \in [0, 1]$. If δ_t is $RV(\alpha)$, then the equivalent vector process Z_t in (3.2) has iid components with regularly varying tails, and we will also say Z_t is $RV(\alpha)$. If $\alpha > 2$, then Z_t belongs to the domain of attraction of a multivariate normal law whose components are iid univariate normal. Loretan and Philips (1994) find that the price fluctuations of currency exchange rates and stock prices often follow an $RV(\alpha)$ model with $2 < \alpha < 4$. In this case, $E\delta_t^2 < \infty$, but $E\delta_t^4 = \infty$. Since the innovation ε_t has a finite variance, the sample autocorrelations for the stationary moving average model are asymptotically normal; see for example, Brockwell and Davis (1991, Proposition 7.3.8). The following results show that when $2 < \alpha < 4$, the sample autocorrelations of the periodic moving average model are asymptotically stable.

Theorem 4.2 *Suppose that X_t is periodic moving average (3.1) of an $RV(\alpha)$ sequence $\delta_t = \sigma_t^{-1} \varepsilon_t$ with mean zero, variance one, and $2 < \alpha < 4$. Then $\sigma_t^2 = E\varepsilon_t^2 < \infty$, $\hat{\mu}_i, \hat{\gamma}_i(\ell), \hat{\rho}_i(\ell)$*

are consistent estimators of $\mu_i, \gamma_i(\ell), \rho_i(\ell)$, respectively, and

$$N_y^{-1/2}(\hat{\mu} - \mu) \Rightarrow \mathcal{N}(0, C) \quad (4.9)$$

where $C = \left(\sum_j \Psi_j\right) \left(\sum_j \Psi_j\right)'$; for some $a_{N_y} \rightarrow \infty$ we have

$$\begin{aligned} N_y a_{N_y}^{-2} [\hat{\gamma}_i(\ell) - \gamma_i(\ell)] &\Rightarrow C_{i\ell} = \sum_{r=0}^{\nu-1} C_r(i, \ell) S_r \\ N_y a_{N_y}^{-2} [\hat{\rho}_i(\ell) - \rho_i(\ell)] &\Rightarrow D_{i\ell} = \sum_{r=0}^{\nu-1} D_r(i, \ell) S_r \end{aligned} \quad (4.10)$$

jointly in $i = 0, \dots, \nu - 1$ and $\ell = 0, \dots, h$, where

$$D_r(i, \ell) = \frac{\rho_i(\ell)}{\gamma_i(\ell)} C_r(i, \ell) - \frac{\rho_i(\ell)}{2\gamma_i(0)} C_r(i, 0) - \frac{\rho_i(\ell)}{2\gamma_{i+\ell}(0)} C_r(i + \ell, 0) \quad (4.11)$$

and $C_r(i, \ell) = \sum_j \sigma_i \psi_i(j\nu + i - r) \sigma_{i+\ell} \psi_{i+\ell}(j\nu + i + \ell - r)$. We can always choose a_N so that $N_y P[|\delta_t| > a_{N_y}] \rightarrow C$ for some $C > 0$. Then $S_0, S_1, \dots, S_{\nu-1}$ are iid $\alpha/2$ stable with mean zero, skewness 1, and dispersion C , $C_{i\ell}$ is $\alpha/2$ with mean zero, skewness 1, and dispersion $C \sum_r |C_r(i, \ell)|^{\alpha/2}$; and $D_{i\ell}$ is $\alpha/2$ stable with mean zero, skewness 1, and dispersion $C \sum_r |D_r(i, \ell)|^{\alpha/2}$.

PROOF. See Anderson and Meerschaert (1997,1998) for a proof.

4.4 Checking Heavy Tails in Time Series Data

When $\alpha \in (0, 2)$, we have stable distributions with $f(x) \sim Cq\alpha x^{-\alpha-1}$ as $x \rightarrow -\infty$ and $f(x) \sim Cp\alpha x^{-\alpha-1}$ as $x \rightarrow +\infty$, where $C > 0$ and $0 \leq p, q \leq 1$ with $p+q = 1$. Therefore, it is reasonable to model a stable law as having Pareto tails. In this section we consider in detail the problem of detecting heavy tails in time series data and estimating the tail parameter α . We illustrate the problem with a data set representing monthly flows of the Salt River near Roosevelt, Arizona.

For data analysis applications, the actual distribution is often unknown, so robust estimators are needed to fit the tail parameter α of an unknown heavy tail distribution. Many robust estimators of the tail index have been proposed in the past several years. Hill (1975) proposed a robust estimator based on the asymptotics of extreme values. Hill calculated that

$$\hat{H} = r^{-1} \sum_{i=1}^r \ln X_{(i)} - \ln X_{(r+1)} \quad (4.12)$$

is the conditional maximum likelihood estimator $1/\alpha$ conditional on $X_{(r+1)} \geq D$, where $X_{(1)} \geq X_{(2)} \geq \dots$ are the order statistics of a random sample $X_1 \dots X_n$. We can approximate α by $\hat{\alpha}_r = 1/\hat{H}_r$.

The Hill's estimator for C is given by

$$\hat{C} = \frac{r+1}{n} X_{(r+1)}^{\hat{\alpha}_r} \quad (4.13)$$

For data which are approximately Pareto in the tails, one should choose r small enough so that only the Pareto-like tail is represented.

Resnick and Stărică (1995) show that Hill's estimation procedure yields a consistent estimator of the tail index α for stationary moving average models where the innovations have regularly varying probability tails with index $-\alpha$. Their result can be understood by noting that since the largest observations in a heavy tail time series model tend to be widely spaced in time, they resemble iid observations. Hill's estimator can be applied to a wide variety of distributions, such as the stable and type II extreme value distributions, whose tails are approximately Pareto. Thus, for example, applying Hill estimator to the largest $r = 15$ order statistics of the Salt River data, we obtain $\hat{\alpha} = 2.925$ and $C = 1.40 \times 10^9$ (see Table 1.1). This indicate that the Salt River data have heavy tails with infinite fourth moment and finite

variance. Figure 4.2a shows the plot of $\hat{\alpha}_r$ versus r . Hall (1982) shows that Hill's estimator is consistent and asymptotically normal with variance approximately $\sigma^2 = \alpha^2/r$. For $r = 15$ and $\hat{\alpha} = 2.925$, we obtain $\sigma = 0.755$, and the z test of $H_0 : \alpha = 4$ versus $H_0 : \alpha < 4$ or $H_0 : \alpha = 2$ versus $H_a : \alpha > 2$ has p -value 0.08 and 0.11, respectively, indicating that the null hypothesis is rejected. Increasing r yields a smaller standard deviation, but Hall's theorem assumes that $r/n \rightarrow 0$ as $n \rightarrow \infty$, so that one should base Hill's estimator on a vanishingly small percentage of the data.

The simplest and most straightforward method of estimating the tail index is to plot the

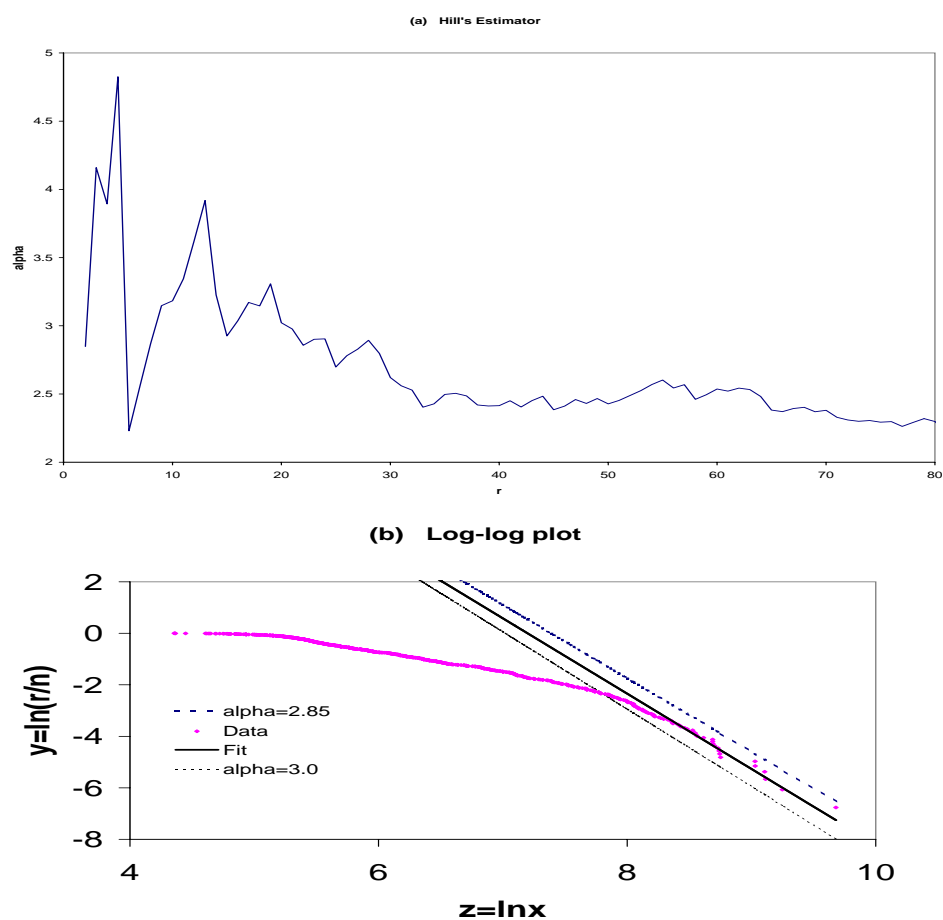


Figure 4.2: (a) Hill's estimator, $\hat{\alpha}$ versus r , and (b) log-log plot for the upper tail of the monthly flows for the Salt River near Roosevelt, AZ.

right tail of the empirical cdf on a log-log scale (see, Figure 3.8, Section 3.4).

Figure 4.2 (b) uses a log-log plot to illustrate the heavy tail distribution of the Salt river data. It shows the Salt River data along with the line that fits a Pareto model with parameters $\alpha = 2.92$ and $C = 1.40 \times 10^9$ (from Hill's estimator for $r=20$). The graph indicates that the true α lies between 2.85 and 3.0. This method is very sensitive to the size of the sample and the choice of the number of observations used in the regression.

Other tail estimators have also been proposed, see for example Dekkers *et al.* (1989), Kratz and Resnick (1996) and Wei (1995), but Hill's estimator continues to be the most popular in practice. McCulloch (1997), Resnick (1997) and Fofack and Nolan (1999) observed that Hill's estimator performs poorly for stable data with $1.5 < \alpha < 2$. Though Hill's estimator is scale invariant, it is not shift invariant. Aban and Meerschaert (2001) modified Hill's estimator to make it shift invariant.

Meerschaert and Scheffler (1998) introduced another robust estimator (hereafter the MS estimator) based on the asymptotics of the sample variance.

Suppose that $\{X_i\}$ is a sequence of n iid random variables that are in the domain of attraction of some stable random variable with $\alpha < 2$. The MS estimator, given in Meerschaert and Scheffler (1998) is

$$\hat{\beta}_n = \frac{\gamma + \ln_+ \sum_{i=1}^n (X_i - \bar{X}_n)^2}{2(\gamma + \ln n)} \quad (4.14)$$

where $\ln_+(x) = \max(\ln(x), 0)$, \bar{X} is the sample mean and $\gamma \doteq 0.5772$. $\hat{\beta}_n$ is the MS estimator for $1/\alpha$. This quadratic estimator performs well in many cases where Hill's estimator is ineffective, such as stable distributions with tail index $1.5 < \alpha < 2$. The quadratic estimator is shift invariant, but not scale invariant. In practical applications, one has to manually scale the data before applying this estimator in order to get useful results. Bianchi and Meerschaert (2004) applied two modifications of the quadratic estimator that makes it both

shift and scale variant.

When $0 < \alpha < 2$, the above formula (4.14) yields a consistent estimator for $1/\alpha$ for a broad range of time series models, including periodic moving averages, whose innovations have regularly varying tails with index α . If $\alpha > 2$, then the estimator $\hat{\gamma} \rightarrow \frac{1}{2}$ in probability as $n \rightarrow \infty$. Since we believe that $2 < \alpha < 4$, we apply the estimator to the squared data, which has tail index $\alpha/2$. The MS estimator is unbiased when the data are Pareto-like with $C = 1$. For Pareto this can be accomplished by dividing by $C^{1/\alpha}$, so we divide the raw data by $C^{1/\alpha} = 1338$, obtained from Hill's estimator with $r = 15$. Then we applied the above estimator of α and doubled the result (since the index of the squared data is $\alpha/2$) to get 2.623 as the MS estimate of $\hat{\alpha}$ for the raw Salt River data .

These tail estimations indicates that $2 < \alpha < 4$, so the probability distribution of X has heavy tails, with a finite variance but infinite fourth moment.

4.5 Application to Modeling of Natural River Flows

In this section we illustrate the application of heavy tail time series methods by fitting a periodic ARMA model to the Salt River flow data. Let \tilde{X}_t denote the average flow rate in cubic feet per second t months after October 1914 of the Salt River near Roosevelt, Arizona.

Preliminary estimates of the tail index in the previous section indicate that $2 < \alpha < 4$, so the probability distribution of \tilde{X}_t has heavy tails, with a finite variance but infinite fourth moment. A partial plot of the original data is given in Figure 4.3, clearly shows the cyclic behavior of the monthly flows. There are occasional sharp spikes in the time series plot of the data (see Figure 4.3), which indicates the heavy tail characteristics of the data.

The sample mean, standard deviation and autocorrelations at lag 1 and lag 2 are given in Table 4.2. The nonstationarity of the series is apparent since the mean, standard deviation

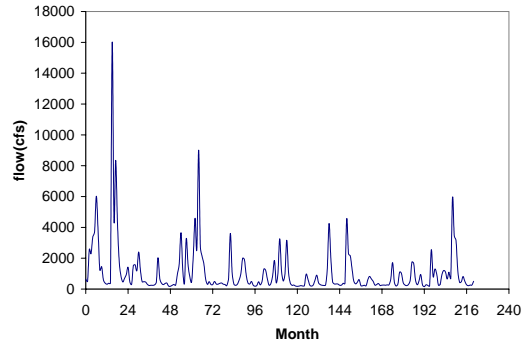


Figure 4.3: Average monthly flows (cfs) for the Salt River near Roosevelt, AZ, indicate a seasonal pattern and heavy tails.

and correlation functions vary widely from month to month. Removing the periodicity in mean and variance will not yield a stationary series. Therefore, a periodically stationary series model is appropriate.

In the following analysis, $i = 0$ corresponds to October and $i = 11$ corresponds to September. Since $2 < \alpha < 4$, Theorem 2.2 is applicable here. After $k = 15$ iterations, the innovations algorithm yields the innovations estimates found in Table 4.3. Following the same procedure as in Section 3.4, we selected a $\text{PARMA}_{12}(1, 1)$ model to the Salt River flow data.

The parameter estimates for this model, obtained using equations (3.44) and (3.45), are summarized in Table 4.4. It must be noted that θ_t in (3.44) must be replaced with $-\theta_t$ so as to be consistent with the representation in (2.1). Model residuals were estimated using equation (3.62). Figure 4.4 shows the ACF and PACF of the model residuals. Although a few values lie slightly outside of the 95% confidence bands, there is no apparent pattern, providing some evidence that the $\text{PARMA}_{12}(1, 1)$ model is adequate.

One reason for carefully modeling the river flow time series is to develop the ability to generate synthetic river flows for further analysis. This requires a realistic distributional

Table 4.2: Sample mean, standard deviation and autocorrelation at lag 1 and 2 of average monthly flow series for the Salt River near Roosevelt, AZ.

month	Parameter			
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\rho}(1)$	$\hat{\rho}(2)$
OCT	464.43	796.38	0.315	0.173
NOV	373.96	361.90	0.740	0.230
DEC	802.15	1289.14	0.273	0.336
JAN	1012.64	2055.94	0.354	0.530
FEB	1372.51	1792.97	0.486	0.479
MAR	1910.04	1800.47	0.619	0.562
APR	2107.85	1697.76	0.897	0.871
MAY	1064.12	1059.41	0.939	0.280
JUN	371.80	303.08	0.298	-0.120
JUL	341.93	396.27	0.340	0.143
AUG	589.57	481.41	0.415	-0.015
SEP	449.97	332.06	0.140	0.207

model for the residuals that can be used to simulate the innovations. A number of possible standard distributions (normal, lognormal, gamma, Weibull, etc.) were tried to model the residuals. It seems that none of them gave an adequate fit both at the center and the tails. For example, a summary (histogram) of the residuals as well as probability plot for a three parameter lognormal distribution are shown in Figure 4.5. The plots clearly indicate that normal and lognormal are not a good fit for the residuals.

The Hill's estimators for the PARMA₁₂(1, 1) Model Residuals are shown in Figure 4.6. Although Hill's estimator indicates $2 < \alpha < 4$, we investigate use of a stable distribution

Table 4.3: Moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, 2, \dots, 10$, after $k = 15$ iterations of the innovations algorithm applied to average monthly flow series for the Salt River near Roosevelt, AZ.

i	$\hat{\psi}_i(1)$	$\hat{\psi}_i(2)$	$\hat{\psi}_i(3)$	$\hat{\psi}_i(4)$	$\hat{\psi}_i(5)$	$\hat{\psi}_i(6)$	$\hat{\psi}_i(7)$	$\hat{\psi}_i(8)$	$\hat{\psi}_i(9)$	$\hat{\psi}_i(10)$...
0	0.410	-0.122	0.002	2.637	0.347	-0.131	0.007	0.014	0.051	0.103	...
1	0.168	0.182	0.019	0.544	0.072	-0.164	-0.032	0.088	0.016	0.000	...
2	3.064	0.221	0.507	0.178	1.103	1.631	-0.636	-0.070	0.285	0.021	...
3	0.368	1.060	-0.189	-0.489	-0.032	1.673	4.882	0.381	0.227	0.140	...
4	0.147	0.084	1.454	0.136	0.529	0.032	2.283	4.711	-0.924	0.309	...
5	0.459	0.440	0.379	1.898	0.283	-0.488	-0.170	0.801	3.354	-0.131	...
6	0.236	0.489	0.366	0.423	2.724	0.225	0.309	-0.088	0.142	4.074	...
7	0.542	0.111	0.322	0.162	0.211	1.580	0.444	-0.036	-0.225	0.210	...
8	0.265	0.123	0.032	0.097	0.060	0.075	0.457	0.085	-0.019	-0.023	...
9	1.056	-0.055	0.126	0.009	0.025	0.003	0.026	0.051	-0.006	-0.056	...
10	0.559	0.785	0.082	-0.064	-0.057	-0.031	-0.011	-0.033	-0.161	-0.045	...
11	0.256	0.184	0.096	0.049	-0.029	-0.031	-0.004	0.011	0.017	-0.158	...

($\alpha < 2$). It has been observed that Hill's estimator can overestimate α for stable data (see, for example, McCulloch, 1997).

There are several methods for estimating stable parameters, but we focus on the three general methods (see, for example, Nolan, 1999): the maximum likelihood (ML) approach maximizes the likelihood numerically, the quantile method (QM) tries to match certain data quantiles with those of stable distributions, the characteristic function based methods fit the empirical characteristic function.

The model residuals were then fitted with a stable distribution $S1(\alpha, \beta, \gamma, \eta)$ and the

Table 4.4: Parameter estimates for $\text{PARMA}_{12}(1, 1)$ model of average monthly flow series for the Salt River near Roosevelt, AZ.

month	$\hat{\phi}$	$\hat{\theta}$	$\hat{\sigma}$
OCT	-0.476	0.886	696.20
NOV	0.443	-0.275	229.67
DEC	1.315	1.749	779.23
JAN	0.346	0.022	1781.03
FEB	0.227	-0.080	1263.85
MAR	3.000	-2.541	1291.93
APR	1.065	-0.829	1044.32
MAY	0.472	0.071	378.06
JUN	0.226	0.039	75.54
JUL	-0.208	1.265	346.75
AUG	0.743	-0.184	409.58
SEP	0.329	-0.074	295.17

parameters were estimated using three methods: ML, QM and the characteristic function based methods. The results are summarized in Table 4.5. The computer program STABLE 3.04 have been used for all the analysis (Nolan, 1999).

Since the estimates from the three methods are close for $\text{PARMA}_{12}(1, 1)$ model residuals, the residuals can be considered as stably distributed (see Nolan, 1999) for diagnostics of assessing stability). The next step is to do the smoothed density, q - q plot and the variance stabilized p - p plot based on the ML estimates. Figure 4.7 shows the smoothed density, q - q plot and the variance stabilized p - p plot for the model residuals. The q - q plots shows that the extreme tails of the data are lighter than the stable model. The p - p plot shows that the

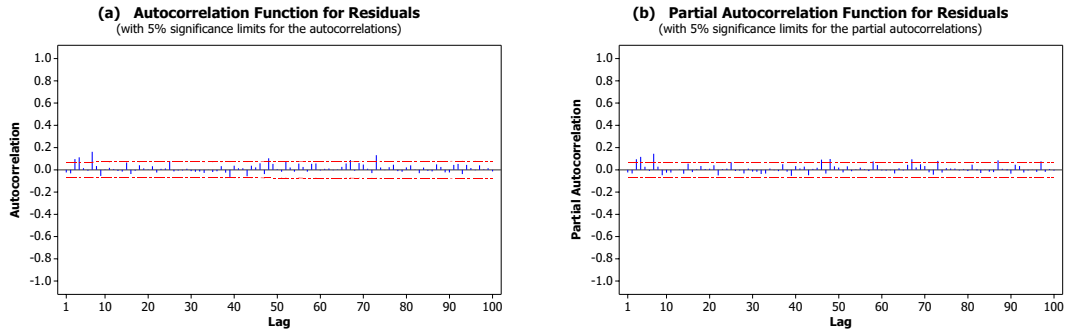


Figure 4.4: (a) ACF and (b) PACF for $PARMA_{12}(1, 1)$ model residuals, showing the 95% confidence bounds $\pm 1.96/\sqrt{N}$.

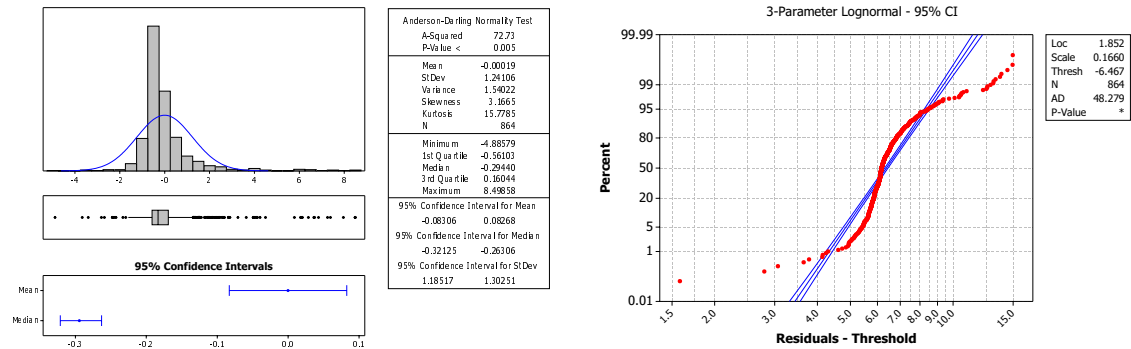


Figure 4.5: $PARMA_{12}(1, 1)$ Model residuals of the Salt River near Roosevelt, AZ: Statistical summary (left) and three parameter lognormal probability plot (right).

data look stable. The plots show a very close stable fit at the center (body), but not good at the tails. This suggest for the use of a mixture of stable with appropriate tail distributions.

Hill's estimators (4.12) and (4.13) give $\hat{\alpha} = 1.661$ and $\hat{C} = 0.163$ for the upper tail ($r = 36$ residuals), and $\hat{\alpha} = 2.320$ and $\hat{C} = 0.043$ for the lower tail ($r = 15$ residuals). Using (3.60), the p -values are 0.018 and 0.393, respectively. The small p -value as well as the visual evidence of the probability plot in Figure 4.8 suggest that the Pareto model is not a good fit for the upper tail of the estimated residuals. Therefore, we decided to fit a truncated Pareto to the upper tail ($x > 2.30$) of the residuals. Since $p > 0.05$, Pareto is a good fit for the lower

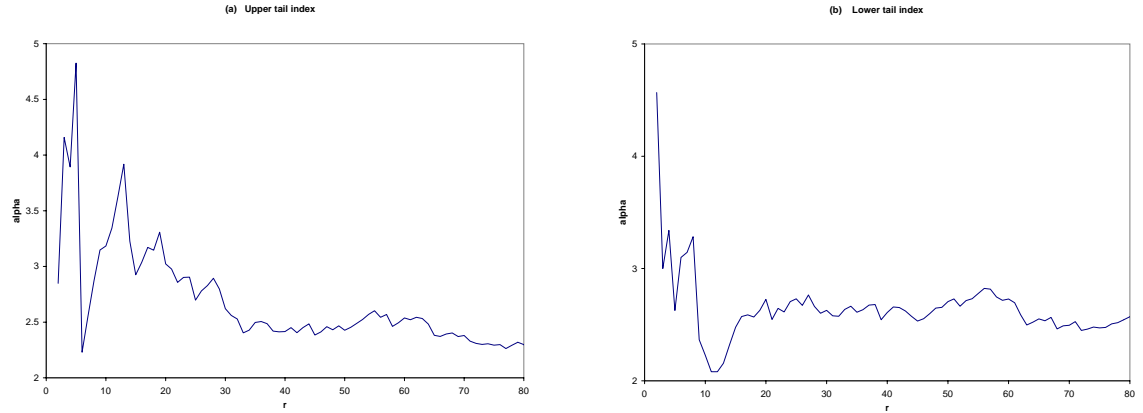


Figure 4.6: Hill's estimator, $\hat{\alpha}$ versus r , of (a) upper tail and (b) lower tail for the $\text{PARMA}_{12}(1, 1)$ model residuals.

Table 4.5: Parameter estimates for $\text{PARMA}_{12}(1, 1)$ model residuals.

Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\eta}$
Quantile	1.1762	0.6319	0.3220	0.3388
Sample Characteristic	1.2061	0.6264	0.3306	0.2238
Maximum Likelihood	1.2081	0.6129	0.3260	0.2414

tail. However, we chose a truncated Pareto for the lower tails ($x < -1.52$) of the residuals because river flows are always bounded from below (i.e. river flows are always greater than zero). Note that this means $E\varepsilon_t^4 < \infty$ so that results from chapter 3 also apply here. Next we determined that $r = 36$ residuals exceed 2.30 (which is the 96 percentile of the stable distribution), and $r = 15$ residuals fall below -1.52 (which equals the 2.26 percentile of the stable distribution). Then the MLE from Theorem 3.13 was used to estimate the parameters ($\hat{\beta} = 8.499$, $\hat{\gamma} = 0.003$, $\hat{\alpha} = 0.345$) of the best fitting truncated Pareto distribution, and the theoretical distribution tail $P(R > r)$ was plotted over the 36 largest positive residuals in Figure 4.8 (left). In Figure 4.8 (right), we used the same method to fit a truncated Pareto ($\hat{\beta} = 4.886$, $\hat{\gamma} = 0.093$, $\hat{\alpha} = 1.401$) to the 15 largest negative residuals, after a change of sign.

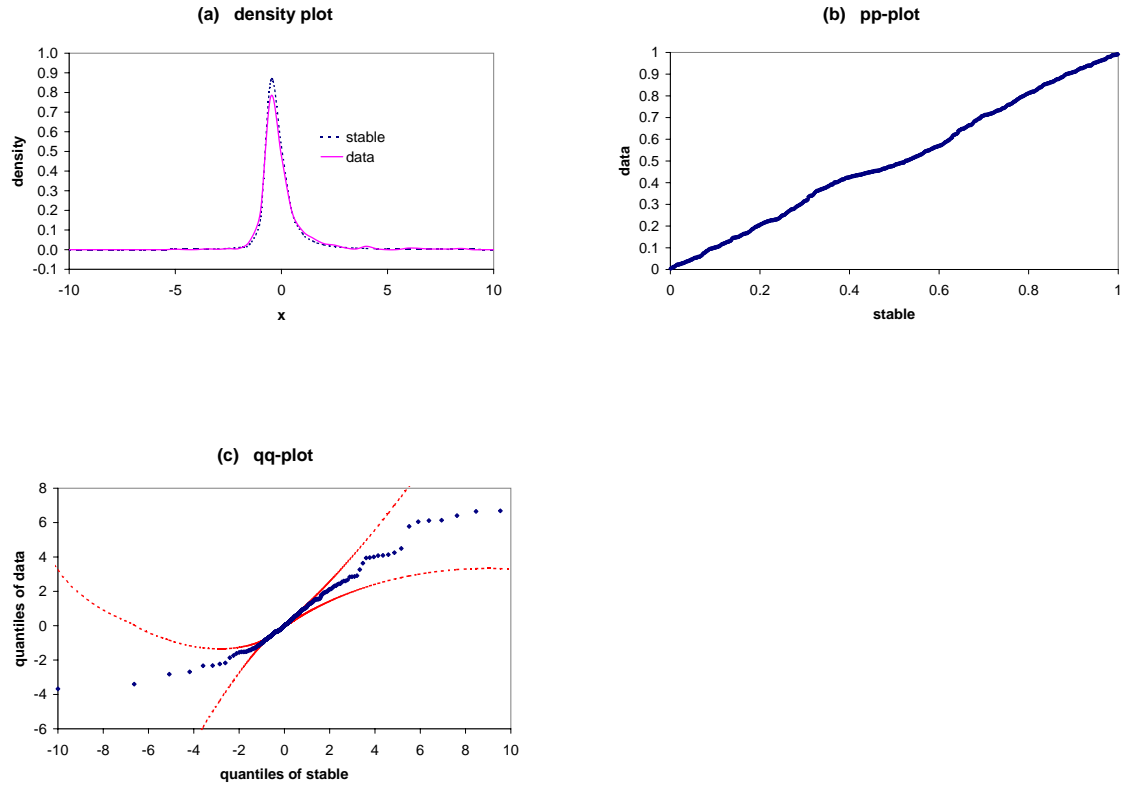


Figure 4.7: Density, variance stabilized p - p plot and q - q plot for the $\text{PARMA}_{12}(1, 1)$ model residuals.

Both of the plots in Figure 4.8 indicate an adequate fit.

A mixture distribution with stable body and truncated Pareto tails was used to simulate the innovations. The mixture has cumulative distribution function (cdf)

$$P(\delta \leq r) = \begin{cases} F_-(r) & \text{if } r < -1.52 \\ F_0(r) & \text{if } -1.52 \leq r \leq 2.30 \\ F_+(r) & \text{if } r > 2.30 \end{cases} \quad (4.15)$$

where F_0 is the cdf of the stable, and F_+ , F_- are truncated Pareto cdfs of the positive and negative tails, respectively. The truncated Pareto distributions were slightly shifted (by $s = -0.057$ on the positive tail and $s = 0.271$ on the negative tail) to make the mixture cdf continuous. Now stable innovations are simulated from $S1(1.2081, 0.6129, 0.3260, 0.2414)$

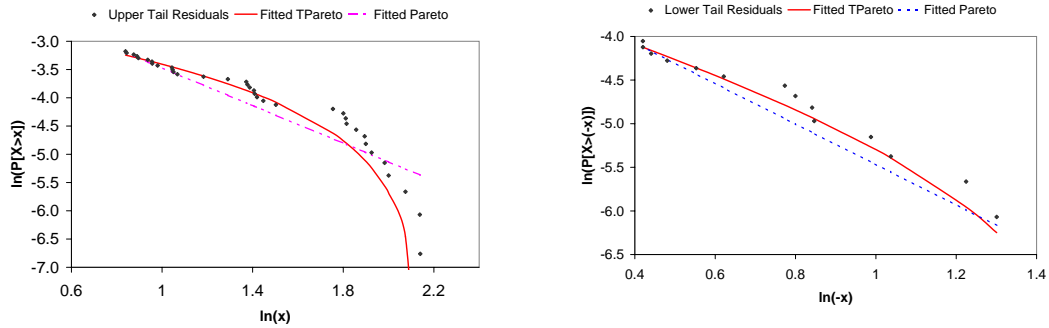


Figure 4.8: Log-log plot of upper (left) and lower (right) residual tails and fitted Pareto and truncated Pareto distributions, Salt River near Roosevelt, AZ.

by using STABLE 3.04 and innovations at the tails are obtained by using the inverse cumulative distribution function method. First, stable random variates were calculated from $S1(1.2081, 0.6129, 0.3260, 0.2414)$ using STABLE 3.04. If $R > 2.30$, the 96% quantile of stable, we generated a uniform $(0, 1)$ random variate U and substituted $\delta = F_+^{-1}(0.96 + 0.04U)$. If $R < -1.52$, the 2.26 % quantile of the stable, we substituted $\delta = F_-^{-1}(0.0226U)$. This gives simulated innovations δ with the mixture distribution (4.15).

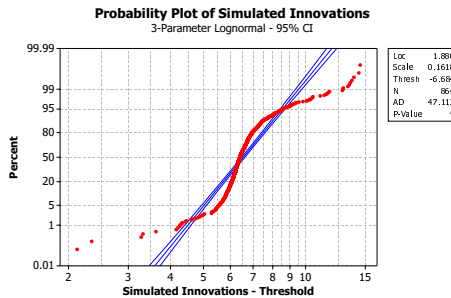


Figure 4.9: Probability plot of simulated innovations using the mixed stable and truncated Pareto distributions. Compare Figure 4.5 (right).

Figure 4.9 shows a probability plot for $N = N_y \nu$ simulated innovations (for $\nu = 12$ months and $N_y = 100$ years) from the mixture distribution (4.15). Comparison with Figure

4.5 (a) shows that the simulated innovations are statistically identical to the computed model residuals in terms of distribution. Substituting the simulated innovations into the model residuals in terms of distribution. Substituting the simulated innovations into the model (3.63) generates N_y years of simulated river flow. In rare cases, a simulated flow comes out negative. When this occurs, we resample (draw a different random innovation) for that period. It is advantageous to simulate several extra years of river flows and throw out the initial years (100 years in this case), since we did not simulate X_t for $t < 0$. This ensures that the simulated series is periodically stationary. Figure 4.10 gives a side-by-side comparison of the observed and simulated river flow time series, illustrating that the synthetic river flow appears statistically similar to the observed flow. Figure 4.11 shows the main statistical characteristics (mean, standard deviation and autocorrelations) of a typical synthetic river flow time series obtained by this method, as well as the same statistical measures for the observed time series. It is apparent that this procedure closely reproduces the main statistical characteristics, indicating that the modeling procedure is trustworthy for generating synthetic river flows. Such synthetic river flows are useful for design of hydraulic structures, for optimal operation of reservoir systems, for calculating the risk of failure of water supply systems, etc.

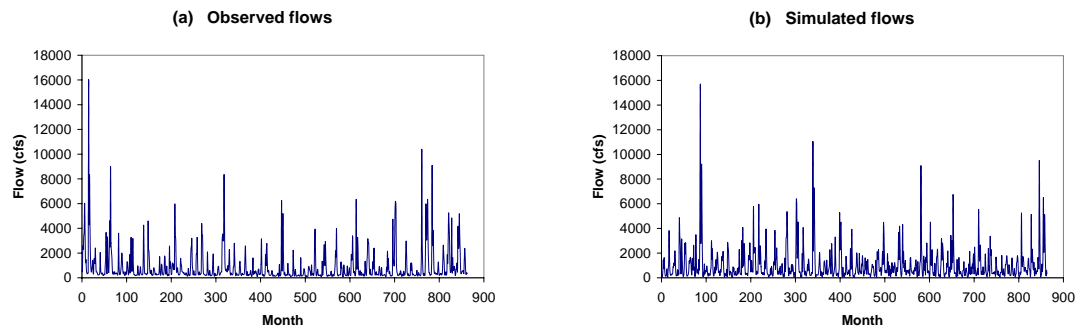


Figure 4.10: Plot of (a) observed and (b) simulated monthly river flows for the Salt River near Roosevelt, AZ, indicating statistical similarity.

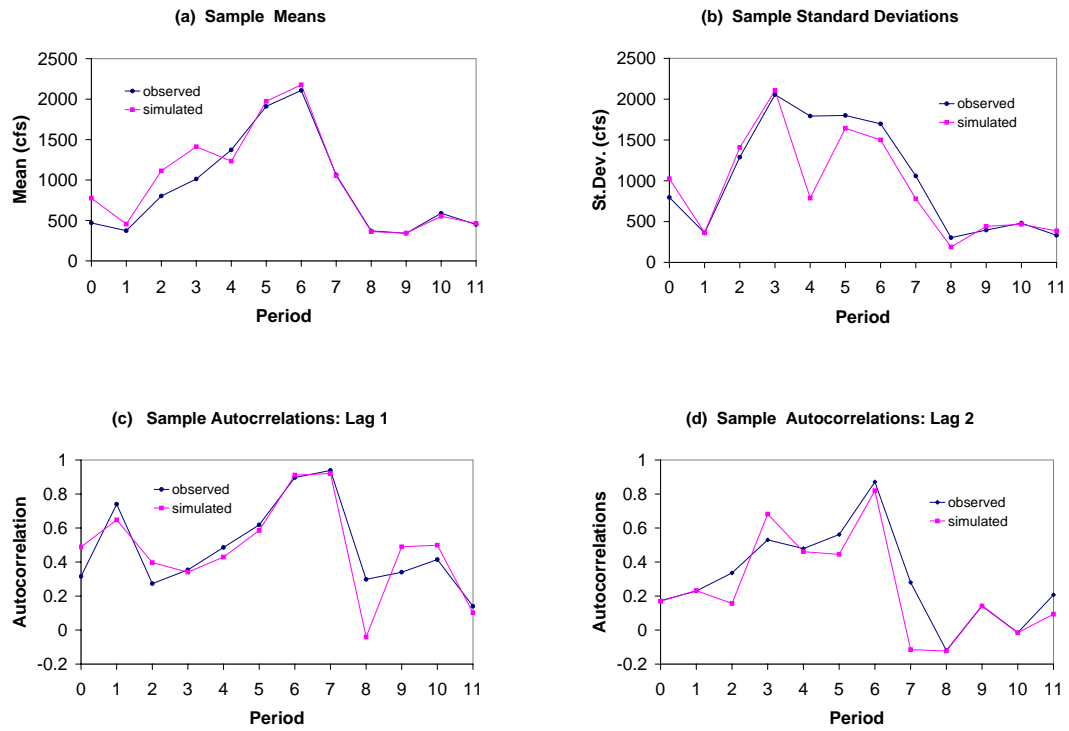


Figure 4.11: Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed monthly river flow data for the Salt River near Roosevelt, AZ.

5 Estimation of Periodic Parameters by Fourier Series

5.1 Overview

In most cases, PARMA models have been applied to monthly or quarter time series. However, when the number of periods is large (for instance, in weekly or daily data), PARMA models require an estimation of too many parameters, which violates the principle of statistical parsimony (model with minimum number of parameters). In this chapter, model parsimony is achieved by expressing the periodic model parameters in terms of their discrete Fourier transforms.

In most of the water resources systems design and operation studies the periodic phenomena have been represented by Fourier functions. Quimpo (1967) has applied Fourier analysis to daily river flow sequences in order to detect significant harmonic components embodied within the sequence considered. Since then Fourier analysis has become a standard tool in any hydrologic study concerning periodicity. Salas *et al.* (1980) proposed a Fourier series approach for reducing the number of parameters in PAR or PARMA models. Vecchia (1985a) also adopted the same approach but used Akaike's information criterion (AIC) for the selection of significant harmonics. Experience in using Fourier analysis for estimating periodic parameters of hydrologic time series shows that for small time interval series, such as daily and weekly series, only the first few harmonics are necessary for a good Fourier series fit in the periodic estimate of model parameters (Salas *et al.*, 1980). This practical criteria should be supplemented by more precise analysis and tests. For instance, Anderson and Vecchia (1993) uses asymptotic properties of the discrete Fourier transform of the estimated periodic autocovariance and autocorrelation function for selecting the harmonics in the PARMA model parameters.

In an effort to obtain a parsimonious model for periodically stationary series, we develop the asymptotic distribution of the discrete Fourier transform of the innovations estimates (ψ -weights of the periodic moving average processes) and then determine those statistically significant Fourier coefficients. We also extend these results to other periodic model parameters. We demonstrate the effectiveness of the technique using simulated data from different PARMA models. An application of the technique is demonstrated through the analysis of a weekly river flow series for the Fraser River, British Columbia.

5.2 Selection of Significant Harmonics and Fourier Coefficients

The $\text{PARMA}_\nu(p, q)$ model (1) has $(p + q + 1)\nu$ total parameters (see Lund and Basawa, 2000; Salas *et al.*, 1980). For example, for a monthly series ($\nu = 12$) where $p = q = 1$, there are 36 parameters. When the period ν is large it is reasonable to assume that the model parameters vary smoothly with respect to time and can therefore be explained by only a few of their non-zero discrete Fourier coefficients.

Experience in using Fourier analysis for estimating periodic parameters of hydrologic time series shows that for small time interval series, such daily and weekly series, only the first few harmonics are necessary for a good Fourier series fit in the periodic estimate of $\theta_t(j)$, $\theta_t(j)$ and σ_t (Salas *et al.*, 1980). This practical criteria should be supplemented by more precise analysis and tests.

More precise analysis for selecting the harmonics in Fourier series fit of a periodic estimate is based on a theorem for detecting periodic variation in the model parameters in (2.1) and (2.2). This theorem is derived by first obtaining the asymptotic distribution of the discrete Fourier transform of the parameters from the innovations algorithm.

Write the moving average parameters in (2.2) at lag j in the form

$$\psi_t(j) = c_0(j) + \sum_{r=1}^k \left\{ c_r(j) \cos\left(\frac{2\pi r t}{\nu}\right) + s_r(j) \sin\left(\frac{2\pi r t}{\nu}\right) \right\} \quad (5.1)$$

where $c_r(j)$ and $s_r(j)$ are the Fourier coefficients, r is the harmonic and k is the total number of harmonics which is equal to $\nu/2$ or $(\nu - 1)/2$ depending on whether ν is even or odd, respectively.

Write the vector of Fourier coefficients at lag j in the form

$$f(j) = \begin{cases} [c_0(j), c_1(j), s_1(j), \dots, c_{(\nu-1)/2}(j), s_{(\nu-1)/2}(j)]^T & (\nu \text{ odd}) \\ [c_0(j), c_1(j), s_1(j), \dots, s_{(\nu/2-1)}(j), c_{(\nu/2)}(j)]^T & (\nu \text{ even}) \end{cases} \quad (5.2)$$

where $[\cdot]$ denotes the greatest integer function. Similarly, define \hat{f}_j to be the vector of Fourier coefficients for the innovations estimates $\hat{\psi}_t(j)$, defined by replacing $\psi_t(j)$ by $\hat{\psi}_t(j)$, $c_r(j)$ by $\hat{c}_r(j)$, and $s_r(j)$ by $\hat{s}_r(j)$ in (5.1) and (5.2). We wish to describe the asymptotic distributional properties of these Fourier coefficients, to determine those that are statistically significantly different from zero. These are the coefficients that will be included in our model.

In order to compute the asymptotic distribution of the Fourier coefficients, it is convenient to work with the complex discrete Fourier transform (DFT) and its inverse

$$\begin{aligned} \psi_r^*(j) &= \nu^{-1/2} \sum_{t=0}^{\nu-1} \exp\left(\frac{-2i\pi r t}{\nu}\right) \psi_t(j) \\ \psi_t(j) &= \nu^{-1/2} \sum_{r=0}^{\nu-1} \exp\left(\frac{2i\pi r t}{\nu}\right) \psi_r^*(j) \end{aligned} \quad (5.3)$$

and similarly $\hat{\psi}_r^*(j)$ is the complex DFT of $\hat{\psi}_m(j)$. The complex DFT can also be written in matrix form. Recall from (3.30) the definition

$$\begin{aligned} \hat{\psi}(\ell) &= [\hat{\psi}_0(\ell), \hat{\psi}_1(\ell), \dots, \hat{\psi}_{\nu-1}(\ell)]^T \\ \psi(\ell) &= [\psi_0(\ell), \psi_1(\ell), \dots, \psi_{\nu-1}(\ell)]^T \end{aligned}$$

and similarly define

$$\begin{aligned}\hat{\psi}^*(j) &= [\hat{\psi}_0^*(j), \hat{\psi}_1^*(j), \dots, \hat{\psi}_{\nu-1}^*(j)]^T \\ \psi^*(j) &= [\psi_0^*(j), \psi_1^*(j), \dots, \psi_{\nu-1}^*(j)]^T\end{aligned}\tag{5.4}$$

noting that these are all ν -dimensional vectors. Define a $\nu \times \nu$ matrix U with complex entries

$$U = \nu^{-1/2} \left(e^{\frac{-i2\pi r t}{\nu}} \right)_{r,t=0,1,\dots,\nu-1}\tag{5.5}$$

so that $\psi^*(j) = U\psi(j)$ and $\hat{\psi}^*(j) = U\hat{\psi}(j)$. This matrix form is useful because it is easy to invert. Obviously $\psi^*(j) = U\psi(j)$ is equivalent to $\psi(j) = U^{-1}\psi^*(j)$ since the matrix U is invertible. This is what guarantees that there exists a unique vector of complex DFT coefficients $\psi^*(j)$ corresponding to any vector $\psi(j)$ of moving average parameters. But in this case, the matrix U is also unitary (i.e, $U\tilde{U}^T = I$) which means that $U^{-1} = \tilde{U}^T$, and the latter is easy to compute. Then we also have $\psi(j) = \tilde{U}^T\psi^*(j)$ which is the matrix form of the second relation in (5.3).

Next we convert from complex to real DFT, and it is advantageous to do this in a way that also involves a unitary matrix. Define

$$\begin{aligned}a_r(j) &= 2^{-1/2}\{\psi_r^*(j) + \psi_{\nu-r}^*(j)\} & (r = 1, 2, \dots, [(\nu-1)/2]) \\ a_r(j) &= \psi_r^*(j) & (r = 0 \text{ or } \nu/2) \\ b_r(j) &= i2^{-1/2}\{\psi_r^*(j) - \psi_{\nu-r}^*(j)\} & (r = 1, 2, \dots, [(\nu-1)/2])\end{aligned}\tag{5.6}$$

and let

$$e(j) = \begin{cases} [a_0(j), a_1(j), b_1(j), \dots, a_{(\nu-1)/2}(j), b_{(\nu-1)/2}(j)]^T & (\nu \text{ odd}) \\ [a_0(j), a_1(j), b_1(j), \dots, b_{(\nu/2-1)}(j), a_{(\nu/2)}(j)]^T & (\nu \text{ even}) \end{cases}\tag{5.7}$$

and likewise for the coefficients of $\hat{\psi}_t(j)$. These relations (5.6) define another $\nu \times \nu$ matrix P with complex entries such that

$$e(j) = PU\psi(j) = P\psi^*(j),\tag{5.8}$$

and it is not hard to check that P is also unitary, so that $\psi^*(j) = P^{-1}e(j) = \tilde{P}^T e(j)$, the latter form being most useful for computations. The DFT coefficients $a_r(j)$ and $b_r(j)$ are not the same as the coefficients $c_r(j)$ and $s_r(j)$ in (5.1) but they are closely related. Substitute the first line of (5.3) into (5.6) and simplify to obtain

$$\begin{aligned} a_r(j) &= \nu^{-1/2} \sum_{t=0}^{\nu-1} \cos\left(\frac{2\pi r t}{\nu}\right) \psi_t(j) & (r = 0 \text{ or } \nu/2) \\ a_r(j) &= \sqrt{\frac{2}{\nu}} \sum_{t=0}^{\nu-1} \cos\left(\frac{2\pi r t}{\nu}\right) \psi_t(j) & (r = 1, 2, \dots, [(\nu-1)/2]) \\ b_r(j) &= \sqrt{\frac{2}{\nu}} \sum_{tm=0}^{\nu-1} \sin\left(\frac{2\pi r t}{\nu}\right) \psi_t(j) & (r = 1, 2, \dots, [(\nu-1)/2]). \end{aligned} \quad (5.9)$$

Inverting the relations (5.6) or, equivalently, using the matrix equation $\psi^*(j) = \tilde{P}^T e(j)$, we obtain $\psi_r^*(j) = a_r(j)$ for $r = 0$ or $\nu/2$ and

$$\psi_r^*(j) = 2^{-1/2} \{a_r(j) - ib_r(j)\} \quad \text{and} \quad \psi_{\nu-r}^*(j) = \tilde{\psi}_r^*(j) = 2^{-1/2} \{a_r(j) + ib_r(j)\}$$

for $r = 1, \dots, k = [(\nu-1)/2]$. Substitute these relations into the second expression in (5.3) and simplify to obtain

$$\psi_t(j) = \nu^{-1/2} a_0(j) + \sqrt{\frac{2}{\nu}} \sum_{r=1}^k \left\{ a_r(j) \cos\left(\frac{2\pi r t}{\nu}\right) + b_r(j) \sin\left(\frac{2\pi r t}{\nu}\right) \right\}$$

for ν odd and

$$\psi_t(j) = \nu^{-1/2} (a_0(j) + a_k(j)) + \sqrt{\frac{2}{\nu}} \sum_{r=1}^{k-1} \left\{ a_r(j) \cos\left(\frac{2\pi r t}{\nu}\right) + b_r(j) \sin\left(\frac{2\pi r t}{\nu}\right) \right\}$$

for ν even, where k is the total number of harmonics which is equal to $\nu/2$ or $(\nu-1)/2$ depending on whether ν is even or odd, respectively. Comparison with (5.1) reveals that

$$\begin{aligned} c_r &= \sqrt{\frac{2}{\nu}} a_r & (r = 1, 2, \dots, [(\nu-1)/2]) \\ c_r &= \nu^{-1/2} a_r & (r = 0 \text{ or } \nu/2) \\ s_r &= \sqrt{\frac{2}{\nu}} b_r & (r = 1, 2, \dots, [(\nu-1)/2]). \end{aligned} \quad (5.10)$$

Substituting into (5.9) yields

$$\begin{aligned}
c_r(j) &= \nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi r m}{\nu}\right) \psi_m(j) & (r = 0 \text{ or } \nu/2) \\
c_r(j) &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi r m}{\nu}\right) \psi_m(j) & (r = 1, 2, \dots, [(\nu-1)/2]) \\
s_r(j) &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \sin\left(\frac{2\pi r m}{\nu}\right) \psi_m(j) & (r = 1, 2, \dots, [(\nu-1)/2])
\end{aligned} \tag{5.11}$$

and likewise for the Fourier coefficients of $\hat{\psi}_m(j)$. Define the $\nu \times \nu$ diagonal matrix

$$L = \begin{cases} \text{diag}(\nu^{-1/2}, \sqrt{2/\nu}, \dots, \sqrt{2/\nu}) & (\nu \text{ odd}) \\ \text{diag}(\nu^{-1/2}, \sqrt{2/\nu}, \dots, \sqrt{2/\nu}, \nu^{-1/2}) & (\nu \text{ even}) \end{cases} \tag{5.12}$$

so that in view of (5.10) we have $f(j) = Le(j)$ and $\hat{f}(j) = L\hat{e}(j)$. Substituting into (5.8) we obtain

$$f(j) = LPU\psi(j) \quad \text{and} \quad \hat{f}(j) = LPU\hat{\psi}(j). \tag{5.13}$$

Theorem 5.1 *For any positive integer j*

$$N_y^{1/2} [\hat{f}(j) - f(j)] \Rightarrow \mathcal{N}(0, R_V) \tag{5.14}$$

where

$$R_V = LPUV_{jj}\tilde{U}^T\tilde{P}^T L^T. \tag{5.15}$$

PROOF. From Theorem 3.7 we have

$$N_y^{1/2} [\hat{\psi}(j) - \psi(j)] \Rightarrow \mathcal{N}(0, V_{jj}) \tag{5.16}$$

where $V_{\ell k}$ is from (3.31). Define $B = LPU$ so that $f(j) = B\psi(j)$ and $\hat{f}(j) = B\hat{\psi}(j)$ using (5.13). Apply Proposition 3.8 with $\mathbf{X}_n = \hat{\psi}(j)$ and $g(\mathbf{X}) = B\mathbf{X}$, so that $\mu = \psi(j)$, $g(\mu) = B\psi(j) = f(j)$, and $D = [\partial g_i / \partial x_j] = B$. Then $N_y^{1/2}[B\hat{\psi}(j) - B\psi(j)] \Rightarrow \mathcal{N}(0, BV_{jj}B^T)$ or in other words $N_y^{1/2}[\hat{f}(j) - Bf(j)] \Rightarrow \mathcal{N}(0, BV_{jj}B^T)$. Although P and U are complex matrices, the product $B = LPU$ is a real matrix, and therefore $B^T = \tilde{B}^T = \tilde{U}^T\tilde{P}^T L^T$. Then (5.14) and (5.15) follow, which finishes the proof.

Theorem 5.2 *Let $X_t = \tilde{X}_t - \mu_t$, where X_t is the periodic moving average process (3.1) and μ_t is a periodic mean function with period ν . Then, under the null hypothesis that the mean-standardized process is stationary with $\psi_t(h) = \psi(h)$ and $\sigma_t = \sigma$, the elements of (5.2) are asymptotically independent with*

$$\begin{aligned} N_y^{1/2}\{\hat{c}_m(h) - \mu_m(h)\} &\Rightarrow \mathcal{N}(0, \nu^{-1}\eta_V(h)) && (m = 0 \text{ or } \nu/2) \\ N_y^{1/2}\{\hat{c}_m(h) - \mu_m(h)\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_V(h)) && (m = 1, 2, \dots, [(\nu - 1)/2]) \\ N_y^{1/2}\{\hat{s}_m(h) - \mu_m(h)\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_V(h)) && (m = 1, 2, \dots, [(\nu - 1)/2]) \end{aligned} \quad (5.17)$$

for all $h \geq 1$, where

$$\mu_m(h) = \begin{cases} \psi(h) & (m = 0) \\ 0 & (m > 0) \end{cases} \quad (5.18)$$

$$\eta_V(h) = \sum_{n=0}^{h-1} \psi^2(n). \quad (5.19)$$

PROOF. Under the null hypothesis, $\psi_t(h) = \psi(h)$ and $\sigma_t = \sigma$, is constant in t for each h and hence the F_n and B_n matrices in (3.32) become, respectively, a scalar multiple of the identity matrix: $F_n = \psi(n)I$, and an identity matrix: $B_n = I$. Then, from (3.31),

$$V_{hh} = \sum_{n=0}^{h-1} \psi(n)I\Pi^{(\nu-1)(h-1-n)}I\{\psi(n)I\Pi^{(\nu-1)(h-1-n)}\}^T = \sum_{n=0}^{h-1} \psi^2(n)I = \eta_V(h)I$$

is also a scalar multiple of the identity matrix. Hence, since scalar multiples of the identity matrix commute in multiplication with any other matrix, we have from (5.15) that

$$PUV_{hh}\tilde{U}^T\tilde{P}^T = V_{hh}PU\tilde{U}^T\tilde{P}^T = V_{hh}$$

since P and U are unitary matrices (i.e, $\tilde{P}^T P = I$ and $U\tilde{U}^T = I$). Then in Theorem 5.1 we have

$$N_y^{1/2} \left[\hat{f}(h) - f(h) \right] \Rightarrow \mathcal{N}(0, R_V) \quad (5.20)$$

where $R_V = LPUV_{hh}\tilde{U}^T\tilde{P}^T L^T = V_{hh}LL^T$, so that

$$R_V = \begin{cases} \eta_V(h) \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, 2\nu^{-1}) & (\nu \text{ odd}) \\ \eta_V(h) \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, \nu^{-1}) & (\nu \text{ even}). \end{cases}$$

Under the null hypothesis, $f(h) = [\psi(h), 0, \dots, 0]^T$ and then the theorem follows by considering the individual elements of the vector convergence (5.20).

Under the hypothesis that $c_m(h)$ and $s_m(h)$ are zero for all $m \geq 1$ and $h \neq 0$, Theorem 5.2 tells us that, for example, if ν is odd then $\{\hat{c}_1(j), \hat{s}_1(h), \dots, \hat{c}_{(\nu-1)/2}(h), \hat{s}_{(\nu-1)/2}(h)\}$ form $\nu - 1$ independent and normally distributed random variables with mean zero and standard error $(2\nu^{-1}\hat{\eta}_V(h)/N_y)^{1/2}$. The Bonferroni¹⁴ α -level test statistic rejects the null hypothesis that $c_m(h)$ and $s_m(h)$ are zero for all $m \geq 1$ if $|Z_c| > z_{\alpha'/2}$ and $|Z_s| > z_{\alpha'/2}$, respectively, and

$$Z_c = \frac{\hat{c}_m(h)}{(\lambda\hat{\eta}_V(h)/N_y)^{1/2}}, \quad Z_s = \frac{\hat{s}_m(h)}{(\lambda\hat{\eta}_V(h)/N_y)^{1/2}} \quad (5.21)$$

where

$$\lambda = \begin{cases} \nu^{-1} & (m = \nu/2) \\ 2\nu^{-1} & (m = 1, 2, \dots, [(\nu - 1)/2]) \end{cases}$$

$$\hat{\eta}_V(h) = \sum_{n=0}^{h-1} \hat{\psi}^2(n)$$

and $Z \sim N(0, 1)$, $\alpha' = \alpha/(\nu - 1)$. When $\alpha = 5\%$ and $\nu = 12$, $\alpha' = 0.05/11 = 0.0045$, $z_{\alpha'/2} = z_{0.0023} = 2.84$, and the null hypothesis is rejected when any $|Z_{c,s}| > 2.84$, indicating that the corresponding Fourier coefficients is statistically significantly different from zero.

¹⁴See Section A-6, Appendix A.

5.3 Estimation of Fourier Coefficients of Model Parameters

For large ν , it is anticipated that PARMA model parameters $\phi_t(\ell)$, $\theta_t(\ell)$ and σ_t , will vary smoothly with respect to t , and can therefore be explained by a few of their non-zero Fourier coefficients. For a $\text{PARMA}_\nu(p, q)$ model, the Fourier series representation of the $\phi_t(\ell)$, $\theta_t(\ell)$ and σ_t , can be obtained by

$$\begin{aligned}\theta_t(\ell) &= c_{a0}(\ell) + \sum_{r=1}^k \left\{ c_{ar}(\ell) \cos\left(\frac{2\pi r t}{\nu}\right) + s_{ar}(\ell) \sin\left(\frac{2\pi r t}{\nu}\right) \right\} \\ \phi_t(\ell) &= c_{b0}(\ell) + \sum_{r=1}^k \left\{ c_{br}(\ell) \cos\left(\frac{2\pi r t}{\nu}\right) + s_{br}(\ell) \sin\left(\frac{2\pi r t}{\nu}\right) \right\} \\ \sigma_t &= c_{d0} + \sum_{r=1}^k \left\{ c_{dr} \cos\left(\frac{2\pi r t}{\nu}\right) + s_{dr} \sin\left(\frac{2\pi r t}{\nu}\right) \right\}\end{aligned}\tag{5.22}$$

$c_{ar,br,dr}$ and $s_{ar,br,dr}$ are the Fourier coefficients, r is the harmonic and k is the total number of harmonics (as in (5.1)). For instance, for monthly series where $\nu = 12$, we have $k = 6$; for weekly series with $\nu = 52$, $k = 26$ and for daily series with $\nu = 365$, $k = 182$. In practice, a small number of harmonic $k^* < k$ is used.

For general Fourier analysis of $\text{PARMA}_\nu(p, q)$ model, it is desirable to represent (3.38) in terms of the Fourier coefficients of the model parameters and the Fourier transform of $\psi_t(j)$ (see Appendix B for a detailed discussion). However, we develop asymptotic distributions of the discrete Fourier coefficients of $\text{PARMA}_\nu(1, 1)$ model parameters by using the relationship in (3.44) and (3.45).

Consider again the $\text{PARMA}_\nu(1, 1)$ model given in (3.42). To simplify notation, we will express the model parameters, along with their Fourier coefficients, in terms of vector notation. Let $\theta = [\theta_0, \theta_1, \dots, \theta_{\nu-1}]^T$, $\phi = [\phi_0, \phi_1, \dots, \phi_{\nu-1}]^T$ and $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{\nu-1}]^T$ be the vector of $\text{PARMA}_\nu(1, 1)$ model parameters. These model parameters may be defined in terms of their complex finite Fourier coefficients θ_t^* , ϕ_t^* and σ^* as follows

$$\begin{aligned}
\theta^*(\ell) &= U\theta(\ell) \quad \text{and} \quad \theta(\ell) = \tilde{U}^T\theta^*(\ell) \\
\phi^*(\ell) &= U\phi(\ell) \quad \text{and} \quad \phi(\ell) = \tilde{U}^T\phi^*(\ell) \\
\sigma^* &= U\sigma \quad \text{and} \quad \sigma = \tilde{U}^T\sigma^*
\end{aligned} \tag{5.23}$$

where “ \sim ” denotes the complex conjugate, U is the $\nu \times \nu$ Fourier transform matrix defined in (5.5) and

$$\begin{aligned}
\theta^* &= [\theta_0^*, \theta_1^*, \dots, \theta_{\nu-1}^*]^T \\
\phi^* &= [\phi_0^*, \phi_1^*, \dots, \phi_{\nu-1}^*]^T \\
\sigma^* &= [\sigma_0^*, \sigma_1^*, \dots, \sigma_{\nu-1}^*]^T
\end{aligned}$$

As in Theorem 5.1 let the vector form for transformed θ and ϕ be given by

$$\begin{aligned}
f_\theta &= LP\theta^* = LPU\theta \\
f_\phi &= LP\phi^* = LPU\phi
\end{aligned} \tag{5.24}$$

where

$$f_\theta = \begin{cases} [c_{a0}, c_{a1}, s_{a1}, \dots, c_{a(\nu-1)/2}, s_{a(\nu-1)/2}]^T & (\nu \text{ odd}) \\ [c_{a0}, c_{a1}, s_{a1}, \dots, s_{a(\nu/2-1)}, c_{a(\nu/2)}]^T & (\nu \text{ even}) \end{cases} \tag{5.25}$$

$$f_\phi = \begin{cases} [c_{b0}, c_{b1}, s_{b1}, \dots, c_{b(\nu-1)/2}, s_{b(\nu-1)/2}]^T & (\nu \text{ odd}) \\ [c_{b0}, c_{b1}, s_{b1}, \dots, s_{b(\nu/2-1)}, c_{b(\nu/2)}]^T & (\nu \text{ even}) \end{cases} \tag{5.26}$$

$$\begin{aligned}
c_{ar} &= \nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi rm}{\nu}\right) \theta_m & (r = 0 \text{ or } \nu/2) \\
c_{ar} &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi rm}{\nu}\right) \theta_m & (r = 1, 2, \dots, [(\nu-1)/2]) \\
s_{ar} &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \sin\left(\frac{2\pi rm}{\nu}\right) \theta_m & (r = 1, 2, \dots, [(\nu-1)/2])
\end{aligned} \tag{5.27}$$

$$\begin{aligned}
c_{br} &= \nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi r m}{\nu}\right) \phi_m & (r = 0 \text{ or } \nu/2) \\
c_{br} &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \cos\left(\frac{2\pi r m}{\nu}\right) \phi_m & (r = 1, 2, \dots, [(\nu-1)/2]) \\
s_{br} &= 2\nu^{-1} \sum_{m=0}^{\nu-1} \sin\left(\frac{2\pi r m}{\nu}\right) \phi_m & (r = 1, 2, \dots, [(\nu-1)/2])
\end{aligned} \tag{5.28}$$

and likewise for the Fourier coefficients of $\hat{\theta}_m$ and $\hat{\phi}_m$. Note that $[\cdot]$ denotes the greatest integer function. We wish to describe the asymptotic distributional properties of the elements of (5.25) and (5.26) .

Theorem 5.3 *Regarding the Fourier transformed ϕ and θ , we have*

$$N^{1/2} [\hat{f}_\theta - f_\theta] \Rightarrow \mathcal{N}(0, R_S) \tag{5.29}$$

$$N^{1/2} [\hat{f}_\phi - f_\phi] \Rightarrow \mathcal{N}(0, R_Q)$$

$$f_\theta = LP\theta^* = LPU\theta$$

$$R_S = LPUS\tilde{U}^T\tilde{P}^T L^T \tag{5.30}$$

$$f_\phi = LP\phi^* = LPU\phi$$

$$R_Q = LPUQ\tilde{U}^T\tilde{P}^T L^T$$

where “ \sim ” denotes the complex conjugate, Q is given by (3.48) and S is given by (3.56).

PROOF. Same procedure as in proof of Theorem 5.1 by applying Proposition 3.8 along with Theorem 3.9 and Theorem 3.11.

Theorem 5.4 *Let $X_t = \tilde{X}_t - \mu_t$, where X_t is the $PARMA_\nu(1, 1)$ process (3.42) and μ_t is a periodic mean function with period ν . Then, under null hypothesis that the mean-standardized process is stationary with $\phi_t = \phi$, $\theta_t = \theta$ and $\sigma_t = \sigma$, the elements of \hat{f}_ϕ , defined by (5.26) with c_{br} replaced by \hat{c}_{br} and s_{br} replaced by \hat{s}_{br} , are asymptotically independent with*

$$\begin{aligned}
N_y^{1/2}\{\hat{c}_{bm} - \mu_{bm}\} &\Rightarrow \mathcal{N}(0, \nu^{-1}\eta_Q) & (m = 0 \text{ or } \nu/2) \\
N_y^{1/2}\{\hat{c}_{bm} - \mu_{bm}\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_Q) & (m = 1, 2, \dots, [(\nu - 1)/2]) \\
N_y^{1/2}\{\hat{s}_{bm} - \mu_{bm}\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_Q) & (m = 1, 2, \dots, [(\nu - 1)/2])
\end{aligned} \tag{5.31}$$

where

$$\mu_{bm} = \begin{cases} \phi & (m = 0) \\ 0 & (m > 0) \end{cases} \tag{5.32}$$

$$\eta_Q = \psi^{-4}(1) \left\{ \psi^2(2) \left(1 - \frac{2\psi^2(1)}{\psi(2)} \right) + \psi^2(1) \sum_{n=0}^1 \psi^2(n) \right\} \tag{5.33}$$

$\psi(1) = \phi - \theta$, and $\psi(2) = \phi\psi(1)$.

PROOF. The proof follows along the same lines as Theorem 3.9 and hence we adopt the same notation. As in proof of Theorem 5.2, we have $B_n = I$ and $F_n = \psi(n)I$ in (3.32) and so (3.31) implies ($x = \text{Min}(h, j)$)

$$\begin{aligned}
V_{jh} &= \sum_{n=0}^{x-1} \psi(j-1-n) \Pi^{(\nu-1)(j-1-n)} I \Pi^{(\nu-1)(h-1-n)T} \psi(h-1-n) \\
&= \sum_{n=0}^{x-1} \psi^2(j-1-n) I \quad \text{if } j = h
\end{aligned}$$

since then $\Pi^{(\nu-1)(j-1-n)} \Pi^{(\nu-1)(h-1-n)T} = I$ so

$$V_{11} = \psi^2(0)I = I$$

$$V_{22} = [\psi^2(1) + \psi^2(0)]I = [\psi^2(1) + 1]I$$

Note that, for $h \neq j$, $\Pi^{(\nu-1)(j-1-n)} I \Pi^{(\nu-1)(h-1-n)T} = \Pi^{(\nu-1)(j-1-n)} \Pi^{-(\nu-1)(h-1-n)} = \Pi^{(\nu-1)(j-h)}$

Then

$$V_{12} = \psi(0) \Pi^{(\nu-1)(1-2)} \psi(1) = \psi(1) \Pi$$

$$V_{21} = \psi(1) \Pi^{-1} \psi(0) = \psi(1) \Pi^{-1} = \psi(1) \Pi^T$$

and

$$Q = \begin{pmatrix} H_1 & H_2 \end{pmatrix} \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{pmatrix} H_1^T \\ H_2^T \end{pmatrix}$$

where $V_{21} = V_{12}^T \Rightarrow Q$ is symmetric.

Since X_t is stationary, every $\psi_\ell(t) = \psi(t)$ in (3.1) and so

$$H_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & \frac{-\psi(2)}{\psi^2(1)} \\ \frac{-\psi(2)}{\psi^2(1)} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \frac{-\psi(2)}{\psi^2(1)} & 0 \end{pmatrix} = \frac{-\psi(2)}{\psi^2(1)} \Pi^{-1}$$

and

$$H_2 = \frac{1}{\psi(1)} I$$

$$V_{11} = B_1 = I$$

$$V_{22} = B_2 + F_1 \Pi^{-1} B_1 \Pi F_1 = I + \psi^2(1) \Pi^{-1} I \Pi = [\psi^2(1) + 1] I$$

$$V_{12} = B_1 \Pi F_1 = \Pi F_1 = \psi(1) \Pi$$

$$V_{21} = V_{12}^T = \psi(1) \Pi^{-1} = \psi(1) \Pi^T$$

$$\begin{aligned} Q &= H_1 V_{11} H_1^T + H_2 V_{21} H_1^T + H_1 V_{12} H_2^T + H_2 V_{22} H_2^T \\ &= \frac{-\psi(2)}{\psi^2(1)} \Pi^{-1} I \left(\frac{-\psi(2)}{\psi^2(1)} \Pi \right) + \frac{-\psi(2)}{\psi^2(1)} \Pi^{-1} \psi(1) \Pi \frac{1}{\psi(1)} \\ &\quad + \frac{1}{\psi(1)} I \psi(1) \Pi^{-1} \left(\frac{-\psi(2)}{\psi^2(1)} \Pi \right) + \frac{1}{\psi(1)} I [\psi^2(1) + 1] I \frac{1}{\psi(1)} I \\ Q &= \left\{ \frac{\psi^2(2) - 2\psi^2(1)\psi(2) + [\psi^2(1) + 1]\psi^2(1)}{\psi^4(1)} \right\} I \end{aligned}$$

So $Q = \eta_Q I$ is actually a scalar multiple of the identity matrix I . Then $PUQ\tilde{U}^T\tilde{P}^T = QPU\tilde{U}^T\tilde{P}^T = Q$ and hence $R_Q = LQL^T = QLL^T$ or in other words

$$R_Q = \begin{cases} \eta_Q \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, 2\nu^{-1}) & (\nu \text{ odd}) \\ \eta_Q \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, \nu^{-1}) & (\nu \text{ even}). \end{cases}$$

Under the null hypothesis, $f_\phi = [\phi, 0, \dots, 0]^T$ and then the theorem follows by considering the individual elements of the vector convergence from the second line of (5.29).

Under the hypothesis that c_{bm} and s_{bm} are zero for all $m \geq 1$ and $h \neq 0$, Theorem 5.4 tells us that, for example, if ν is odd then $\{\hat{c}_{b1}, \hat{s}_{b1}, \dots, \hat{c}_{b(\nu-1)/2}, \hat{s}_{b(\nu-1)/2}\}$ form $\nu - 1$ independent and normally distributed random variables with mean zero and standard error $(2\nu^{-1}\hat{\eta}_S/N_y)^{1/2}$. The Bonferroni α -level test statistic rejects the null hypothesis that c_{bm} and s_{bm} are zero for all $m \geq 1$ and $h \neq 0$ if $|Z_c| > z_{\alpha'/2}$ and $|Z_s| > z_{\alpha'/2}$, respectively. Here, $Z \sim N(0, 1)$, $\alpha' = \alpha/(\nu - 1)$, and

$$Z_c = \frac{\hat{c}_{bm}(h)}{(\lambda\hat{\eta}_Q/N_y)^{1/2}}, \quad Z_s = \frac{\hat{s}_{bm}(h)}{(\lambda\hat{\eta}_Q/N_y)^{1/2}} \quad (5.34)$$

where

$$\lambda = \begin{cases} \nu^{-1} & (m = \nu/2) \\ 2\nu^{-1} & (m = 1, 2, \dots, [(\nu - 1)/2]) \end{cases}$$

$$\hat{\eta}_Q = \hat{\psi}^{-4}(1) \left\{ \hat{\psi}^2(2) \left(1 - \frac{2\hat{\psi}^2(1)}{\hat{\psi}(2)} \right) + \hat{\psi}^2(1) \sum_{n=0}^1 \hat{\psi}^2(n) \right\}.$$

Theorem 5.5 *Let $X_t = \tilde{X}_t - \mu_t$, where X_t is the $PARMA_\nu(1, 1)$ process (3.42) and μ_t a periodic mean function with period ν . Then, under null hypothesis that the mean-standardized process is stationary with $\theta_t = \theta$, $\phi_t = \phi$ and $\sigma_t = \sigma$, the elements of \hat{f}_θ , defined by (5.25) with c_{br} replaced by \hat{c}_{br} and s_{br} replaced by \hat{s}_{br} , are asymptotically independent with*

$$\begin{aligned} N_y^{1/2}\{\hat{c}_{am} - \mu_{am}\} &\Rightarrow \mathcal{N}(0, \nu^{-1}\eta_S) & (m = 0 \text{ or } \nu/2) \\ N_y^{1/2}\{\hat{c}_{am} - \mu_{am}\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_S) & (m = 1, 2, \dots, [(\nu - 1)/2]) \\ N_y^{1/2}\{\hat{s}_{am} - \mu_{am}\} &\Rightarrow \mathcal{N}(0, 2\nu^{-1}\eta_S) & (m = 1, 2, \dots, [(\nu - 1)/2]) \end{aligned} \quad (5.35)$$

where

$$\mu_{am} = \begin{cases} \theta & (m = 0) \\ 0 & (m > 0) \end{cases} \quad (5.36)$$

$$\eta_S = \psi^{-4}(1) \left\{ \psi^2(2) \left(1 - \frac{2\psi^2(1)}{\psi(2)} \right) + \sum_{j=1}^2 \psi^{4/j}(1) \sum_{n=0}^{j-1} \psi^2(n) \right\} \quad (5.37)$$

$\psi(1) = \phi - \theta$, and $\psi(2) = \phi\psi(1)$.

PROOF. The proof follows along the same lines as Theorem 3.11 and hence we adopt the same notation. Note that $S = Q + \bar{S}$ where

$$\bar{S} = V_{11} - H_1 V_1 - V_{11} H_1^T - H_2 V_{21} - V_{12} H_2^T = I$$

so as in Theorem 5.4 proof we get $R_S = LPUS\tilde{U}^T\tilde{P}^T L^T = SLPU\tilde{U}^T\tilde{P}^T L^T = SLL^T$ where

$$S = R_Q - I = \left\{ \frac{\psi^2(2) - 2\psi^2(1)\psi(2) + [\psi^2(1) + 1]\psi^2(1) + \psi^4(1)}{\psi^4(1)} \right\} I.$$

so that

$$R_S = \begin{cases} \eta_S \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, 2\nu^{-1}) & (\nu \text{ odd}) \\ \eta_S \text{diag}(\nu^{-1}, 2\nu^{-1}, \dots, 2\nu^{-1}, \nu^{-1}) & (\nu \text{ even}). \end{cases}$$

Under the null hypothesis, $f_\theta = [\theta, 0, \dots, 0]^T$ and then the theorem follows by considering the individual elements of the vector convergence from the first line of (5.29).

Under the hypothesis that c_{am} and s_{am} are zero for all $m \geq 1$ and $h \neq 0$, Theorem 5.5 tells us that, for example, if ν is odd then $\{\hat{c}_{a1}, \hat{s}_{a1}, \dots, \hat{c}_{a(\nu-1)/2}, \hat{s}_{a(\nu-1)/2}\}$ form $\nu - 1$ independent and normally distributed random variables with mean zero and standard errors obtained from $(\lambda\hat{\eta}_S/N_y)^{1/2}$ where λ is defined below. The Bonferroni α -level test statistic rejects the null hypothesis that c_{am} and s_{am} are zero for all $m \geq 1$ and $h \neq 0$ if $|Z_c| > z_{\alpha'/2}$ and $|Z_s| > z_{\alpha'/2}$, respectively. Here $Z \sim N(0, 1)$, $\alpha' = \alpha/(\nu - 1)$, and

$$Z_c = \frac{\hat{c}_{am}(h)}{(\lambda \hat{\eta}_S / N_y)^{1/2}}, \quad Z_s = \frac{\hat{s}_{am}(h)}{(\lambda \hat{\eta}_S / N_y)^{1/2}} \quad (5.38)$$

where

$$\lambda = \begin{cases} \nu^{-1} & (m = \nu/2) \\ 2\nu^{-1} & (m = 1, 2, \dots, [(\nu - 1)/2]) \end{cases}$$

$$\hat{\eta}_S = \hat{\psi}^{-4}(1) \left\{ \hat{\psi}^2(2) \left(1 - \frac{2\hat{\psi}^2(1)}{\hat{\psi}(2)} \right) + \sum_{j=1}^2 \hat{\psi}^{4/j}(1) \sum_{n=0}^{j-1} \hat{\psi}^2(n) \right\}.$$

5.4 Simulation Study

A detailed simulation study was conducted to demonstrate the effectiveness of the results of the previous section using simulated data from different $\text{PARMA}_\nu(p, q)$ models with finite fourth moment. For each model, individual realizations of $N_y = 50, 100, 300,$ and 500 years of data (i.e., sample size of $N = N_y\nu$) were simulated and the innovations algorithm was used to obtain parameter estimates for each realization. In each case, estimates were obtained for $k = 15$ iterations. Then discrete Fourier transformed innovation estimates and model parameters were obtained from (5.11), (5.27) and (5.28). The test statistics for the Fourier coefficients were computed using (5.21), (5.34) and (5.38) to identify those estimates that were statistically significant. A FORTRAN program was used to simulate the PARMA samples as well as to make all the necessary calculations.

As an example, we summarize here the results of two particular cases of a $\text{PARMA}_{12}(p, q)$ model and one case of a $\text{PARMA}_{52}(p, q)$ model. We first consider the following $\text{PARMA}_{12}(0, 1)$ model

$$X_{k\nu+i} = \varepsilon_{k\nu+i} + \theta_i \varepsilon_{k\nu+i-1} \quad (5.39)$$

where

$$\begin{aligned}\theta_i &= c_{a0} + \sum_{r=1}^6 \left\{ c_{ar} \cos\left(\frac{2\pi r i}{\nu}\right) + s_{ar} \sin\left(\frac{2\pi r i}{\nu}\right) \right\} \\ \sigma_i &= c_{d0} + \sum_{r=1}^6 \left\{ c_{dr} \cos\left(\frac{2\pi r i}{\nu}\right) + s_{dr} \sin\left(\frac{2\pi r i}{\nu}\right) \right\}\end{aligned}$$

The periodic notation $X_{k\nu+i}$ refers to the mean zero simulated data during the i^{th} period of cycle k . The innovations $\{\delta_t = \sigma_i^{-1} \varepsilon_{k\nu+i}\}$ form an iid normal sequence of random variables with mean zero and standard deviation one.

For example, suppose we wish to test the hypothesis that θ_i is dominated by two harmonics of period 12,

$$\theta_i = c_{a0} + \sum_{r=1}^2 \left\{ c_{ar} \cos\left(\frac{2\pi r i}{12}\right) + s_{ar} \sin\left(\frac{2\pi r i}{12}\right) \right\} \quad (5.40)$$

For $c_{a0} = 0.45$, $c_{a1} = 0.25$, $s_{a1} = 0.75$, $c_{a2} = 0.80$ and $s_{a2} = 0.50$, the values of θ_i are calculated from (5.40), and assume $\sigma_i = 2.00 + 0.15 \cos \frac{2\pi i}{12} + 0.90 \sin \frac{2\pi i}{12}$. Then using (5.39), a single realization with $N_y = 500$ years of data (i.e., sample size of $N = 6000$) was generated. Table 5.2 shows the results after $k = 15$ iterations of the innovations algorithm.

Table 5.1: Model parameters and estimates for simulated $\text{PARMA}_{12}(0, 1)$ data.

	c_{a0}	c_{a1}	s_{a1}	c_{a2}	s_{a2}
True Value	0.45	0.25	0.75	0.80	0.50
Estimated	0.463	0.282	0.725	0.806	0.519

Fourier coefficients with test statistics $z > 3.32$ (for $\alpha = 1\%$) are considered to be significant. It seems that the ψ -weights are dominated by a few of Fourier coefficients; especially $c_m(j)$ and $s_m(j)$ where $m = 1, 2$. Besides, the Fourier coefficients are statistically insignificant after lag 1, which indicate that the simulated data are from $\text{PARMA}_{12}(0, 1)$. The estimated standard deviation ($\hat{\sigma}_i \approx 1.92 + 0.15 \cos \frac{2\pi i}{12} + 0.88 \sin \frac{2\pi i}{12}$) is close to the true

values. The results of the residuals analysis also indicate that they look independent and normally distributed with mean zero and variance one. Those significant Fourier coefficients and the true values are summarized in Table 5.1 (neglecting the smaller coefficients). Again, the results show the closeness of the estimates.

The second model for which simulation results will be presented is a $\text{PARMA}_{12}(1,1)$ model

$$X_{k\nu+i} = \phi_i X_{k\nu+i-1} + \varepsilon_{k\nu+i} + \theta_i \varepsilon_{k\nu+i-1} \quad (5.41)$$

For example, suppose we wish to test the hypothesis that θ_i and ϕ_i are dominated by two harmonics of period 12,

$$\phi_i = c_{b0} + \sum_{k=1}^2 \left\{ c_{bk} \cos\left(\frac{2\pi ki}{12}\right) + s_{bk} \sin\left(\frac{2\pi ki}{12}\right) \right\} \quad (5.42)$$

For coefficients given in Table 5.5, the values of θ_i and ϕ_i are calculated from (5.41), and (5.42), respectively. Here we assume constant innovation variance ($\sigma_i^2 = 1$ for all i). From the above model, a single realization with $N_y = 500$ years of quarterly data (sample size of $N = 6000$) was generated. Table 5.3 shows the results after $k = 15$ iterations of the innovations algorithm. Since the discrete Fourier transform of $\hat{\psi}_i$ weights do not generally cut-off to (statistically) zero at a certain lag, we choose a parsimonious mixed model that captures the periodic behavior as well as the exponential decay evidenced in the autocorrelation function. After making all the diagnostics, we found that a $\text{PARMA}_{12}(1,1)$ model is adequate. The discrete Fourier transform of parameter estimates for this model, obtained using equations (5.27) and (5.28), are summarized in Table 5.4. It seems that estimated model parameters ($\hat{\theta}_t, \hat{\phi}_t$) of $\text{PARMA}_{12}(1,1)$ model are approximately dominated by two harmonics, and these Fourier coefficients as well as true values are summarized in Table 5.5. The result shows the

closeness of the estimates.

The last model used to illustrate the Fourier analysis technique is $\text{PARMA}_{52}(1, 1)$. The same Fourier coefficients as for $\text{PARMA}_{12}(1, 1)$ model are used for calculating the $\text{PARMA}_{52}(1, 1)$ model parameters, and then a single realization with $N_y = 500$ years of weekly data (sample size of $N = 26000$) was generated. The Fourier transformed values of the estimated $\text{PARMA}_{52}(1, 1)$ model parameters, the standard errors and the test statistics are summarized in Table 5.6 and Table 5.7. Fourier coefficients with test statistics $z > 3.72$ (for $\alpha = 1\%$) are considered to be significant. They are roughly dominated by two harmonics, and these Fourier coefficients as well as true values are summarized in Table 5.8. The result shows again the closeness of the estimates.

5.5 Application to Modeling of River Flows

First, we present a detailed Fourier analysis of monthly Fraser river flow series. We found that $\text{PARMA}_{12}(1, 1)$ is a reasonable fit for monthly flows (see Section 4.5). The monthly variation of the parameters (see Figure 5.1) is fairly smooth indicating that the Fourier transform of $\hat{\theta}_t$ and $\hat{\phi}_t$ may be dominated by a few components. The Fourier transformed values of these model parameters, the standard errors and the test statistics are summarized in Table 5.9. Based on Theorem 5.4 and Theorem 5.5, Fourier coefficients with test statistics $z > 3.32$ (for $\alpha = 1\%$) are considered to be significant. It seems that the $\text{PARMA}_{12}(1, 1)$ model parameters are dominated by few Fourier coefficients (see Table 5.9). Those coefficient with a test statistic close to 3.32 (s_{a2}, s_{a4}, s_{b4}) can also be considered statistically insignificant so as to achieve model parsimony (see the effect in Figure 5.1)

Using those significant Fourier coefficients ($c_{a0} = 0.304$, $s_{a1} = -0.426$, $c_{a3} = 0.665$, $c_{b0} = 0.337$, $s_{b1} = 0.466$, $c_{b2} = 0.408$, $s_{b2} = 0.355$ and $c_{b3} = -0.649$, see Table 5.9), the

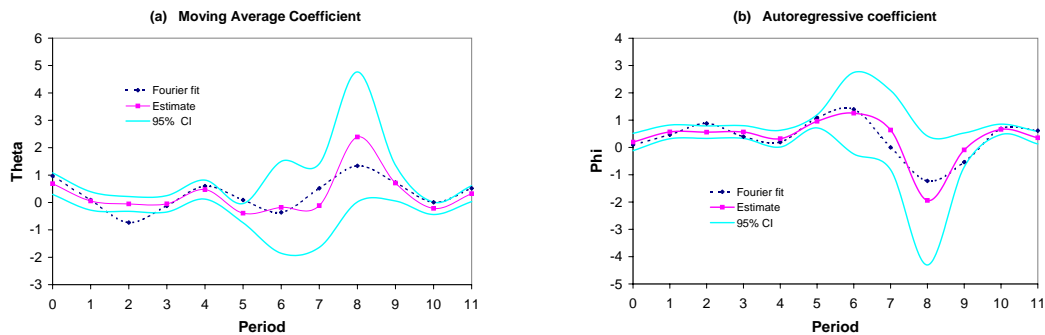


Figure 5.1: Plot of $\text{PARMA}_{12}(1,1)$ model parameters (with their Fourier fit) of average monthly flow series for the Fraser River at Hope, BC.

residuals are calculated again using (3.62). The ACF and PACF (not shown) indicate that there is no serial dependence. Statistical summary and lognormal probability plot of these residuals are shown in Figure 5.2, indicating a mixture of lognormal with an appropriate tail distribution (compare with Section 4.5).

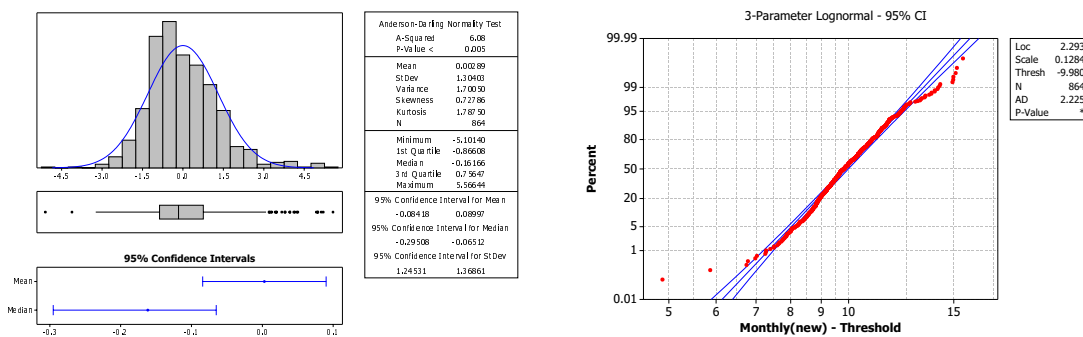


Figure 5.2: Statistical summary (a) and lognormal probability plot (b) for model residuals, Fourier model parameters, monthly Fraser river data. Compare Figure 3.7 (a).

Next we model the 72 years of weekly river flow data from October 1912 to September 1984, and use as a second example for Fourier analysis. The weekly river flow series are obtained by averaging daily measurements over 7 days (8 or 9 days at the end of a each water year) starting from October of each year. In this analysis, $\nu = 0$ corresponds to the

first 7 days of October and $\nu = 51$ corresponds to the last 7 to 9 days of September. The weekly statistics such as the mean, standard deviation, and lag-1 and lag-2 serial correlation coefficient are displayed in Figure 5.3. Following similar reasoning as in Section 4.5, we fit $\text{PARMA}_{52}(1, 1)$ model to the data. The parameter estimates for this model, obtained using equations (3.44) and (3.45), are shown in Figure 5.6. It must be noted that θ_t in (3.44) must be replaced with $-\theta_t$ so as to be consistent with the representation in (2.1). Model residuals were estimated using equation (3.62). Figure 5.4 shows the ACF and PACF of the model residuals. Although a few values lie slightly outside of the 95% confidence bands, there is no apparent pattern, providing some evidence that the $\text{PARMA}_{52}(1, 1)$ model is adequate.

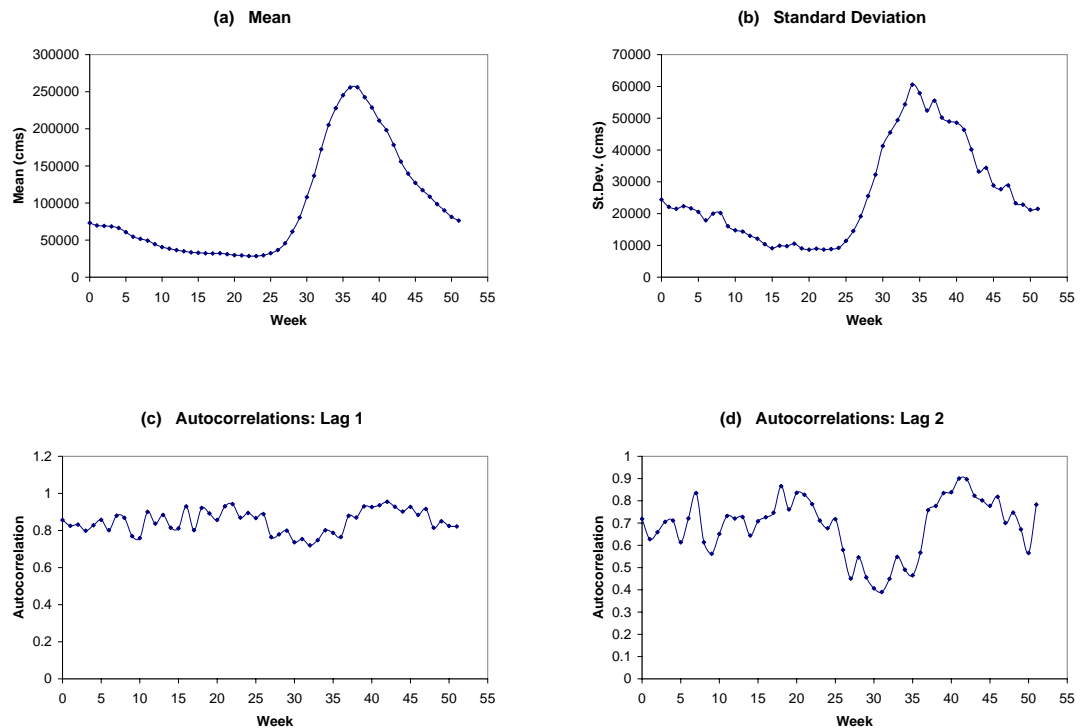


Figure 5.3: Sample Statistics (a) mean (b) standard deviation and autocorrelations at lag 1 (c) and lag 2 (d) of weekly river flow data for the Fraser River at Hope, BC.

Statistical summary and lognormal probability plot of model residuals are also shown

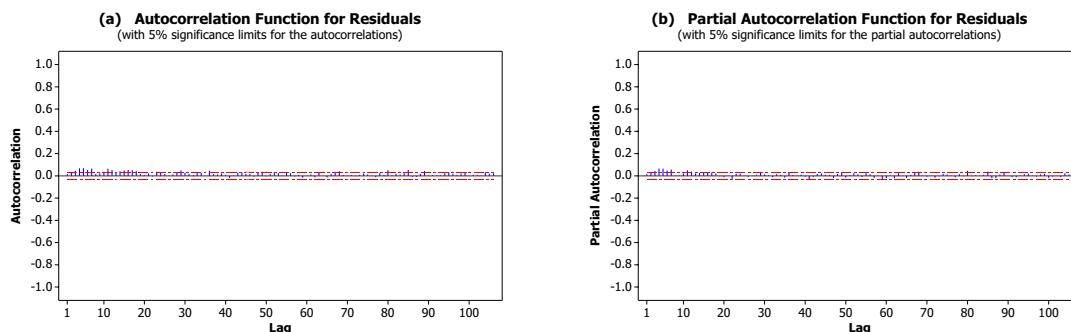


Figure 5.4: ACF and PACF for model residuals, showing the bounds $\pm 1.96/\sqrt{N}$, indicate no serial dependence.

in Figure 5.5, indicating a mixture of lognormal with an appropriate tail distribution (see Section 4.5). The purpose here is to test the hypothesis that the model parameters ($\hat{\theta}_t$ and $\hat{\phi}_t$) are dominated by few Fourier coefficients. The Fourier transformed values of these model parameters, the standard errors and the test statistics are summarized in Table 5.10 and Table 5.11. Based on Theorem 5.4 and Theorem 5.5, Fourier coefficients with test statistics $z > 3.72$ (for $\alpha = 1\%$) are considered to be significant. It is true that the $\text{PARMA}_{52}(1, 1)$ model parameters are dominated by few Fourier coefficients (see Table 5.10 and Table 5.11).

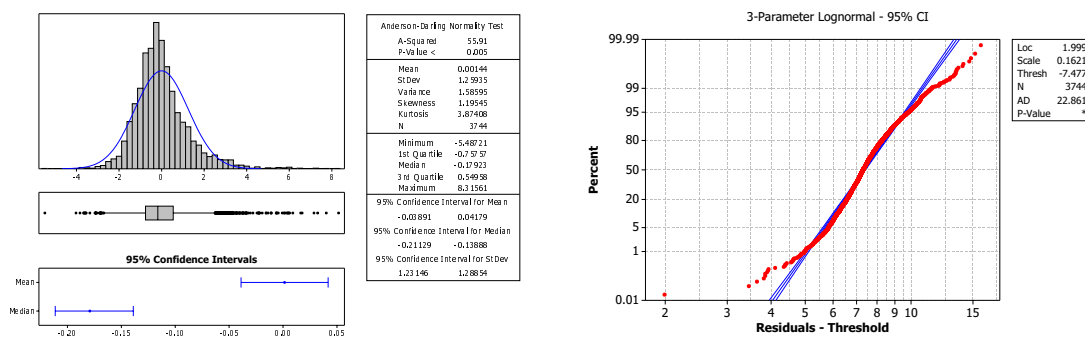


Figure 5.5: Statistical summary (left) and lognormal probability plot (right) for model residuals, weekly Fraser river data. Compare Figure 3.7

The correlation analysis of residuals (obtained from (3.62), for example), using the coef-

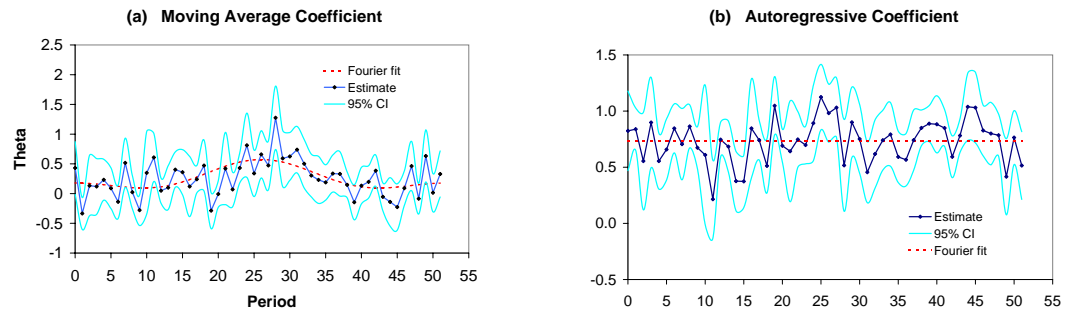


Figure 5.6: Plot of $\text{PARMA}_{52}(1, 1)$ model parameters (with their Fourier fit) of average weekly flow series for the Fraser River at Hope, BC.

coefficients: $c_{a0} = 0.255$, $c_{a1} = -0.197$, $c_{a2} = 0.121$ and $c_{b0} = 0.732$) indicate that there is no serial dependence (see Chapter 6).

Table 5.2: Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 500$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in (\cdot) is the test statistic (5.21).

lag	Coefficient	harmonic m						
		0	1	2	3	4	5	6
1	$\hat{c}_m(1)$	0.463	0.282*	0.806*	-0.013	-0.008	-0.000	0.029
			(15.434)	(44.143)	(-0.700)	(-0.436)	(-0.016)	(2.230)
	$\hat{s}_m(1)$		0.725*	0.519*	0.005	-0.028	-0.013	
			(39.711)	(28.441)	(0.288)	(-1.544)	(-0.719)	
	SE		0.018	0.018	0.018	0.018	0.018	0.013
2	$\hat{c}_m(2)$	0.015	0.053	0.027	0.049	0.001	-0.011	-0.015
			(2.678)	(1.354)	(2.466)	(0.054)	(-0.536)	(-1.073)
	$\hat{s}_m(2)$		0.001	-0.004	-0.037	-0.008	0.006	
			(0.058)	(-0.210)	(-1.887)	(-0.456)	(0.299)	
	SE		0.020	0.020	0.020	0.020	0.020	0.014
3	$\hat{c}_m(3)$	0.010	0.035	0.091*	0.049	0.038	0.006	0.017
			(1.750)	(4.604)	(2.495)	(1.939)	(0.319)	(1.219)
	$\hat{s}_m(3)$		-0.024	-0.004	0.031	-0.015	-0.007	
			(-1.211)	(-0.220)	(1.579)	(-0.741)	(-0.370)	
	SE		0.020	0.020	0.020	0.020	0.020	0.014
4	$\hat{c}_m(4)$	0.007	0.027	0.045	-0.007	-0.008	-0.018	-0.021
			(1.349)	(2.270)	(-0.352)	(-0.408)	(-0.917)	(-1.490)
	$\hat{s}_m(4)$		0.016	0.067*	0.022	0.036	0.010	
			(0.798)	(3.421)	(1.091)	(1.815)	(0.523)	
	SE		0.020	0.020	0.020	0.020	0.020	0.014
:	:	:	:	:	:	:	:	:

*Fourier coefficients with test statistic ≥ 3.32

Table 5.3: Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 500$ years of simulated PARMA₁₂(1, 1) data. Note that the value in (.) is the test statistic (5.21).

lag	Coefficient	harmonic m						
		0	1	2	3	4	5	6
1	$\hat{c}_m(1)$	0.698	0.410*	0.683*	-0.037	-0.021	0.018	-0.001
			(22.452)	(37.417)	(-2.052)	(-1.129)	(0.989)	(-0.085)
	$\hat{s}_m(1)$		0.762*	0.180*	0.015	-0.024	-0.012	
			(41.727)	(9.871)	(0.833)	(-1.337)	(-0.664)	
	SE		0.018	0.018	0.018	0.018	0.018	0.013
2	$\hat{c}_m(2)$	0.378	0.221*	0.226*	-0.050	0.147*	-0.010	0.004
			(10.464)	(10.724)	(-2.373)	(6.963)	(-0.472)	(0.271)
	$\hat{s}_m(2)$		0.378*	0.226*	0.277*	0.112*	0.006	
			(17.897)	(10.726)	(13.131)	(5.302)	(0.286)	
	SE		0.022	0.022	0.022	0.022	0.022	0.015
3	$\hat{c}_m(3)$	0.157	0.055	0.009	-0.077*	-0.061	-0.119*	-0.032
			(2.530)	(0.399)	(-3.533)	(-2.813)	(-5.488)	(-2.057)
	$\hat{s}_m(3)$		0.225*	0.151*	0.157 *	0.067 *	-0.005	
			(10.329)	(3.958)	(7.227)	(3.084)	(-0.229)	
	SE		0.022	0.022	0.022	0.022	0.022	0.015
4	$\hat{c}_m(4)$	0.069	-0.075*	-0.036	-0.106*	-0.045	0.036	0.024
			(-3.406)	(-1.624)	(-4.851)	(-2.070)	(1.637)	(1.561)
	$\hat{s}_m(4)$		0.126*	0.058	0.016	-0.059	-0.037	
			(5.771)	(2.644)	(0.717)	(-2.691)	(-1.707)	
	SE		0.022	0.022	0.022	0.022	0.022	0.015
:	:	:	:	:	:	:	:	:

*Fourier coefficients with test statistic ≥ 3.32

Table 5.4: Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated PARMA₁₂(1, 1) data. Note that the value in (.) is the test statistic (5.34) and (5.38).

Parameter	Statistic	harmonic m						
		0	1	2	3	4	5	6
$\hat{\theta}_t$	\hat{c}_{am}	0.361	0.181*	0.228*	-0.023	-0.048	0.009	-0.020
			(5.711)	(7.204)	(-0.715)	(-1.529)	(0.289)	(-0.895)
	\hat{s}_{am}		0.418*	0.314*	0.021	-0.021	-0.047	
			(13.234)	(9.916)	(0.668)	(-0.652)	(-1.473)	
	SE		0.032	0.032	0.032	0.032	0.032	0.022
$\hat{\phi}_t$	\hat{c}_{bm}	1.165	0.229*	0.455*	-0.015	0.028	0.009	0.019
			(6.280)	(12.471)	(-0.406)	(0.760)	(0.244)	(0.733)
	\hat{s}_{bm}		0.343*	-0.133*	-0.006	-0.004	0.034	
			(9.404)	(-3.652)	(-0.162)	(-0.104)	(0.943)	
	SE		0.022	0.022	0.022	0.022	0.022	0.015

*Fourier coefficients with test statistic ≥ 3.32

Table 5.5: Significant discrete Fourier transform coefficients for simulated PARMA₁₂(1, 1) model parameters.

		c_{a0}	c_{a1}	s_{a1}	c_{a2}	s_{a2}
θ_t	True Value	0.35	0.15	0.40	0.25	0.35
	Estimated	0.361	0.181	0.418	0.228	0.313
		c_{b0}	c_{b1}	s_{b1}	c_{b2}	s_{b2}
ϕ_t	True Value	0.35	0.25	0.35	0.45	-0.15
	Estimated	0.336	0.229	0.343	0.455	-0.133

Table 5.6: Discrete Fourier transform of moving average parameter estimate, $\hat{\theta}_t$ (with standard error, SE = 0.009 for $m = 26$ and SE = 0.013 for $m \neq 26$) for the simulated PARMA₅₂(1, 1) data. Note that the value in (.) is the test statistic (5.38).

	harmonic m									
	0	1	2	3	4	5	6	7	8	
\hat{c}_{am}	0.301	0.148*	0.310*	-0.046*	0.063*	0.007	-0.049*	0.054*	-0.033	
		(11.836)	(24.675)	(-3.691)	(5.014)	(0.559)	(-3.918)	(4.723)	(-2.611)	
\hat{s}_{am}		0.452*	0.315*	0.003	0.028	-0.070*	0.019	0.009	-0.076*	
		(36.031)	(25.097)	(0.259)	(2.235)	(-5.508)	(1.480)	(0.712)	(-6.036)	
	9	10	11	12	13	14	15	16	17	
\hat{c}_{am}	-0.031	0.054*	-0.068*	-0.008	0.036	-0.033	0.028	0.060*	-0.049*	
	(-2.495)	(4.339)	(-5.388)	(-0.626)	(2.901)	(-2.603)	(2.207)	(4.816)	(-3.910)	
\hat{s}_{am}	0.056 *	-0.018	-0.037	0.066*	-0.019	-0.025	0.079*	-0.049*	-0.027	
	(4.451)	(-1.422)	(-2.974)	(5.257)	(-1.544)	(-1.972)	(6.301)	(-3.920)	(-2.174)	
	18	19	20	21	22	23	24	25	26	
\hat{c}_{am}	-0.018	0.056*	-0.039	0.025	0.061*	-0.056*	0.025	0.002	-0.017	
	(-1.432)	(4.428)	(-3.117)	(2.024)	(4.895)	(-4.433)	(2.000)	(0.162)	(-1.868)	
\hat{s}_{am}	0.032	-0.041	0.021	0.069*	-0.074*	-0.015	0.028	-0.070*		
	(2.599)	(-3.244)	(1.687)	(5.475)	(-5.865)	(-1.218)	(2.242)	(-5.544)		

*Fourier coefficients with test statistic ≥ 3.72

Table 5.7: Discrete Fourier transform of autoregressive parameter estimate, $\hat{\phi}_t$ (with standard error, SE = 0.011 for $m = 26$ and SE = 0.015 for $m \neq 26$) for the simulated PARMA₅₂(1, 1) data. Note that the value in (.) is the test statistic (5.34).

	harmonics m									
	0	1	2	3	4	5	6	7	8	
\hat{c}_{am}	0.385	0.240	0.417*	0.053*	-0.055*	-0.007	0.061	-0.058	0.027	
		(15.661)	(27.233)	(3.460)	(-3.576)	(-0.485)	(4.014)	(-3.767)	(1.739)	
\hat{s}_{am}		0.280*	-0.127	-0.004	-0.021	0.064*	-0.038	-0.009	0.077*	
		(18.260)	(-8.308)	(-0.273)	(-1.355)	(4.183)	(-2.512)	(-0.559)	(5.029)	
	9	10	11	12	13	14	15	16	17	
\hat{c}_{am}	0.046	-0.051*	0.057*	0.016	-0.041	0.040	-0.027	-0.071*	0.062*	
	(3.014)	(-3.307)	(3.698)	(1.064)	(-2.697)	(2.609)	(-1.764)	(-4.624)	(4.024)	
\hat{s}_{am}	-0.060*	0.020	0.028	-0.068*	0.009	0.024	-0.070*	0.061*	0.029	
	(-3.920)	(1.308)	(1.813)	(-4.475)	(0.593)	(1.556)	(-4.563)	(3.955)	(1.910)	
	18	19	20	21	22	23	24	25	26	
\hat{c}_{am}	0.022	-0.046	0.048	-0.022	-0.047	0.044	-0.015	-0.011	0.018	
	(1.436)	(-3.012)	(3.154)	(-1.446)	(-3.071)	(2.855)	(-0.976)	(-0.694)	(1.643)	
\hat{s}_{am}	-0.049	0.041	-0.020	-0.073*	0.060*	0.010	-0.047	0.064*		
	(-3.178)	(2.691)	(-1.307)	(-4.750)	(3.902)	(0.676)	(-3.053)	(4.166)		

*Fourier coefficients with test statistic ≥ 3.72

Table 5.8: Significant discrete Fourier transform coefficients for simulated PARMA₅₂(1, 1) model parameters.

		c_{a0}	c_{a1}	s_{a1}	c_{a2}	s_{a2}
θ_t	True Value	0.35	0.15	0.40	0.25	0.35
	Estimated	0.300	0.148	0.452	0.310	0.314
		c_{b0}	c_{b1}	s_{b1}	c_{b2}	s_{b2}
ϕ_t	True Value	0.35	0.25	0.35	0.45	-0.15
	Estimated	0.385	0.240	0.279	0.417	-0.127

Table 5.9: Discrete Fourier transform of model parameters estimates and standard errors (SE) of average monthly flow series for the Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.34) and (5.38).

Parameter	Statistic		m						
	SE		0	1	2	3	4	5	6
$\hat{\theta}_t$	\hat{c}_{am}	0.090	0.304	0.011	-0.253	0.665*	-0.008	-0.244	0.214*
				(0.128)	(-2.816)	(7.406)	(-0.095)	(-2.721)	(3.373)
	\hat{s}_{am}		-0.426*	0.301*	0.039	-0.300*	0.080		
			(-4.741)	(3.351)	(0.433)	(-3.338)	(0.887)		
$\hat{\phi}_t$	\hat{c}_{bm}	0.102	0.337	-0.036	0.408*	-0.649*	0.142	0.159	-0.161
				(-0.355)	(3.998)	(-6.369)	(1.394)	(1.560)	(-2.239)
	\hat{s}_{bm}		0.466*	-0.355*	-0.023	0.327*	-0.157		
			(4.570)	(-3.487)	(-0.230)	(3.204)	(-1.537)		

*Fourier coefficients with test statistic ≥ 3.32

Table 5.10: Discrete Fourier transform of moving average parameter estimate, $\hat{\theta}_t$ (with standard error, SE = 0.017 for $m = 26$ and SE = 0.024 for $m \neq 26$) for PARMA₅₂(1, 1) model of average weekly flow series for Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.38).

	harmonics m									
	0	1	2	3	4	5	6	7	8	
\hat{c}_{am}	0.255	-0.197*	0.121*	-0.040	0.081	-0.022	-0.031	-0.013	-0.078	
		(-8.25)	(5.062)	(-1.670)	(3.388)	(-0.941)	(-1.300)	(-0.531)	(-3.283)	
\hat{s}_{am}		-0.0198	0.081	-0.117*	0.015	-0.044	-0.041	-0.020	-0.016	
		(-0.828)	(3.392)	(-4.89)	(0.633)	(-1.823)	(-1.731)	(-0.832)	(-0.680)	
	9	10	11	12	13	14	15	16	17	
\hat{c}_{am}	0.044	0.052	0.023	-0.071	0.009	0.021	0.139*	-0.053	0.017	
	(1.833)	(2.169)	(0.957)	(-2.996)	(0.392)	(0.880)	(5.836)	(-2.240)	(0.696)	
\hat{s}_{am}	0.018	0.027	-0.059	-0.034	-0.051	-0.071	0.041	0.073	-0.009	
	(0.753)	(1.118)	(-2.471)	(-1.436)	(-2.154)	(-2.993)	(1.714)	(3.078)	(-0.396)	
	18	19	20	21	22	23	24	25	26	
\hat{c}_{am}	0.031	-0.007	0.113*	0.019	0.075	-0.016	0.018	-0.065	0.019	
	(1.335)	(-0.280)	(4.71)	(0.795)	(3.151)	(-0.660)	(0.769)	(-2.741)	(1.110)	
\hat{s}_{am}	-0.014	-0.040	-0.076	-0.018	-0.091	-0.041	-0.016	0.008		
	(-0.569)	(-1.696)	(-3.177)	(-0.748)	(-3.829)	(-1.716)	(-0.669)	(0.351)		

*Fourier coefficients with test statistic ≥ 3.72

Table 5.11: Discrete Fourier transform of autoregressive parameter estimate, $\hat{\phi}_t$ (with standard error, SE = 0.023 for $m = 26$ and SE = 0.033 for $m \neq 26$) for PARMA₅₂(1, 1) model of average weekly flow series for the Fraser River at Hope, BC. Note that the value in (.) is the test statistic (5.34) .

	harmonics m									
	0	1	2	3	4	5	6	7	8	
\hat{c}_{am}	0.732	-0.021	0.051	-0.076	-0.005	-0.045	0.052	0.002	0.082	
		(-0.621)	(1.536)	(-2.302)	(-0.174)	(-1.362)	(1.557)	(0.058)	(2.483)	
\hat{s}_{am}		-0.065	-0.049	0.063	0.019	-0.009	0.020	0.043	0.000	
		(-1.957)	(-1.470)	(1.902)	(0.583)	(-0.275)	(0.588)	(1.305)	(0.000)	
	9	10	11	12	13	14	15	16	17	
\hat{c}_{am}	0.009	-0.043	-0.018	0.032	0.024	0.021	-0.055	0.044	0.010	
	(0.275)	(-1.298)	(-0.539)	(0.959)	(0.711)	(0.648)	(-1.667)	(1.319)	(0.287)	
\hat{s}_{am}	0.005	-0.011	0.024	-0.041	0.040	0.016	-0.064	-0.026	0.042	
	(0.150)	(-0.330)	(0.714)	(-1.226)	(1.194)	(0.493)	(-1.931)	(-0.797)	(1.279)	
	18	19	20	21	22	23	24	25	26	
\hat{c}_{am}	-0.019	0.036	-0.014	0.023	-0.003	0.000	-0.025	0.037	-0.000	
	(-0.580)	(1.071)	(-0.418)	(0.695)	(-0.103)	(0.000)	(-0.756)	(1.105)	(-0.007)	
\hat{s}_{am}	0.035	-0.001	0.066	0.004	0.071	0.043	-0.010	-0.009		
	(1.054)	(-0.019)	(1.999)	(0.119)	(2.124)	(1.283)	(-0.304)	(-0.282)		

*Fourier coefficients with test statistic ≥ 3.72

6 Application of PARMA Models: Hydrological Drought Analysis

6.1 General

Obtaining drought properties is important for planning and management of water resources system. For example, the design of water supply capacity of a given city may be based on meeting water demands during a critical drought that may occur in a specified planning horizon (Frick *et al.*, 1990). Moreover, the estimation of return periods associated to severe droughts can provide useful information in order to improve water systems management under drought condition.

It is difficult to give a universal definition since there exists widely diverse views about the interpretation of droughts among the scientists of different disciplines. In a broader sense, a drought is defined as a deficit of water in time. It can be divided into three categories- meteorological, hydrological, and agricultural drought-which may occur as a result of precipitation, river flow, and soil moisture deficits, respectively (see for example, Dracup *et al.* 1980; Sen 1980, etc). In this study the emphasis is on hydrological drought, which is defined as uninterrupted sequences of river flows below a threshold level. It has three components- duration, intensity, and severity. Since river flow sequences are stochastic variables, the corresponding drought properties are random and must be described in probabilistic terms.

Yevjevich (1967) proposed the theory of runs as a major tool to use in objectively defining droughts and studying their statistical properties. Many studies of drought properties followed Yevjevich's definition of droughts. For example, probabilistic behavior of drought properties has been derived analytically, assuming a given stochastic structure of the underlying hydrological series (e.g., Downer *et al.*, 1967; Llamas and Siddiqui, 1969;

Sen, 1976,1977,1980a; Guven, 1983; Zelenhastic and Salvai, 1987,2002; Mathier *et al.* 1992; Sharma, 1995; Chung and Salas, 2000; Shiau and Shen, 2001; Cancelliere and Salas, 2004).

Another subject of much interest when analyzing extreme events such as drought is the determination of the return period. The return period can be defined in different ways for different applications. Some authors (e.g., Lloyd, 1970; Loaiciga and Mariño, 1991; Fernandez and Salas, 1999a,b; Shiau and Shen, 2001) have assumed the return period as average elapsed time between occurrences of specified events (i.e., droughts exceeding a critical value). An alternative definition of return period is the average number of trials (or time periods) required to the first occurrence of a specified drought event (e.g. Vogel, 1987; Bras, 1990; Fernandez and Salas,1999a,b; Douglas *et al.*, 2002).

Hydrologic droughts can be determined from the historical record alone by using non-parametric methods but, because the number of drought events that can be drawn from the historical sample is generally small, the “historical” drought properties have a large degree of uncertainty. Other alternatives for finding drought properties include using stochastic models that can represent the underlying river flows, simulating long records of such river flows, and then deriving droughts properties from the simulated samples based on the theory of runs. In this chapter, we use PARMA models to model river flows and generate a large sequence of synthetic flows, and then drought properties are derived from the simulated samples based on the theory of runs. The applicability of these methods is demonstrated by using weekly river flow data for the Fraser River near Hope in British Colombia.

6.2 Definitions

An objective definition (as suggested by Yevjevich, 1967) will be made on the basis of a given sequence of river flows $\{X_t\}$ and a preselected level of reference X_{0t} , see Figure

6.1. A hydrologic drought event is defined as uninterrupted sequences of river flows below the preselected level X_{0t} (also called negative run). This preselected reference level can correspond to various hydrological quantities (see, for example, Sen, 1976; Dracup et al. 1980a). In case of water supply, X_{0t} may be considered equal to demand, which may be constant or variable. Thus the drought duration L (negative run length) is the number of consecutive time intervals in which $X_t < X_{0t}$, preceded and followed by (at least one interval where) $X_t \geq X_{0t}$. Likewise, a deficit at time t has a magnitude $S_t = X_{0t} - X_t$ so that the accumulated deficit (also called severity) D is the sum of deficit S_t over the drought duration L . The drought intensity is the average accumulated deficit, i.e., $I = D/L$. Periods of deficit and surplus are sometimes referred to as dry and wet periods, respectively. Such an objective and simple drought definition may enable one to characterize droughts using stochastic approaches.

Another subject of much interest when analyzing extreme events such as drought is the determination of the return period. In this analysis, return period of a hydrologic drought is defined as the mean interarrival time of hydrologic droughts with magnitude that are equal or greater than a certain arbitrary value M_0 (usually critical drought).

6.3 Methodology

Drought properties of various return periods are needed to assess the degree to which a water supply system will be able to cope with future droughts and accordingly, to plan alternate water supply strategies. They can be determined from the historical record alone by using non-parametric methods but, because the number of drought events that can be drawn from the historical sample is generally small, the “historical” drought properties have a large degree of uncertainty. In such cases, drought properties can be derived by synthetically

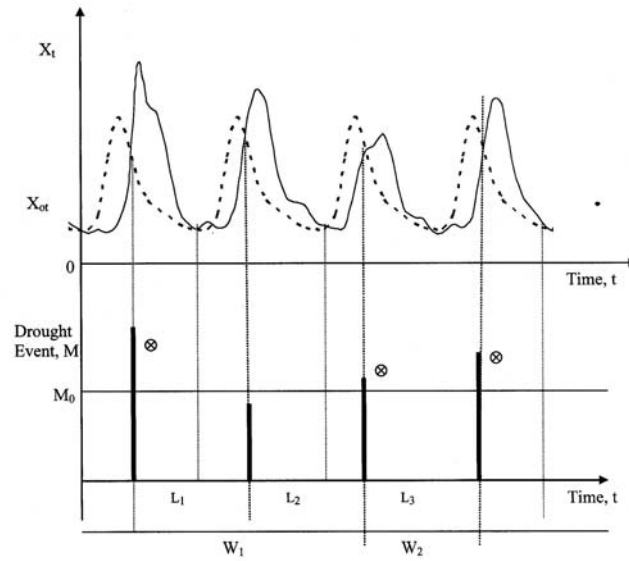


Figure 6.1: Definition of hydrological drought properties for a periodic river flow series X_t (solid line) and a periodic water demand X_{0t} (broken line); and definition of variables for estimation of return period.

generating river flows at key points in the water supply system under consideration.

Estimation Drought properties

Both drought duration and severity are random in character and, therefore, probabilistic distributions are useful to describe them. In general, several droughts result in a time series of given demand level and sample size N . Assume that m droughts with durations L_1, \dots, L_m and severities D_1, \dots, D_m occur. The means (\bar{L} , \bar{D}), standard deviation (\bar{s}_L , \bar{s}_D) and the maximum (L^* , D^*) of drought duration and severity are important characteristics describing a given time series.

$$\begin{aligned}
\bar{L} &= m^{-1} \sum_{j=1}^m L_j & \text{and} & & \bar{D} &= m^{-1} \sum_{j=1}^m D_j \\
\bar{s}_L &= \frac{\left[\sum_{j=1}^m (L_j - \bar{L})^2 \right]^{1/2}}{m-1} & \text{and} & & \bar{s}_D &= \frac{\left[\sum_{j=1}^m (D_j - \bar{D})^2 \right]^{1/2}}{m-1} \\
L^* &= \max[L_1, \dots, L_m] & \text{and} & & D^* &= \max[D_1, \dots, D_m]
\end{aligned} \tag{6.1}$$

It must be noted that for a given N and demand level X_{0t} , the above characteristics are random variables. In this analysis, the hydrologic drought properties are derived from the times series model (PARMA model) fitted to the historical river flows series by Monte Carlo simulation. That is, first a PARMA model is fitted to the observed river flow data, and a synthetic river flow series is generated from the PARMA model. Then, drought properties (e.g. maximum drought duration and severity) are determined by the runs analysis. This process is repeated several times to find the array of drought properties (e.g. maximum drought duration) whose average is an estimate of that particular property.

Estimation Return period

The return period of hydrologic droughts, based on the concept of stochastic processes, is derived in the following manner. Drought events M (severity D or duration L) are assumed to be independent and identically distributed. The occurrence of hydrologic drought event is observed beginning at time 0, one arbitrary point on the time axis, and illustrated in Figure 6.1. The interarrival time of drought events is defined as the period of time between two successive drought events (L_i), namely, the time elapsing from the initiation of a drought to the beginning of the next drought. In Figure 6.1, hydrologic drought events with magnitude equal to or greater than any arbitrary value (usually critical drought) M_0 , $M \geq M_0$, are denoted by \otimes . It is assumed that there occurs a hydrologic event with $M \geq M_0$ at any time.

Once a hydrologic drought event with $M \geq M_0$ occurs, let N be the number of hydrologic drought events until the occurrence of the next hydrologic drought event with $M \geq M_0$. Let W denote the interarrival time between two successive hydrologic drought events with $M \geq M_0$. Then

$$W = \sum_{i=1}^N L_i \quad (6.2)$$

where L_i is the drought interarrival time between any two successive drought events.

According to the above definition, the return period T of a hydrologic drought event with $M \geq M_0$ is the expected value of W . If the drought interarrival times, L_i are assumed to have the same distribution, then

$$T = EW = E\left(\sum_{i=1}^N L_i\right) = E(N)E(L_i). \quad (6.3)$$

In this analysis, the return period can be determined by Monte Carlo simulation. Recall that the return period is defined as the expected time between successive occurrences of a drought greater or equal to a specified critical drought, M_0 . Start the simulation and continue until two successive droughts occur with magnitude greater or equal to the critical drought. Denote the time between these two droughts (interarrival time) by W_1 . Repeat the simulation n times, to obtain a random sequence W_1, W_2, \dots, W_n . Then, the estimate of the mean return period is $T = n^{-1} \sum W_i$.

6.4 Example: Hydrologic Drought Analysis for the Fraser River

6.4.1 Modeling of PARMA₅₂(1, 1) Model Residuals

We present a detailed analysis of weekly river flow series for drought analysis. As observed in Chapter 5, the PARMA₅₂(1, 1) model parameters are dominated by few Fourier coefficients, and are obtained from

$$\begin{aligned}\hat{\theta}_i &= 0.255 - 0.197\cos\left(\frac{\pi i}{26}\right) + 0.121\cos\left(\frac{2\pi i}{26}\right) \\ \hat{\phi}_i &= 0.732\end{aligned}\tag{6.4}$$

and see Table 6.1 for other parameters such as $\hat{\sigma}_t$ and $\hat{\mu}_t$. The standardized residuals for $\text{PARMA}_\nu(1, 1)$ are computed using (3.62). Inspection of the ACF and PACF of these residuals indicate that they are iid (see Figure 6.2).

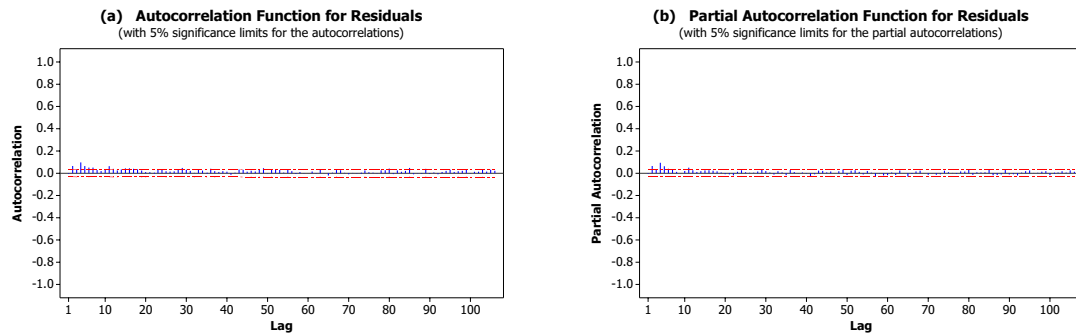


Figure 6.2: (a) ACF and (b) PACF for $\text{PARMA}_{52}(1, 1)$ model residuals of the Fraser River at Hope, BC.

A summary (histogram) of the residuals as well as probability plot for a three parameter lognormal distribution are shown in Figure 6.3, indicating that lognormal (scale = 0.094, location = 2.561 and threshold = -13.01) gives adequate fit except at the tails. This suggest for the use of a mixture of three parameter lognormal with appropriate tail distributions.

Hill's estimator (4.12) gives $\hat{\alpha} = 2.512$ and $\hat{C} = 0.379$ for the upper tail ($r = 215$ residuals), and $\hat{\alpha} = 3.518$ and $\hat{C} = 0.252$ for the lower tail ($r = 207$ residuals). Using (3.70), the p -values are 0.001 and 0.588, respectively. The small p -value as well as the visual evidence of the probability plot in Figure 6.4 suggests that the Pareto model is not a good fit for the upper tail of the estimated residuals. Therefore, we decided to fit a truncated Pareto to roughly the upper 5% of the residuals. Since $p > 0.05$, Pareto is a good fit for the

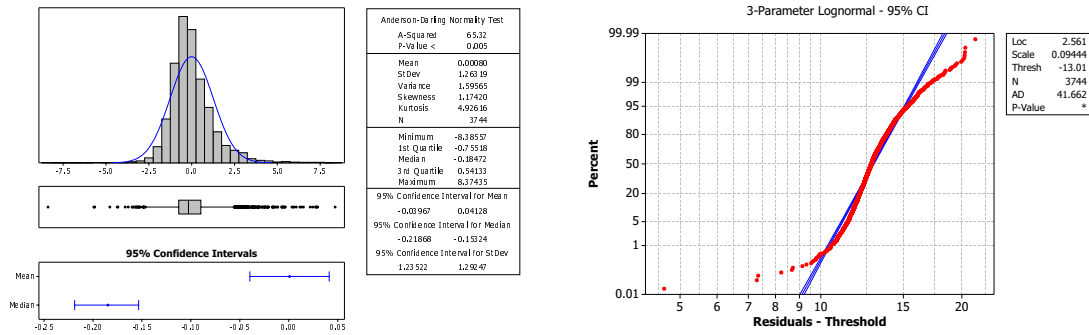


Figure 6.3: Statistical summary (left) and lognormal probability plot (right) for $\text{PARMA}_{52}(1,1)$ model residuals of the weekly Fraser River data, after applying significant Fourier coefficients. Compare Figure 5.5 (right).

lower tail. However, we chose a truncated Pareto for the lower 10% of the residuals because river flows are bounded from below. Then we computed $-1.537 =$ the 10th percentile and $2.115 =$ the 95th percentile of the three parameter lognormal distribution we fit to the body of the residuals. Next we determined that $r = 215$ residuals exceed the 95th percentile, and $r = 207$ residuals fall below the 10th percentile. Then the MLE from Theorem 3.13 was used to estimate the parameters ($\hat{\beta} = 8.374$, $\hat{\gamma} = 0.547$, $\hat{\alpha} = 2.068$) of the best fitting truncated Pareto distribution, and the theoretical distribution tail $P(R > r)$ was plotted over the 215 largest positive residuals in Figure 6.4 (left). In Figure 6.4 (right), we used the same method to fit a truncated Pareto ($\hat{\beta} = 8.386$, $\hat{\gamma} = 0.667$, $\hat{\alpha} = 3.459$) to the 207 largest negative residuals, after a change of sign. Both of the plots in Figures 6.4 indicate an adequate fit.

A mixture distribution with lognormal body and truncated Pareto tails was used to simulate the innovations. The mixture has cumulative distribution function (cdf)

$$P(\delta \leq r) = \begin{cases} F_-(r) & \text{if } r < -1.537 \\ F_0(r) & \text{if } -1.537 \leq r \leq 2.115 \\ F_+(r) & \text{if } r > 2.115 \end{cases} \quad (6.5)$$

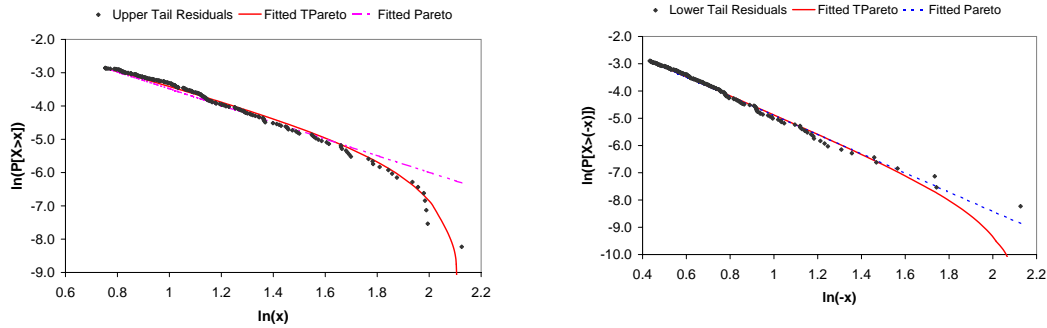


Figure 6.4: Log-log plot of upper (left) and lower (right) residual tails and fitted Pareto and truncated Pareto distributions, Fraser river at Hope, BC.

where F_0 is the cdf of the lognormal, and F_+, F_- are truncated Pareto cdfs of the positive and negative tails, respectively. The truncated Pareto distributions were slightly shifted (by $s = -0.141$ on the positive tail and $s = 0.240$ on the negative tail) to make the mixture cdf continuous. Now innovations could be simulated by the inverse cumulative distribution function method $\delta = F^{-1}(U)$ where U is a pseudorandom number uniformly distributed on the unit interval $(0, 1)$. However, this is impractical in the present case since the lognormal cdf is not analytically invertible. Instead, we used the Box-Müller method to generate standard normal random variates Z (see Gentle, 2003). Then lognormal random variates were calculated using $\delta = 2.561 + \exp(2.561 + 0.094Z)$. If $R > 2.115$, the 95th percentile of the lognormal, we generated another uniform $(0, 1)$ random variate U and substituted $\delta = F_+^{-1}(0.95 + 0.05U)$. If $R < -1.537$, the 10th percentile of the lognormal, we substituted $\delta = F_-^{-1}(0.1U)$. This gives simulated innovations δ with the mixture distribution (6.5).

Figure 6.5 shows a probability plot for $N = N_y\nu$ simulated innovations (for $\nu = 52$ weeks and $N_y = 100$ years) from the mixture distribution (6.5). Comparison with Figure 6.3 (right) shows that the simulated innovations are statistically identical to the computed model residuals in terms of distribution. Substituting the simulated innovations into the

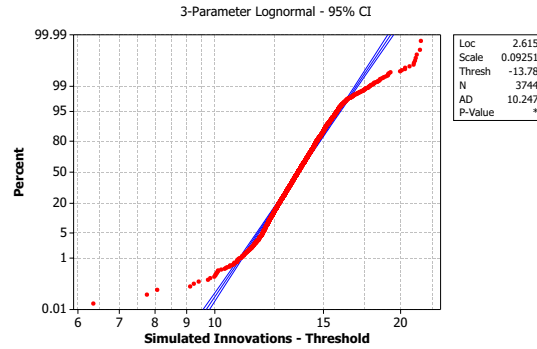


Figure 6.5: Probability plot of simulated innovations using the mixed three parameter lognormal and truncated Pareto distributions. Compare Figure 6.3 (right).

model (3.63) generates N_y years of simulated river flow. It is advantageous to simulate several extra years of river flows and throw out the initial years (100 years in this case), since we did not simulate X_t for $t < 0$. This ensures that the simulated series is periodically stationary (see Figure 6.6). Figure 6.7 shows the main statistical characteristics (mean, standard deviation and autocorrelations) of a typical synthetic river flow time series obtained by this method, as well as the same statistical measures for the observed time series. It is apparent that $\text{PARMA}_{52}(1, 1)$ model closely reproduces the main statistical characteristics, indicating that the model can be used for generating synthetic river flows in drought analysis of the Fraser River.

6.4.2 Drought Statistics Estimation

The main purpose of this section is to illustrate the proposed methodology for estimating drought properties based on simulated weekly river flow series of the Fraser River at Hope for a given water demand. Just as there are criteria for defining flood levels and event, criteria for defining drought years must also be developed. A more useful approach would be to define a drought year by a minimum flow required to meet (a) all water right priorities or

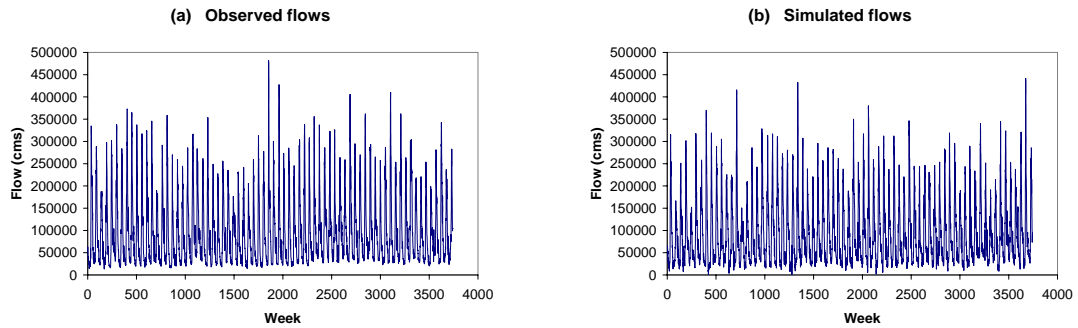


Figure 6.6: Plot of (a) observed and (b) simulated weekly river flows for the Fraser River at Hope, BC, indicating similarity.

(b) just municipal water requirements, and by how many years would some portion of that flow would not be met. In general, water demand is high during summer months and low in winter months. Therefore, we assumed a hypothetical sinusoidal demand curve given by

$$D_i = 34300 + 8000\cos\left(\frac{2\pi i}{52}\right) - 191000\sin\left(\frac{2\pi i}{52}\right) \quad (6.6)$$

Using $\text{PARMA}_{52}(1, 1)$ model, 1000 years and 30 realizations of weekly synthetic flow sequences for the Fraser River at Hope are generated. These synthetic flows were analyzed to ensure that the historical statistics were reproduced, and then equations (6.1) and (6.3) were applied to derive the necessary drought statistics (for example, maximum duration, severity and return periods for different values of L_0 and M_0) for the assumed water demand (6.6). The resulting droughts (a total of 30 for each drought property) were averaged and summarized in Table 6.2. The historical data were also analyzed to determine the drought statistics (see Table 6.2), indicating few drought events (which are not useful in obtaining drought events of large return periods, say 100 years). However, we were able to determine the return periods of drought duration and severity up to 385 years and 333 years, respectively, from the generated data set.

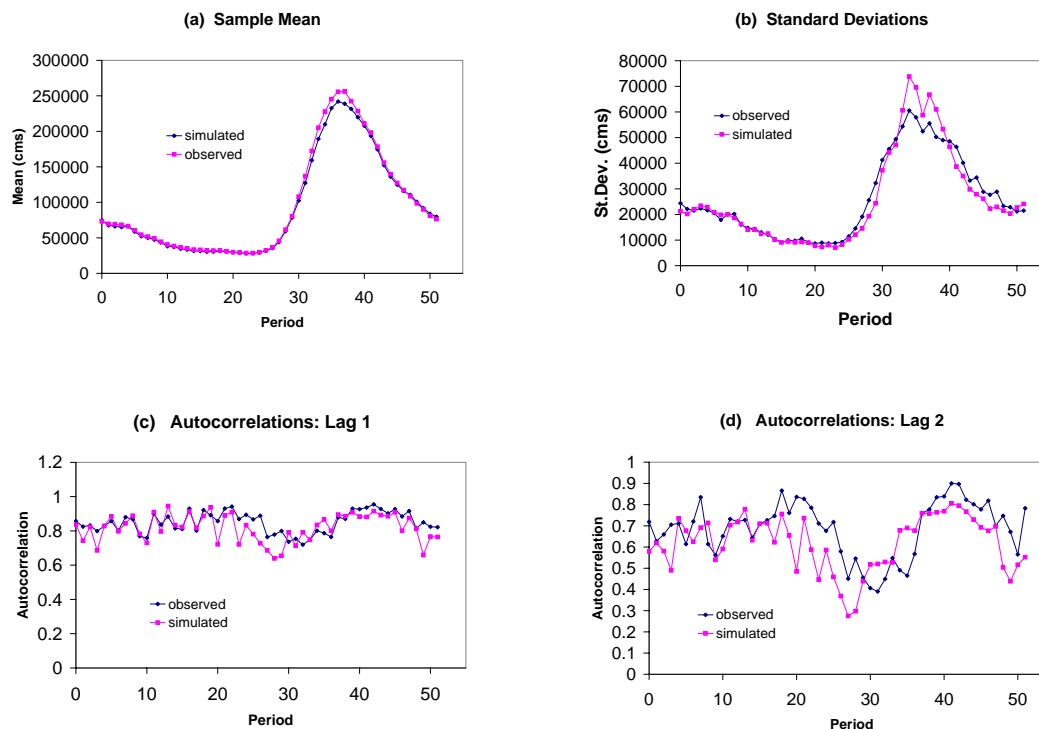


Figure 6.7: Comparison of mean, standard deviation, and autocorrelations for simulated vs. observed weekly river flow data for the Fraser River at Hope, BC.

It is observed that in 22 realizations a drought with the largest severity is different from the drought with the longest drought duration. However, in the remaining 8 realization, a drought with the worst severity is the same as the drought with the longest duration. It must be noted the result depends on the assumed demand curve and number of years of simulation.

In summary, a reliable analysis of the drought statistics based on historical data series can not be properly carried out due to the limited number of drought events in historical data series. It is essential for the water resource planner to use stochastic models such as PARMA models to generate long records and obtain more drought events for statistical analysis of droughts.

Table 6.1: Other parameter estimates for PARMA₅₂(1, 1) model of average weekly flow series for the Fraser River at Hope, BC.

Parameter	Season ν							
	0	1	2	3	4	5	6	7
$\hat{\sigma}_t$	11341.4	9365.7	10674.6	11173.0	11442.3	10321.3	7949.6	9161.4
$\hat{\mu}_t$	73139.4	69715.0	69133.0	68332.7	66348.2	60664.8	54528.4	51692.9
	8	9	10	11	12	13	14	15
$\hat{\sigma}_t$	8064.8	6196.3	8288.3	7374.0	4567.7	5751.0	3381.8	4368.1
$\hat{\mu}_t$	49291.5	44493.0	40590.5	38387.8	36543.0	35021.9	33342.0	32970.1
	16	17	18	19	20	21	22	23
$\hat{\sigma}_t$	4699.0	3014.5	4614.0	3065.2	2229.2	3958.7	2543.7	2073.1
$\hat{\mu}_t$	32180.6	31958.4	32220.3	30982.1	29701.5	29361.2	28435.7	28440.0
	24	25	26	27	28	29	30	31
$\hat{\sigma}_t$	3805.1	4267.0	5566.1	7582.3	11230.3	17216.4	19467.5	23820.6
$\hat{\mu}_t$	29564.3	32309.3	36677.3	45680.3	61496.0	80419.7	107893.3	136771.3
	32	33	34	35	36	37	38	39
$\hat{\sigma}_t$	28120.6	29511.4	35410.9	29142.7	26127.0	29320.7	18830.2	20127.7
$\hat{\mu}_t$	172356.4	205165.0	227797.0	245244.2	255747.8	256049.1	242545.2	228583.8
	40	41	42	43	44	45	46	47
$\hat{\sigma}_t$	15124.9	14232.0	11492.2	9000.8	10763.7	10741.4	8428.7	11202.4
$\hat{\mu}_t$	211048.4	198195.4	178312.8	155861.5	139551.5	127168.7	117237.5	108533.9
	48	49	50	51				
$\hat{\sigma}_t$	8588.8	10147.6	9350.6	9845.6				
$\hat{\mu}_t$	98497.6	90001.4	81269.0	76323.0				

Table 6.2: Drought statistics estimation of the historical and simulated data for the Fraser River at Hope, BC. Note that the drought duration (L), severity (D) and return period (T) are in weeks, in cms and in years, respectively.

Statistics	Generated				Historical			
Max. Duration	13.3				9			
Max. Severity*	167729				43331			
Return Period	L_0	T	D_0^*	T	L_0	T	D_0^*	T
	4	6.2	20,000	5.1	4	6.0	20,000	4.0
	6	14.3	40,000	8.6			40,000	12.0
	8	56.5	60,000	26.5				
	10	197.4	80,000	67.6				
	12	385.4	100,000	132.5				
			120,000	211.0				
			140,000	333.2				

* The values can be multiplied by 604800 sec. so as to obtain Severity in (m^3)

7 Summary and Conclusions

Generation of synthetic river flow data is important in planning, design and operation of water resources systems. River flow series usually exhibit both heavy tails and periodical stationarity; that is, their mean and covariance functions are periodic with respect to time. The common procedure in modeling such periodic river flow series is first to standardize or filter the series and then fit an appropriate stationary stochastic model to the reduced series. However, standardizing or filtering most river flow series may not yield stationary residuals due to periodic autocorrelations. Periodic autoregressive moving average (PARMA) models provide a powerful tool for the modeling of periodic hydrologic series in general and river flow series in particular. PARMA models are extensions of commonly used ARMA models that allow parameters to depend on season. The PARMA modeling procedure involves iterative steps of model identification, parameter estimation, model diagnosis and fitting the residuals (noise) with a probability distribution function (pdf). The opposite process to step-by-step modeling is the use of models to generate (simulate) new samples or a long sample of the process. One starts with the random noise and its pdf by generating its sample(s). Then generate the corresponding data samples by using the fitted PARMA model.

The main purposes of this dissertation were to fit a PARMA model to represent a given river flow data, estimate parameters, check for goodness of fit to the data, model the residuals, and to use the fitted model for generating synthetic river flows and apply them to extreme analysis such as droughts. The innovations algorithm is used to obtain estimates of the PARMA model parameters. Asymptotic distributions for the innovations estimates and model parameters (which are developed here) provide us with a general technique for identifying PARMA models with the minimum number of parameters.

First the innovations algorithm estimation procedure, as well as model identification

using simulated data from different Gaussian $\text{PARMA}_\nu(p, q)$ models, is discussed in detail. A simulation study demonstrates that the innovations algorithm and asymptotic distributions are an efficient and reliable technique for parameter estimation and model identification of PARMA models, respectively. However, such techniques often do not adequately identify the order of the model when the model is mixed in the sense that both the autoregressive (AR) and moving average (MA) components are present. In this case, the model identification technique can be supplemented by modeler experience or the Akaike's information criterion (AIC). We used monthly river flow data for the Fraser River in British Columbia in order to illustrate the model identification procedure and residual modeling procedure, and to prove the ability to generate realistic synthetic river flows. This example illustrates that the innovations algorithm is useful for modeling river flow time series. For monthly average river flow data for the Fraser River, a first order periodic autoregressive moving average model is adequate to capture the essential features. A mixture of three parameter lognormal body and truncated Pareto tails fits the model residuals nicely. This mixture model is then applied with satisfactory results to generate synthetic monthly river flow records.

Secondly, we applied the same methods to monthly flow data for the Salt River in Arizona, which has heavy tails. In this case, we are able to generate realistic synthetic river flows by developing a mixture probability model (stable with truncated Pareto tails) for the residuals. The methodology presented herein provides a useful tool for river flow modeling and synthetic river flow data generation. The results allow practitioners and planners to explore realistic decision-making scenarios for a given water resource system.

Thirdly, for analysis and design of water resources systems, it is sometimes required to generate river flow data with high resolution (that is, weekly or daily values). Fitting the PARMA model to historical weekly or daily data, however, requires estimation of too

many parameters which violates the principle of parsimony (model with minimum number of parameters). In an effort to obtain a parsimonious model representing these periodically stationary series, we develop the asymptotic distribution of the discrete Fourier transform of the innovation estimates and then determine those statistically significant Fourier coefficients, which are directly applicable to the selection of significant Fourier coefficients of the parameters of pure $\text{PMA}_\nu(q)$ and $\text{PAR}_\nu(1)$ processes. For higher order PARMA models, it is more difficult to obtain an explicit representation for the Fourier transform of the model parameters, and hence the selection of significant Fourier coefficients is a complicated problem. However, we developed the asymptotic distribution of the Fourier coefficients of the model parameters for $\text{PARMA}_\nu(1, 1)$. The simulation study demonstrates that these asymptotic results are efficient and reliable technique for identification of significant Fourier coefficients of PARMA model parameters. We demonstrated its application for modeling monthly and weekly Fraser River flow time series. It seems that even for monthly time series the parameters of $\text{PARMA}_{12}(1, 1)$ indeed vary smoothly with respect to t (dominated by a few Fourier coefficients). For weekly data series, we also found that $\text{PARMA}_{52}(1, 1)$ is adequate to capture the essential features. It is true that the $\text{PARMA}_{52}(1, 1)$ model parameters are dominated by few Fourier coefficients. This analysis can be easily extended to historical daily or weekly time series.

Finally, obtaining drought properties is important for planning and management of water resources system. For example, hydrologic drought properties (severity and duration) of various return periods are needed to assess the degree to which a water supply system will be able to cope with future droughts and, accordingly, to plan alternative water supply strategies. They can be determined from the historical record alone by using non-parametric methods but, because the number of drought events that can be drawn from the historical sample is

generally small, the “historical” drought properties have a large degree of uncertainty. Other alternatives for finding drought properties include using stochastic models (such as PARMA models) that can represent the underlying river flows, simulating long records of such river flows, and then deriving droughts properties from the simulated samples based on the theory of runs. In this study, we used lower order PARMA models such as $\text{PARMA}_{52}(1, 1)$ and illustrated the methodology for assessing drought properties. It seems that a $\text{PARMA}_{52}(1, 1)$ model is a good fit for the weekly Fraser River flows. Therefore, the $\text{PARMA}_{52}(1, 1)$ model was used to generate a number of synthetic flow sequences of 1000 years. These synthetic flows were analyzed to ensure that the historical statistics were reproduced, and then equations (6.1) and (6.3) were applied to derive the necessary drought statistics for the Fraser River at Hope. We were able to determine the return periods of drought duration and severity up to 385 years and 333 years, respectively, from the generated data set. We found that the methodology is useful for drought analysis in rivers where the high resolution data series (such as weekly) are described by low order PARMA models. In some cases, however, low order PARMA models may not be good fit to high resolution data series (e.g. weekly flows for the Truckee River at Farad) due to long term dependence of the weekly data series (long memory property of time series). Fractionally differenced low order PARMA models can be useful for river flows with such long memory property. This procedure can also be extended to flood frequency analysis.

The techniques presented in this research allow hydrologists, practitioners and planners to explore realistic decision-making scenarios for a given water resource system. As part of future work, comparison of this new method with the traditional approach in water resources will be made; a guideline on how (and when) to use the techniques effectively to water resources application, will be prepared.

Directions for future research include the development of asymptotic distributions for the parameters of higher order PARMA models, sensitivity analysis, more general discrete Fourier analysis for PARMA models and fractionally integrated PARMA models. Demonstration of the PARMA modeling procedure for other applications such as reservoir operation, flood analysis and river flow forecasting will also be part of future work.

As for drought analysis, the focus was on the illustration the proposed methodology for estimating drought properties based on simulated weekly river flow of the Fraser River. In the future, detailed investigation of the empirical probability distribution of the drought properties (duration and severity) both for the historical and simulated data, the goodness of fit as well as the sensitivity of the drought properties to the number of years and number of realizations will be carried out. As another case example, a detailed drought study of the Truckee River (which the main source of water supply in Truckee Meadows) will be conducted. It involves the determination of the prolonged drought characteristics for the river as well as the evaluation of drought impact on the water supply system for both present and future conditions.

References

- [1] Aban, I.B, M.M. Meerschaert and A.K. Panorska (2004), Parameter estimation for the truncated Pareto distribution. (*to appear, Journal of the American Statistical Association: Theory and Methods*).
- [2] Akaike, H. (1974), A new look at the statistical model identification. *IEEE Transaction on Automatic Control AC-19*(6), 716–723.
- [3] Adams, G.J and G.C. Goodwin, (1995), Parameter estimation for periodically ARMA models. *Journal of Time Series Analysis*, 16, 127–145.
- [4] Anderson, P.L., M.M. Meerschaert and A.V. Vecchia (1999), Innovations algorithm for periodically stationary time series. *Stochastic Processes and their Applications*, 83, 149–169.
- [5] Anderson, P.L. (1989), Asymptotic results and identification for cyclostationary time series. PhD. Dissertation, Colorado School of Mines, Golden, Colorado.
- [6] Anderson, P.L. and A.V. Vecchia (1993), Asymptotic results for periodic autoregressive moving -average processes. *Journal of Time Series Analysis*, 14, 1–18.
- [7] Anderson, P.L. and M.M. Meerschaert (1997), Periodic moving averages of random variables with regularly varying tails. *Annals of Statistics*, 25, 771–785.
- [8] Anderson, P.L. and M.M. Meerschaert (1998), Modeling river flows with heavy tails. *Water Resources Research*, 34(9), 2271–2280.
- [9] Anderson, P.L. and M.M. Meerschaert (2003), Parameter Estimates for Periodically Stationary Time Series. *Journal of Time Series Analysis*, to appear.

- [10] Ansley, C.F (1979), An algorithm for the exact likelihood of a mixed autoregressive moving average process. *Biometrika*, 66, 59–65.
- [11] Barnes, F.B. (1954), Storage required for a city water supply. *J. Inst. Eng., Australia*, 26, 198–203.
- [12] Bentarzi, M. and M. Hallin (1993), On the invertibility of periodic moving-average models. *Journal of Time Series Analysis*, 15, 263–268.
- [13] Bhansali, R. (1993), Estimation of the impulse response coefficients of a linear process with infinite variance. *J. Multivariate Anal.*, 45, 274–290.
- [14] Bianchi, K.L and M.M. Meerschaert (2004), Scale and shift invariant estimators for the heavy tail index alpha. *Journal of Statistical planning and inference*, submitted.
- [15] Bianchi, K.L (2003), Shift and scale invariant estimators of heavy tail index alpha. MS Thesis, University of Nevada, Reno.
- [16] Box, G.E., and G.M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.
- [17] Bras, R.L. (1990), *Hydrology: An Introduction to Hydrologic Science*, Addison-Wesley-Longman, Reading, Mass.
- [18] Brockwell, P.J. and R.A. Davis (1991), *Time Series: Theory and Methods*, 2nd ed. New York: Springer-Verlag.
- [19] Brockwell, P.J. and R.A. Davis (1996), *Introduction to Time Series and Forecasting*. New York: Springer-Verlag.

- [20] Burroughs, S.M, and S.F. Tebbens (2002), The upper-truncated power law applied to earthquake cumulative frequency-magnitude distributions. *Bulletin of the Seismological Society of America*, **92**, No. 8, 2983–2993.
- [21] Burroughs, S.M, and S.F. Tebbens (2001), Upper-truncated power law distributions. *Fractals*, **9**, 209–222.
- [22] Burroughs, S.M, and S.F. Tebbens (2001), Upper-truncated power laws in natural systems. *Journal of Pure and Applied Geophysics*, **158**, 741–757.
- [23] Cancelliere, A. and J.D. Salas (2004), Drought length properties for periodic-stochastic hydrologic data. *Water Resources Research*, *40*, W0250, 1–13.
- [24] Chen, H.-L. and A.R. Rao (2002), Testing hydrologic time series for stationarity. *Journal of Hydrologic Engineering*, *7*(2), 129–136.
- [25] Chung, C. and J.D. Salas (2000), Drought occurrence probabilities and risks of dependent hydrologic processes. *Journal of Hydrologic Engineering*, *5*(3), 259–268.
- [26] Davis, R. and S. Resnick (1985a), Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability*, *13*, 179–195.
- [27] Davis, R. and S. Resnick (1985b), More limit theory for the sample correlation function of moving averages. *Stochastic Processes and Their Applications*, *20*, 257–279.
- [28] Davis, R. and S. Resnick (1986), Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics*, *14*, 533–558.
- [29] Dekkers, A., J. Einmahl and L. de Haan (1989), A moment estimator for the index of an extreme value distribution. *Annals of Statistics*, *17*, 1833–1855.

- [30] Delleur, J.W., P.C. Tao and M.L. Kavvas (1976), an evaluation of the practicality and complexity of some rainfall and runoff times series models. *Water Resources Research*, 12, 953–970.
- [31] Douglas, E.M., R.M. Vogel, and C.N. Kroll (2002), Impact of streamflow persistence on hydrologic design. *J. Hydrol. Eng.*, 7(3), 220–227.
- [32] Downer, R., M. Siddiqui, and V. Yevjevich (1967), Applications of runs to hydrologic drought, paper presented at the International Hydrology Symposium, Colorado State University, Fort Collins, Colorado.
- [33] Dracup, J.A, K.S. Lee and E.G. Paulson (1980a), On the statistical characteristics of drought events. *Water Resources Research*, 16(2), 289–296.
- [34] Dracup, J.A, K.S. Lee and E.G. Paulson (1980b), On the definition of droughts. *Water Resources Research*, 16(2), 297–302.
- [35] Fama, E. (1965), The behavior of stock market prices. *Journal of Business*, 38, 34–105.
- [36] Feller, W. (1971), *An Introduction to Probability Theory and Its Applications*, vol. 2, 2nd ed., John Wiley, New York.
- [37] Fernandez, B., and J.D. Salas (1999a), Return period and risk of hydrologic events. I: Mathematical Formulation. *J. Hydrol. Eng.*, 4(4), 297–307.
- [38] Fernandez, B., and J.D. Salas (1999b), Return period and risk of hydrologic events. I: Applications. *J. Hydrol. Eng.*, 4(4), 308–316.
- [39] Fofack, H. and J.P. Nolan (1999), Tail behavior, modes and other characteristics of stable distributions. *Extremes*, 2, (1), 39–58.

- [40] Frick, D.M., D. Bode, and J.D. Salas (1990), Effect of drought on urban water supplies. Drought analysis. *J. Hydrol. Eng.*, 116, (6), 733–753.
- [41] Gentle, J.E. (2003), *Random Number Generation and Monte Carlo Methods*, 2nd ed. New York: Springer-Verlag.
- [42] Gladyšev, E.G. (1961), Periodically correlated random sequences, *Sov. Math.*, 2(2), 385–388.
- [43] Granger, C.W.J. and Z. Ding (1996), Varieties of long memory models. *Journal of Econometrics*, 73, 61–77.
- [44] Granger, C.W.J. and R. Joyeux (1980), An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–20.
- [45] Gueven, O. (1983), A simplified semiempirical approach to probabilities of extreme hydrologic droughts. *Water Resources Research*, 19(2), 441–453.
- [46] Hall, P. (1982), On some simple estimates of an exponent of regular variation. *J.R. Stat. Soc., Ser. B.*, 44, 37–42.
- [47] Hazen, A. (1914), Storage to be provided in impounding reservoirs for municipal water supply. *Trans. Amer. Soc. Civil Eng.*, 77, 1539–1669.
- [48] Hill, B.M. (1975), A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3(5), 1163–1173.
- [49] Hosking, J.R.M., J.R. Wallis and E.F. Wood (1985), Estimation of the general extreme-value distribution by the probability weighted moments. *Technometrics*, 27(3), 251–261.

- [50] Hosking, J.R.M. and J.R.Wallis (1987), Parameter and quantile estimation of the generalized Pareto distribution. *Technometrics*, 29(3), 339–349.
- [51] Hosking, J.R.M. (1996), Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long-memory time series. *Journal of Econometrics*, 73, 261–284.
- [52] Hurst, H.E. (1951), Long term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116, 770–799.
- [53] Jackson, B.B. (1975), The use of streamflow models in planning. *Water Resources Bulletin*, 11, 54–63.
- [54] Janicki, A. and A. Weron (1994a), Can one see α -stable variables and processes? *Statist. Sci.*, 9, 109–126.
- [55] Janicki, A. and A. Weron (1994b), *Simulation and Chaotic Behavior of α -stable Stochastic Processes*. Dekker, New York.
- [56] Jansen, D. and C. De Vries (1991), On the frequency of large stock market returns: putting booms and busts into perspective. *Review of Economics and Statistics*, 239, 18–24.
- [57] Jones, R. H. and W. M. Brelsford (1967), Times series with periodic structure. *Biometrika*, 54, 403–408.
- [58] Kokoszka, P. and M. Taqqu (1994), Infinite variance stable ARMA processes. *Journal of Time Series Analysis*, 15, 203–220.
- [59] Kokoszka, P. and M. Taqqu (1995), Fractional ARIMA with stable innovations. *Stochastic Processes and their Applications*, 60, 19–47.

- [60] Kokoszka, P. (1996), Prediction of infinite variance fractional ARIMA. *Probab. Math. Statist.*, 16, 65–83.
- [61] Kokoszka, P. and M. Taqqu (1996a), Infinite variance stable moving averages with long memory. *Journal of Econometrics*, 73, 79–99.
- [62] Kokoszka, P. and M. Taqqu (1996b), Parameter estimation for infinite variance fractional ARIMA. *Annals of Statistics*, 24, 1880–1913.
- [63] Kratz, M. and S. Resnick (1996), The qq estimator and heavy tails. *Stochastic Models*, 12, 699–724.
- [64] Lévy, P. (1925), *Calcul des Probabilités*. Gauthier Villars, Paris.
- [65] Li, W.K. (1988), An algorithm for the exact likelihood of periodic moving average models. *Communications in statistics: Simulation and computation*, 17(4), 1483–1494.
- [66] Llamas, J., and M. Siddiqui (1969), Runs of precipitation series. *Hydrology paper 33*, Colorado State University, Fort Collins, Colorado.
- [67] Lloyd, E.H. (1970), Return period in the presence of persistence. *Journal of Hydrology*, 10(3), 202–215.
- [68] Loaiciga, M. and M.A. Mariño (1991), Recurrence interval of geophysical events. *J. Water Resour. Plann. Manage.*, 117(3), 367–382.
- [69] Loretan, M. and P. Phillips (1994), Testing the covariance stationarity of heavy tailed time series. *J. Empirical Finance*, 1, 211–248.
- [70] Loucks, D.P., J.R. Stedinger, and D.A. Haith (1981), *Water Resources Systems Planning and Analysis*. Prentice-Hall, Englewood Cliffs, N.J.

- [71] Lund, R. B. and I. V. Basawa (1999), Modeling and inference for periodically correlated time series. *Asymptotics, Nonparameterics, and Time series*(ed. S. Ghosh), 37–62. Marcel Dekker, New York.
- [72] Lund, R. B. and I. V. Basawa (2000), Recursive prediction and likelihood evaluation for periodic ARMA models. *Journal of Time Series Analysis*, 20(1), 75–93.
- [73] Maidment, D.R. (1993), *Handbook of Hydrology* . New York: McGraw-Hill.
- [74] Mandelbrot, B. (1963), The variation of certain speculative prices. *J. Business*, 36, 394–419.
- [75] Mathier, L., L. Perreault, and B. Bobée (1992), The use of geometric and gamma related distributions for frequency analysis of water deficit. *Stochastic Hydrol. hydraul.*, 6, 239–254.
- [76] McCulloch, J.H. (1997), Measuring tail thickness in order to estimate the stable index α : A critique. *Journal of Business and Economic Statistics*, 15, 74–81.
- [77] Meerschaert, M.M. and H.P. Scheffler (1998), A simple robust estimator for the thickness of heavy tails. *J. Stat. Plann. Inference*, 71, 19–34.
- [78] Meerschaert, M.M. and H.P. Scheffler (2001), *Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice*, Wiley Interscience, New York.
- [79] Mikosch, T., T. Gdrich, C. Klüppenberg and R. Adler (1995), Parameter estimation for ARMA models with infinite variance inovations. *Annals of Statistics*, 23, 305–326.
- [80] Mittnik, S., and S. Rachev (1995), *Modelling Finiancial Assets with Alternative Stable Models*. Wiley, New York.

- [81] Moss, M.E. and M.C. Bryson (1974), Autocorrelation structure of monthly streamflows. *Water Resources Research* 10(4),737–744.
- [82] Nikias, C. and M. Shao (1995), *Signal Processing with Alpha Stable Distributions and Applications*. Wiley, New York.
- [83] Noakes, D.J., A.L. McLeod and K. Hipel (1985), Forecasting monthly riverflow time series. *International Journal of Forecasting*, 1, 179–190.
- [84] Nolan, J.P. (1999), Fitting data and assessing goodness of fit with stable distribution. Department of Mathematics and Statistics, American University.
- [85] Pagano, M. (1978), On periodic and multiple autoregressions. *Annals of Statistics*, 6(6), 1310–1317.
- [86] Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (1996), *Numerical Recipes in Fortran 90*. New York : Cambridge University Press.
- [87] Quimpo, R.G. (1967), Stochastic modeling of daily river flow sequences. *Hydrology Paper* 18, Colorado State University, Fort Collins, Colorado.
- [88] Rao, A.R. and R.L. Kashyap (1974), Stochastic modelling of Riverflows. *IEEE Transaction on Automatic Control* AC-19(6), 874–881.
- [89] Resnick, S and C. Stărică (1995), Consistency of Hill’s estimator for dependent data. *Journal of applied Probability*, 32, 139–167.
- [90] Resnick, S. (1997), Heavy tail modeling and teletraffic data. *Annals of Statistics*, 25, 1805–1869.

- [91] Roesner, L.A. and V.M. Yevjevich (1966), Mathematical models for time series of monthly precipitation and monthly runoff. *Hydrology Paper 15*, Colorado State University, Fort Collins, Colorado.
- [92] Ross, S.M. (2000), *Introduction to Probability Models*, 7th ed. San Diego: Harcourt/Academic Press.
- [93] Rybin, A.K. (1978), Effectiveness and stability of estimates of the parameters of a strong signal in the case of non-Gaussian noises. *Engineering Cybernetics*, 16, 115–129.
- [94] Salas, J.D, J.W. Delleur, V. Yevjevich and W.L. Lane (1980), Applied Modeling of Hydrologic Time Series. *Water Resource Publications*, Littleton, Colorado.
- [95] Salas, J.D, J.T.B Obeysekera and R.A. Smith (1981), Identification of streamflow stochastic models. *ASCE Journal of the Hydraulics Division*, 107(7), 853–866.
- [96] Salas, J.D, D.C. Boes and R.A. Smith (1982), Estimation of ARMA models with seasonal parameters. *Water Resources Research*, 18, 1006–1010.
- [97] Salas, J.D, G.Q. Tabios III and P. Bartolini (1985), Approaches to multivariate modeling of water resources time series. *Water Resources Bulletin*, 21, 683–708.
- [98] Salas, J.D. and J.T.B Obeysekera (1992), Conceptual basis of seasonal streamflow time series models. *Journal of Hydraulic Engineering*, 118(8), 1186–1194.
- [99] Salas, J.D. (1993), *Analysis and modeling of hydrologic time series*. The Mc Graw-Hill Handbook of Hydrology, D.R. Maidment, ed., 19.1–19.71.
- [100] Samorodnitsky, G. and M. Taqqu (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman and Hall, London.

- [101] Sen, Z. (1976), Wet and dry periods of annual flows series. *Journal of Hydr. Div., ASCE*, 102(10), 1503–1514.
- [102] Sen, Z. (1977), Run-sums of annual flow series. *Journal of Hydrolog*, Amsterdam, 35, 311–324.
- [103] Sen, Z. (1978), A mathematical model of monthly flow sequences. *Hydrological Sciences Bulletin*, 23(2), 223–229.
- [104] Sen, Z. (1980a), Statistical analysis of hydrologic critical droughts. *Journal of Hydr. Div., ASCE*, 106(1), 99–115.
- [105] Sen, Z. (1980b), Regional drought and flood frequency analysis: Theoretical consideration. *Journal of Hydrology*, 46, 265–279.
- [106] Sen, Z. (1991), Probabilistic modeling of crossing in small samples and applications of runs to hydrology. *Journal of Hydrology*, 124, 345–362.
- [107] Shao, Q. and R.B. Lund (2004), Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models. *Journal of Time Series Analysis*, 25(3), 359–372.
- [108] Sharma, T. (1995), Estimation of drought severity on independent and dependent hydrologic series. *Water Resources Management*, 11, 35–39.
- [109] Shiau, J. and H.W. Shen (2001), Recurrence analysis of hydrologic droughts of different severity. *Journal Water Resources Planning and Management*, 127(1), 30–40.
- [110] Shumway, R.H and D.S. Stoffer (2000), *Time Series Analysis and Its Applications*. Springer-Verlag New York.

- [111] Stedinger, J.R. and M.R. Taylor (1982), Synthetic streamflow generation 1. Model verification and validation. *Water Resources Research*, 36, 1519–1533.
- [112] Stedinger, J.R., R.M. Vogel and E.F-Georgiou (1993), Frequency analysis of extreme events. The Mc Graw-Hill Handbook of Hydrology, D.R. Maidment, ed., 18.1–19.61.
- [113] Sudler, C.E. (1927), Storage required for regulation of streamflow. *Trans. Amer. Soc. Civil Eng.*, 91, 622–660.
- [114] Tao, P.C. and J.W. Delleur (1976), Seasonal and nonseasonal ARMA models. *ASCE Journal of the Hydraulic Division*, 102(HY10), 1541–1559.
- [115] Thomas, H.A. and Fiering, M.B. (1962), Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation. In: A. Mass et al. (Editors), Design of Water Resource Systems. Harvard University Press, Cambridge, MA.
- [116] Thompstone, R.M., K.W. Hipel and A.I. McLeod (1985), Forecasting quarter-monthly riverflow. *Water Resources Bulletin*, 25(5), 731–741.
- [117] Tiao, G.C. and M.R. Grupe (1980), Hidden periodic autoregressive moving average models in time series data. *Biometrika* 67, 365–373.
- [118] Tjøstheim, D. and J. Paulsen (1982), Empirical identification of multiple time series. *Journal of Time Series Analysis*, 3, 265–282.
- [119] Troutman, B.M. (1979), Some results in periodic autoregression. *Biometrika*, 6, 219–228.
- [120] Ula, T.A. (1990), Periodic covariance stationarity of multivariate periodic autoregressive moving average processes. *Water Resources Research*, 26(5), 855–861.

- [121] Ula, T.A. (1993), Forecasting of Multivariate periodic autoregressive moving average processes. *Journal of Time Series Analysis*, 14, 645–657.
- [122] Ula, T.A. and A.A Smadi (1997), Periodic stationarity conditions for periodic autoregressive moving average processes as eigenvalue problems. *Water Resources Research*, 33(8), 1929–1934.
- [123] Ula, T.A. and A.A Smadi (2003), Identification of periodic moving average models. *Communications in Statistics: Theory and Methods*, 32(12), 2465–2475.
- [124] Vecchia, A.V., J.T.B Obeysekera, J.D. Salas and D.C Boes (1983), Aggregation and estimation for low-order periodic ARMA models. *Water Resources Research*, 19(5), 1297–1306.
- [125] Vecchia, A.V. (1985a), Periodic autoregressive-moving average (PARMA) modelling with applications to water resources. *Water Resources Bulletin* 21, 721–730.
- [126] Vecchia, A.V. (1985b), Maximum likelihood estimation for periodic moving average models. *Technometrics*, 27(4), 375–384.
- [127] Vecchia, A.V. and R. Ballerini (1991), Testing for periodic autocorrelations in seasonal time series data. *Biometrika*, 78(1), 53–63.
- [128] Vogel, R.M. (1987), Reliability indices for water supply systems. *Journal Water Resources Planning and Management*, 113(4), 645–654.
- [129] Wei, X. (1995), Asymptotically efficient estimation of the index of regular variation. *Annals of Statistics*, 23, 2036–2058.
- [130] Werner, T. and C. Upper (2002), Time variation in the tail behaviour of Bund future returns. Economic Research Center of Deutsche Bundesbank, Germany.

- [131] Yevjevich, V.M. (1963), Fluctuations of wet and dry years. Part 1. Research data assembly and mathematical models. textitHydrology Paper 1, Colorado State University, Fort Collins, Colorado.
- [132] Yevjevich, V. (1967), An objective approach to definitions and investigations of continental droughts. Hydrology paper 23, Colorado State University, Fort Collins, Colorado.
- [133] Yevjevich, V.M. and R.I. Jeng (1969), Properties of non-homogeneous hydrologic time series. *Hydrology Paper 32*, Colorado State University, Fort Collins, Colorado.
- [134] Yevjevich, V.M. (1972), Structural analysis of hydrologic time series. *Hydrology Paper 56*, Colorado State University, Fort Collins, Colorado.
- [135] Yevjevich, V.M. (1993), General Introduction of stochastic hydrology in water resources. *NATO ASI Series, Series E, Applied Sciences*, Vol. 237, 3–23.
- [136] Zelenhastic, E., and A. Salvai (1987), A method of streamflow drought analysis. *Water Resources Research*, 23(1), 156–168.
- [137] Zelenhastic, E. (2002), On the extreme streamflow drought analysis. *Water Resources Management*, 16, 105–132.

A Basic Concepts

A-1 Definitions, Theorems, etc.

Time series is a set of observations X_t , each one being recorded at a specific time t . For example, X_t can be river flow (daily, monthly, etc.) quantities at time t . A time series model for the observed data $\{X_t\}$ is a specification of the joint distributions (or possible only the means and covariances) of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization. A mathematical model representing a stochastic process is called “Stochastic model” or “Time Series model”. It has a certain mathematical form or structure and a set of parameters.

A complete probabilistic time series model for the sequences of random variables $\{X_1, X_2, \dots\}$ would specify all of the joint distributions of the random vectors $(X_1, X_2, \dots, X_n)'$, $n = 1, 2, \dots$ or equivalently all of the probabilities

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} \quad -\infty < x_1, \dots, x_n < \infty, n = 1, 2, \dots \quad (\text{A-1})$$

Such specification is rarely used in time series analysis (unless the data are generated by some well-understood simple mechanism), since in general it will contain far too many parameters to be estimated from the available data. Instead we specify on the first- and second-order moments of the joint distributions, i.e., the expected values EX_t and the expected products $EX_{t+h}X_t$, $t = 1, 2, \dots$, $h = 1, 2, \dots$, focusing on the properties of the sequence $\{X_t\}$ which depends only on these. Such properties of $\{X_t\}$ are referred to as second-order properties. In the particular case when all the joint distributions are multivariate normal, the second-order properties of $\{X_t\}$ completely determine the joint distributions and hence give a complete probabilistic characterization of the sequence. In general we shall lose a certain amount of

information by looking at the time series “through second order spectacles”; however, the theory of minimum mean squared error linear prediction depends only on the second order properties, thus providing further justification for the use of the second order characterization of time series models.

If the joint distribution can be factored into the product of the marginal distributions as

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= P\{X_1 \leq x_1\} \cdot P\{X_2 \leq x_2\} \cdot \dots \cdot P\{X_n \leq x_n\} \\ &= F(x_1) \cdot F(x_2) \cdot \dots \cdot F(x_n) \end{aligned} \quad (\text{A-2})$$

then process becomes an independent stochastic process and the series is an independent series. Otherwise there is certain type of serial dependence among the variables and the process is called a serially dependent stochastic process and correspondingly a serially dependent time series.

Definition A.1 (*The Autocovariance Function*). If $\{X_t, t \in T\}$ is a process such that $\text{Var}(X_t) < \infty$ for each $t \in T$, then the autocovariance function $\gamma_X(\cdot, \cdot)$ of $\{X_t\}$ is defined by

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - EX_r)(X_s - EX_s)], \quad r, s \in T \quad (\text{A-3})$$

The autocovariance measures the linear dependence between two points on the same series observed at different times. Very smooth series exhibit autocovariance functions that stay large even when the r and s are far apart, whereas choppy series tend to have autocovariance functions that are nearly zero for large separations. Recall from classical statistics that if $\gamma_X(r, s) = 0$, X_r and X_s are not linearly related, but there still may be some dependence structure between them. If, however, X_r and X_s are bivariate normal, $\gamma_X(r, s) = 0$ ensure their independence.

Definition A.2 (*Stationarity*). The time series $\{X_t, t \in \mathbb{Z}\}$, with index set $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$, is said to be stationary if

1. $E|X_t|^2 < \infty$ for all $t \in \mathbb{Z}$
2. $EX_t = \mu$ for all $t \in \mathbb{Z}$ and
3. $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ for all $r, s, t \in \mathbb{Z}$

Remark A.3 *Stationarity as just defined is frequently referred to in the literature as weak stationarity, covariance stationarity, stationarity in a wider sense or second-order stationarity.*

Remark A.4 *If $\{X_t, t \in \mathbb{Z}\}$ is stationary then $\gamma_X(r, s) = \gamma_X(r - s, 0)$ for all $r, s \in \mathbb{Z}$. It is therefore convenient to redefine the autocovariance function of a stationary process as function of just one variable,*

$$\gamma_X(h) \equiv \gamma_X(h, 0) = \text{Cov}(X_{t+h}, X_t), \quad \text{for all } t, h \in \mathbb{Z} \quad (\text{A-4})$$

The function $\gamma_X(\cdot)$ is referred to as the autocovariance function of $\{X_t\}$ and $\gamma_X(h)$ as its value at “lag” h . The autocorrelation function (ACF) of $\{X_t\}$ is defined analogously as the function whose value at lag h is

$$\rho_X(h) \equiv \gamma_X(h)/\gamma_X(0) = \text{Corr}(X_{t+h}, X_t), \quad \text{for all } t, h \in \mathbb{Z} \quad (\text{A-5})$$

Remark A.5 *It will be noticed that we have defined stationarity only in the case when $T = \mathbb{Z}$. It is not difficult to define stationarity using a more general index set, but for our purposes this will not be necessary. If we wish to model a set of data $\{x_t, t \in T \subset \mathbb{Z}\}$ as a realization of a stationary process, we can always consider it to be part of the a realization of a stationary process $\{X_t, t \in \mathbb{Z}\}$.*

The autocorrelation function measure the linear predictability of the series at time t , say X_t , using the only values X_s . We can easily show that $-1 \leq \rho(s = t + h, t) \leq 1$ using

the Cauchy-Schwarz inequality. If we can predict X_t perfectly from X_s through a linear relationship, $X_t = \beta_0 + \beta_1 X_s$, then the correlation will be 1 when $\beta_1 > 0$, -1 when $\beta_1 < 0$, and 0 when $\beta_1 = 0$. Hence, we have a rough measure of the ability to forecast the series at time t from the value at time s .

Partial autocorrelation Function (PACF)

To define the PACF for mean-zero stationary time series, let X_h^{h-1} denote the best linear predictor of X_h based on the $\{X_1, X_2, \dots, X_{h-1}\}$. We note X_h^{h-1} has the form:

$$X_h^{h-1} = \beta_1 X_{h-1} + \beta_2 X_{h-2} + \dots + \beta_{h-1} X_1, \quad (\text{A-6})$$

where the β 's are chosen to minimize the mean square linear prediction error, $E(X_h - X_h^{h-1})^2$. In addition, let X_0^{h-1} denote the minimum mean square linear predictor of X_0 based on $\{X_1, X_2, \dots, X_{h-1}\}$. X_0^{h-1} can be written as

$$X_0^{h-1} = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{h-1} X_{h-1}. \quad (\text{A-7})$$

Equation (A-6) can be thought of as the linear regression of X_h on the past, $X_{h-1}, X_{h-2}, \dots, X_1$, and (A-7) can be thought of as the linear regression of X_0 on the past, X_1, X_2, \dots, X_{h-1} . The coefficients, X_h on the past, $\beta_1, \beta_2, \dots, \beta_{h-1}$ are the same as in (A-6) and (A-7), which means that, for stationary process, linear prediction backward in time is equivalent to linear prediction forward in time.

Formally, for a stationary time series, X_t , we define the partial autocorrelation function (PACF), ϕ_{hh} , $h = 1, 2, \dots$, by

$$\phi_{11} = \text{Corr}(X_1, X_0) \quad (\text{A-8})$$

and

$$\phi_{hh} = \text{Corr}(X_h - X_h^{h-1}, X_0 - X_0^{h-1}), \quad h \geq 2 \quad (\text{A-9})$$

Both $(X_h - X_h^{h-1})$ and $(X_0 - X_0^{h-1})$ are uncorrelated with $\{X_1, X_2, \dots, X_{h-1}\}$. By stationarity, the PACF, ϕ_{hh} , is the correlation between X_t and X_{t-h} with the linear effect of $\{X_{t-1}, X_{t-2}, \dots, X_{t-(h-1)}\}$, on each, removed. If the process X_t is Gaussian, then

$\phi_{hh} = \text{Corr}(X_t, X_{t-h} | X_{t-1}, X_{t-2}, \dots, X_{t-(h-1)})$. That is, ϕ_{hh} is the correlation coefficient between X_t and X_{t-h} in the bivariate distribution of (X_t, X_{t-h}) conditional on $\{X_{t-1}, X_{t-2}, \dots, X_{t-(h-1)}\}$.

Definition A.6 (*Strict Stationarity*). *The time series $\{X_t, t \in \mathbb{Z}\}$ is said to be strictly stationary if the joint distribution of $(X_{t_1}, \dots, X_{t_k})'$ and $(X_{t_1+h}, \dots, X_{t_k+h})'$ are the same for all positive integers k and for all $t_1, \dots, t_k, h \in \mathbb{Z}$*

Remark A.7 *Definition A.6 is equivalent to the statement that $(X_1, \dots, X_k)'$ and $(X_{1+h}, \dots, X_{k+h})'$ have the same joint distributions for all integers h and $k > 0$.*

If $\{X_t\}$ is strictly stationary it immediately follows, on taking $k = 1$ in Definition A.6, that X_t has the same distribution for each $t \in \mathbb{Z}$. If $E|X_t|^2 < \infty$ this implies in particular that EX_t and $\text{Var}(X_t)$ are both constant. Moreover, taking $k = 2$ in Definition A.6, we find that X_{t+h} and X_t have the same joint distribution and hence the same covariance for all $h \in \mathbb{Z}$. Thus a strictly stationary process with finite second moments is stationary.

The converse of the previous statement is not true. For example if X_t is a sequence of independent random variables such that X_t is exponentially distributed with mean one when t is odd and normally distributed with mean one and variance one when t is even, then X_t is stationary with $\gamma_X(0) = 1$ and $\gamma_X(h) = 0$ for $h \neq 0$. However since X_1 and X_2 have different distributions, and X_t can not be strictly stationary.

There is one important case however in which stationarity does imply strict stationarity.

Definition A.8 (*Gaussian Time Series*). The process X_t is a Gaussian time series if and only if the distribution functions of X_t are all multivariate normal.

If $\{X_t, t \in \mathbb{Z}\}$ is a stationary Gaussian process then $\{X_t\}$ is strictly stationary, since for all $n \in \{1, 2, \dots\}$ and for all $h, t_1, t_1, \dots \in \mathbb{Z}$, the random vectors $(X_{t_1}, \dots, X_{t_n})'$ and $(X_{t_1+h}, \dots, X_{t_n+h})'$ have the same mean and covariance matrix, and hence the same distribution.

Definition A.9 The time series X_t is a linear process if it has the representation

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \quad (\text{A-10})$$

for all t , where $\{Z_t\} \sim WN(0, \sigma^2)$ and $\{\psi_j\}$ is a sequence of constants with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Note that $WN(0, \sigma^2)$ is white noise (sequence of uncorrelated random variables) with mean 0 and variance σ^2 .

A-2 Statistical Properties of Time Series

A stationary process X_t is characterized, at least from second-order point of view, by its mean μ and its autocovariance function $\gamma(\cdot)$. The estimation of $\mu, \gamma(\cdot)$ and the autocorrelation function $\rho(\cdot) = \gamma(\cdot)/\gamma(0)$ from observations X_1, \dots, X_n therefore plays a crucial role in problem of inference and in particular in the problem of constructing an appropriate model for the data.

For a set of observations X_1, \dots, X_n , the sample mean, autocovariance and autocorrelation functions are defined by

$$\begin{aligned}
\hat{\mu} &= \bar{X}_n = n^{-1} \sum_{t=1}^n X_t \\
\hat{\gamma}(h) &= (n)^{-1} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n) \\
\hat{\rho}(h) &= \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}
\end{aligned} \tag{A-11}$$

Definition A.10 Let $\{X_n\}$ be a sequence of random variables, and X a random variable.

Then we say that X_n converges in probability to X if

$$\forall \epsilon > 0 \quad P(|X_n - X| > \epsilon) \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{A-12}$$

We denote convergence in probability by \xrightarrow{P} .

Definition A.11 Let $\{X_n\}$ be a sequence of random variables and X and random variable.

Then we say that X_n converges in distribution to X if

$$F_n(x) \rightarrow F(x) \text{ for all } x \text{ such that } F(x) \text{ is continuous, as } n \rightarrow \infty \tag{A-13}$$

where $F_n(x)$ is the distribution of X_n , and $F(x)$ is the distribution of X . We denote convergence in distribution by \Rightarrow .

Definition A.12 A sequence of normal random variables $\{X_n\}$ is said to be asymptotical normal with “mean” μ_n and “standard deviations” σ_n if $\sigma_n > 0$ for n sufficiently large and

$$\sigma_n^{-1}(X_n - \mu_n) \Rightarrow Z, \quad \text{where } Z \sim \mathcal{N}(0, 1). \tag{A-14}$$

We shall abbreviate this as

$$X_n \sim AN(\mu_n, \sigma_n^2) \tag{A-15}$$

where \sim will denote “is distributed as”.

Theorem A.13 *If X_t is a stationary linear process of the form (A-10) and $\sum_j \psi_j \neq 0$, then*

$$\bar{X}_n \sim \text{AN}(\mu_X, n^{-1}V), \quad (\text{A-16})$$

where

$$V = \sum_{h=-\infty}^{\infty} \gamma(h) = \sigma^2 \left(\sum_{j=-\infty}^{\infty} \psi_j \right)^2 \quad (\text{A-17})$$

and $\gamma(\cdot)$ is the autocovariance function of X_t

PROOF. See Theorem 1.5 of Shumway and Stoffer (2000) for a proof.

Theorem A.14 *If X_t is a stationary linear process of the form (A-10) satisfying the fourth moment condition (i.e., $\text{E}Z_t^4 = \eta\sigma^4 < \infty$ where η is some constant), then*

$$\begin{pmatrix} \hat{\gamma}(0) \\ \hat{\gamma}(1) \\ \vdots \\ \hat{\gamma}(K) \end{pmatrix} \sim \text{AN} \left[\begin{pmatrix} \gamma(0) \\ \gamma(1) \\ \vdots \\ \gamma(K) \end{pmatrix}, n^{-1}V \right] \quad (\text{A-18})$$

where V is the matrix with elements given by

$$v_{pq} = (\eta - 3)\sigma^4\gamma(p)\gamma(q) + \sum_{u=-\infty}^{\infty} [\gamma(u)\gamma(u-p+q) + \gamma(u+q)\gamma(u-p)] \quad (\text{A-19})$$

PROOF. See Theorem 1.6 of Shumway and Stoffer (2000) for a proof.

Theorem A.15 *If X_t is a stationary linear process of the form (A-10) satisfying the fourth moment condition (i.e., $\text{E}Z_t^4 = \eta\sigma^4 < \infty$ where η is some constant), then*

$$\begin{pmatrix} \hat{\rho}(0) \\ \hat{\rho}(1) \\ \vdots \\ \hat{\rho}(K) \end{pmatrix} \sim \text{AN} \left[\begin{pmatrix} \rho(0) \\ \rho(1) \\ \vdots \\ \rho(K) \end{pmatrix}, n^{-1}W \right] \quad (\text{A-20})$$

where W is the matrix with elements given by

$$w_{pq} = \sum_{u=1}^{\infty} [\rho(u+p) + \rho(u-p) - 2\rho(p)\rho(u)] \times [\rho(u+p) + \rho(u-q) - 2\rho(q)\rho(u)] \quad (\text{A-21})$$

PROOF. See Theorem 1.7 of Shumway and Stoffer (2000) for a proof.

A-3 Spectral Representation of a Stationary Process

Suppose that $\{X_t\}$ is a non-zero stationary time series with autocovariance function $\gamma(\cdot)$ satisfying that $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$. The spectral density of $\{X_t\}$ is the function $f(\cdot)$ defined by

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \text{Exp}(-ih\lambda)\gamma(h), \quad -\infty < \lambda < \infty \quad (\text{A-22})$$

where $\text{Exp}(-i\lambda) = \cos(\lambda) + i \sin(\lambda)$ and $i = \sqrt{-1}$.

Basic properties of f

1. f is even, i.e., $f(\lambda) = f(-\lambda)$
2. $f(\lambda) \geq 0$ for all $\lambda \in (-\pi, \pi]$ and
3. $\gamma(k) = \int_{-\pi}^{\pi} \text{Exp}(ik\lambda)f(\lambda)d\lambda = \int_{-\pi}^{\pi} \cos(k\lambda)f(\lambda)d\lambda$.

Definition A.16 A function f is the spectral density of a stationary time series $\{X_t\}$ with ACVF $\gamma(\cdot)$ if

1. $f(\lambda) \geq 0$ for all $\lambda \in (0, \pi]$
2. $\gamma(h) = \int_{-\pi}^{\pi} \text{Exp}(ih\lambda)f(\lambda)d\lambda$ for all integers h

Remark A.17 Spectral densities are essentially unique. That is, if f and g are two spectral densities corresponding the autocovariance function $\gamma(\cdot)$, i.e., $\gamma(h) = \int_{-\pi}^{\pi} \text{Exp}(ih\lambda)f(\lambda)d\lambda =$

$\int_{-\pi}^{\pi} \text{Exp}(ih\lambda)g(\lambda)d\lambda$ for all integers h , then f and g have the same Fourier coefficients and hence are equal.

Proposition A.18 *A real-valued function f defined on $(-\pi, \pi]$ is the spectral density of a stationary process if and only if*

1. $f(\lambda) = f(-\lambda)$
2. $f(\lambda) \geq 0$ and
3. $\int_{-\pi}^{\pi} f(\lambda)d\lambda < \infty$

PROOF. See Proposition 4.1.1 of Brockwell and Davis (1996) for a proof.

Corollary A.19 *An absolutely summable function $\gamma(\cdot)$ is the autocovariance function of a stationary time series if and only if it is even and*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \text{Exp}(-ih\lambda)\gamma(h), \quad -\infty < \lambda < \infty \quad (\text{A-23})$$

in which case $f(\cdot)$ is the spectral density of $\gamma(\cdot)$

PROOF. See Corollary 4.1.1 of Brockwell and Davis (1996) for a proof.

Theorem A.20 (*Spectral representation of the ACVF*) *A function $\gamma(\cdot)$ defined on the integers is the AVCF of a stationary time series if and only if there exists a right-continuous, nondecreasing, bounded function F on $[-\pi, \pi]$ with $F(-\pi) = 0$ such that*

$$\gamma(h) = \int_{(-\pi, \pi]} \text{Exp}(ih\lambda)dF(\lambda) \quad (\text{A-24})$$

for all integers h . (For real-valued time series, F is symmetric in the sense that $\int_{(a, b]} dF(x) = \int_{[-b, -a)} dF(x)$ for all a and b such that $0 < a < b$.)

PROOF. See Theorem 4.1.1 of Brockwell and Davis (1996) for a proof.

Remark A.21 *The function F is a generalized distribution function on $[-\pi, \pi]$ if $G(\lambda) = F(\lambda)/F(\pi)$ is a probability distribution function on $[-\pi, \pi]$. Note that since $F(\pi) = \gamma(0) = \text{Var}(X_1)$, the ACF of $\{X_t\}$ has spectral representation*

$$\rho(h) = \int_{(-\pi, \pi]} \text{Exp}(ih\lambda) dG(\lambda). \quad (\text{A-25})$$

The function F in (A-25) is called the spectral distribution function of $\gamma(\cdot)$. If $F(\lambda)$ can be expressed as $F(\lambda) = \int_{-\pi}^{\lambda} f(y) dy$ for all $\lambda \in [-\pi, \pi]$, then f is the spectral density function and the time series is said to have a continuous spectrum. If F is a discrete distribution (i.e., if G is a discrete probability distribution), then the time series is said to have a discrete spectrum.

A-4 Simulation of Continuous Random Variates

Generating Uniform Random Variates

Consider uniform continuous-values deviates X that lie in the range $[0, 1]$. The probability density function (pdf) of X is

$$f_X(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{elsewhere} \end{cases} \quad (\text{A-26})$$

The mean and standard deviation for this distribution are $\mu = 1/2, \sigma^2 = 1/12$

Because many statistical methods rely on random samples, applied statisticians often need a source of “random numbers”. Nowadays, a computer is used to generate “random” (psuedorandom) numbers directly. Psuedorandom numbers are meant to simulate random numbers.

A number of different methods for generating pseudo-random uniform numbers (Gentle, 2000). We describe here a linear congruential generators. The linear congruential method produces a sequence of integers X_1, X_2, \dots between zero and $m-1$ according to the following recursive relationship: $X_{i+1} = (aX_i + c) \bmod m, i = 0, 1, 2, \dots$. The initial value X_0 is called the seed, a is called the constant multiplier, c is the increment, and m is the modulus. All values are integers. If $c \neq 0$, the form is called the mixed congruential method. If $c = 0$, the form is known as the multiplicative congruential method.

Each X_i is scaled into a unit interval $(0, 1)$ by division by m , that is

$$u_i = \frac{X_i}{m} \tag{A-27}$$

If a and m are properly chosen, the u_i s will “look like” they are randomly and uniformly distributed between 0 and 1.

The selection of the values for a, c, m and X_0 are of vital importance to the quality of the generator. If they are not chosen properly, the generator may not have the longest possible period, or generated numbers may not exhibit good randomness, or the generator may not be efficiently implementable. For example, the book by Press *et al.* (1996), suggests that $a = 16807$ and $m = 2^{31} - 1$.

The Inverse Transformation Method

A general method for simulating a random variable having a continuous distribution - called the inverse transformation method-is based on the following proposition.

Proposition A.22 *Let U be a uniform $(0, 1)$ random variable. For any continuous distribution function F if we define the random variable X by*

$$X = F^{-1}(U) \tag{A-28}$$

then the random variable X has the distribution function F . [$F^{-1}(u)$ is defined to equal that value x for which $F(x) = u$]

PROOF. See Proposition 11.1 Ross (2000) for a proof.

Generating Normal Random Variates A simple method for generating standard normal random variates (i.e., with zero mean and unit standard deviation) is based on Box-Müller method arising from a polar transformation: If U_1 and U_2 are independently distributed as $U(0,1)$, and

$$\begin{aligned} X_1 &= \sqrt{-2 \log(U_1)} \cos(2\pi U_2) \\ X_2 &= \sqrt{-2 \log(U_1)} \sin(2\pi U_2) \end{aligned} \tag{A-29}$$

then X_1 and X_2 are independently distributed as $\mathcal{N}(0, 1)$.

A-5 Akaike Information Criterion

A mathematical formulation which considers the principle of parsimony in model building is the Akaike Information Criterion(AIC) proposed by Akaike (1974). For comparing among competing ARMA (p, q) models he used

$$\text{AIC}(p, q) = N \ln(\hat{\sigma}_\varepsilon^2) + 2(p + q)$$

where N is the sample size and $\hat{\sigma}_\varepsilon^2$ is the maximum likelihood estimate of the residual variance. Akaike suggested such criterion to select the correct model among competing ARMA models. Under this criterion the model which gives the minimum AIC is the one to be selected. In case of PARMA $_\nu(p, q)$, the AIC is given within an additive constant(C) by

$$\text{AIC} = -2 \ln L(\phi(\ell), \theta(\ell), \sigma) + C$$

where $\phi(\ell) = [\phi_0(\ell), \phi_1(\ell), \dots, \phi_{\nu-1}(\ell)]^T$, $1 \leq \ell \leq p$; $\theta(\ell) = [\theta_0(\ell), \theta_1(\ell), \dots, \theta_{\nu-1}(\ell)]^T$, $1 \leq \ell \leq q$; $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{\nu-1}]^T$ and $L(\phi(\ell), \theta(\ell), \sigma)$ is the likelihood function. See, for example,

Lund and Basawa (2000) for the Gaussian likelihood of the PARMA model.

A-6 The Bonferroni Inequality and Simultaneous Confidence Intervals

Proposition A.23 (*The Bonferroni Inequality*). *If A_1, \dots, A_k are events (not necessarily independent) which have probability $1 - p$ of occurring, then the probability they all occur is at least $1 - kp$.*

PROOF: The probability that A_1, \dots, A_k all occur is

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = 1 - P(A_1^c \cup A_2^c \cup \dots \cup A_k^c) \quad (\text{A-30})$$

i.e, the probability they all occur is one minus the probability at least one does not occur.

Now,

$$P(A_1^c \cup A_2^c \cup \dots \cup A_k^c) \leq P(A_1^c) + P(A_2^c) \dots + P(A_k^c) \quad (\text{A-31})$$

The right side of the above equal to kp since $P[A_i^c] = 1 - (1 - p) = p$. Therefore,

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = 1 - P(A_1^c \cup A_2^c \cup \dots \cup A_k^c) \geq 1 - kp \quad (\text{A-32})$$

Use with the confidence intervals

A $1 - \alpha'$ confidence interval for a population parameter is constructed such that, before we collect the data, there is a probability $1 - \alpha$ that the confidence interval will contain the parameter. However, if we are going to construct k confidence intervals, each with level $1 - \alpha$, for different parameters from the same data set, the probability that the confidence intervals all contain their respective parameters is less than $1 - \alpha$. One way to ensure that this probability is at least $1 - \alpha$ to make each confidence interval individually have level $1 - \alpha/k$. Then, by the Bonferroni inequality, the probability they will all contain their respective

parameters is at least $1 - k(\alpha/k) = 1 - \alpha$. When we construct confidence intervals such that the confidence that they are all correct is $1 - \alpha$, we call these 'simultaneous $1 - \alpha$ level confidence intervals'. For example, suppose we wish to construct 95% confidence intervals for the population means of 5 variables measured on the same sample. Then $\alpha = 0.05$ so we need to make each confidence interval separately a $1 - (0.05/5) = 0.99$ level confidence interval. We use the t procedure to construct a 99% confidence interval for each mean. We then call the set of 5 confidence intervals 'simultaneous 99% confidence intervals for the population means' or, to be more specific, 'Benferroni simultaneous 99% confidence intervals'. You could also call these 5 confidence intervals 'individual 99% confidence intervals'. Including the word 'individual' indicates that the confidence level applies to each interval separately and not to the whole group simultaneously.

B Vector Difference Equation for Discrete Fourier Coefficients

We wish to relate the Fourier coefficients of $\psi(j)$ to the Fourier coefficients of the model parameters. Let $\phi(\ell) = [\phi_0(\ell), \phi_1(\ell), \dots, \phi_{\nu-1}(\ell)]^T$, $1 \leq \ell \leq p$, be the vector of periodic autoregressive parameters at lag ℓ , $\theta(\ell) = [\theta_0(\ell), \theta_1(\ell), \dots, \theta_{\nu-1}(\ell)]^T$, $1 \leq \ell \leq q$, be the vector of periodic moving average parameters at lag ℓ , and let $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{\nu-1}]^T$ denote the periodic residual standard deviation parameters. These model parameters may be defined in terms of their finite Fourier coefficients $\phi_t^*(\ell)$, $\theta_t^*(\ell)$ and σ_t^* as follows

$$\begin{aligned} \phi^*(\ell) &= U\phi(\ell) \quad \text{and} \quad \phi(\ell) = \tilde{U}^T \phi^*(\ell) \\ \theta^*(\ell) &= U\theta(\ell) \quad \text{and} \quad \theta(\ell) = \tilde{U}^T \theta^*(\ell) \\ \sigma^* &= U\sigma \quad \text{and} \quad \sigma = \tilde{U}^T \sigma^* \end{aligned} \tag{B-1}$$

where ” \sim ” denotes the complex conjugate and U is the $\nu \times \nu$ Fourier transform matrix defined in (5.5).

Using the Fourier transformed of $\psi^*(\ell) = U\psi(\ell)$ and $\psi(\ell) = \tilde{U}^T \psi^*(\ell)$, then (3.38) leads to

$$\begin{cases} \psi^*(j) - \sum_{k=1}^p U A_k \Pi^{-k} \tilde{U}^T \psi^*(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi^*(j) - \sum_{k=1}^j U A_k \Pi^{-k} \tilde{U}^T \psi^*(j-k) = -\theta^*(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \tag{B-2}$$

Note that $\psi_t^*(0) = 1$.

Define

$$D_j = \nu^{-1/2} \text{diag}\{1, w^j, w^{2j}, \dots, w^{(\nu-1)j}\} \tag{B-3}$$

where $w^q = e^{\left(\frac{-i2\pi q}{\nu}\right)}$ and $w^{-q} = w^{\nu-q}$ with $e^{iy} = \cos y + i \sin y$ (the complex exponential function). By properties of the Fourier transform matrix U (Brockwell and Davis, 1991), we

have

$$D_{v+j} = D_j = \nu^{-1/2} U \Pi^{-j} \tilde{U}^T \quad (\text{B-4})$$

also note that U is the unitary matrix ($U \tilde{U}^T = I$).

Theorem B.1 *If we define D_j as in (B-3), then*

$$D_j = \Pi^m U \tilde{D}_m \Pi^{-j} \tilde{U}^T \quad (\text{B-5})$$

where $\tilde{D}_j = \nu^{-1/2} \text{diag}\{1, w^{\nu-j}, w^{\nu-2j}, \dots, w^{\nu-(\nu-1)j}\}$

PROOF: See p.25 of Anderson, 1989.

Writing

$$A_\ell = \sum_{m=0}^{\nu-1} \phi_m^*(\ell) \tilde{D}_m \quad (\text{B-6})$$

where $\phi_m^*(\ell)$ is the m^{th} element of $\phi^*(\ell)$. Then (B-2) becomes

$$\begin{cases} \psi^*(j) - \sum_{k=1}^p (\sum_{m=0}^{\nu-1} \phi_m^*(k) \Pi^{-m}) D_k \psi^*(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi^*(j) - \sum_{k=1}^j (\sum_{m=0}^{\nu-1} \phi_m^*(k) \Pi^{-m}) D_k \psi^*(j-k) = -\theta^*(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \quad (\text{B-7})$$

Thus equation (B-7) gives the desired difference equation relating the Fourier transformed $\psi(j)$ to the Fourier coefficients of the model parameters. It may be further simplified by defining

$$P_\ell = \left(\sum_{m=0}^{\nu-1} \phi_m^*(\ell) \Pi^{-m} \right)^T \quad (\text{B-8})$$

then observing that

$$P_\ell = \text{circ}\{\phi_0^*(\ell), \phi_1^*(\ell), \dots, \phi_{\nu-1}^*(\ell)\} \quad (\text{B-9})$$

Because the model parameters are real-valued, it follows that $\phi_k^*(j) = \tilde{\phi}_{\nu-k}^*(j)$ (Fuller 1996). This implies that P_ℓ is circulant as well as Hermitian (i.e, $P_\ell = \tilde{P}_\ell^T$). Using this fact along the substitution of (B-8) into (B-7) gives

$$\begin{cases} \psi^*(j) - \sum_{k=1}^p \tilde{P}_k D_k \psi^*(j-k) = 0 & j \geq \text{Max}(p, q+1) \\ \psi^*(j) - \sum_{k=1}^j \tilde{P}_k D_k \psi^*(j-k) = -\theta^*(j) & 0 \leq j \leq \text{Max}(p, q+1) \end{cases} \quad (\text{B-10})$$

(a) Periodic Moving Average Processes

The vector difference equation concerning the Fourier transformed innovation estimates of a PMA(q) process given by (B-10) is

$$\begin{cases} \psi^*(j) = -\theta^*(j) & 0 \leq j \leq q \\ \psi^*(j) = 0 & j > q \end{cases} \quad (\text{B-11})$$

(b) Periodic Autoregressive Processes

The vector difference equation concerning the Fourier transformed innovation estimates of a PAR(p) process given by (B-10) is

$$\psi^*(j) = \sum_{k=1}^p \tilde{P}_k D_k \psi^*(j-k) \quad j \geq p \quad (\text{B-12})$$

C Figures and Tables

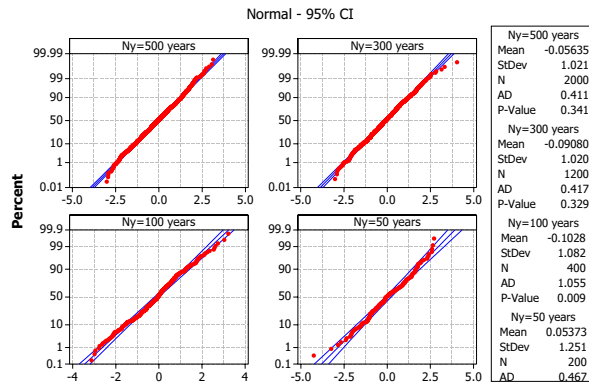


Figure C-1: Normal probability plots of the residuals (for $k=15$) from $\text{PARMA}_4(0, 1)$, support the normality assumptions of residuals.

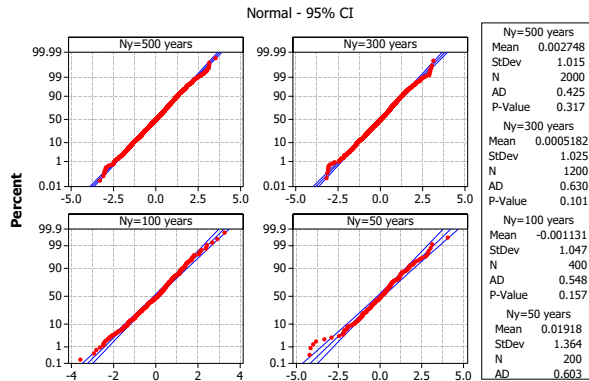


Figure C-2: Normal probability plots of the residuals (for $k=15$) from $\text{PARMA}_4(1, 1)$, support the normality assumptions of residuals.

Table C-1: Moving average parameter estimates and p -values after $k = 20$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated $\text{PARMA}_4(0, 1)$ data.

Parameter	Season 0	Season 1	Season 2	Season 3
θ	0.25	0.65	0.90	0.35
σ	0.90	1.90	0.50	1.20

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.253	0.002	0.425	0.178	0.889	0.000	0.621	0.212
2	0.247	0.430	-0.066	0.718	-0.279	0.332	-0.039	0.734
3	-0.036	0.610	1.215	0.058	-0.114	0.490	0.294	0.238
4	0.230	0.138	0.190	0.204	1.218	0.038	-0.355	0.014
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.630		1.399		0.313		1.099	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.179	0.012	0.687	0.000	0.932	0.000	0.394	0.088
2	-0.020	0.904	0.004	0.976	0.051	0.794	0.081	0.250
3	-0.022	0.668	-0.223	0.522	-0.009	0.952	0.064	0.652
4	-0.039	0.704	0.077	0.466	-0.160	0.624	-0.038	0.712
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.824		1.661		0.503		1.162	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.250	0.000	0.640	0.000	0.899	0.000	0.399	0.008
2	0.289	0.016	0.061	0.516	-0.009	0.936	0.064	0.092
3	0.041	0.174	0.054	0.826	0.044	0.604	-0.121	0.122
4	-0.049	0.424	0.078	0.198	-0.043	0.842	-0.042	0.472
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.844		1.736		0.428		1.123	

Table C-2: Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated $\text{PARMA}_4(0, 1)$ data.

Parameter	Season 0		Season 1		Season 2		Season 3	
θ	0.25		0.65		0.90		0.35	
σ	0.90		1.90		0.50		1.20	

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.217	0.012	0.692	0.032	0.879	0.000	0.703	0.106
2	0.171	0.536	-0.241	0.234	-0.099	0.734	-0.057	0.590
3	0.018	0.788	1.108	0.078	-0.090	0.610	0.279	0.250
4	0.228	0.132	0.190	0.218	0.747	0.168	-0.326	0.028
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.671		1.530		0.364		1.115	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.372	0.000	0.687	0.000	0.881	0.000	0.268	0.276
2	0.108	0.596	0.162	0.266	0.027	0.872	-0.114	0.124
3	0.014	0.818	0.380	0.294	0.061	0.632	0.105	0.448
4	0.046	0.682	0.011	0.920	0.209	0.502	-0.019	0.850
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.836		1.541		0.459		1.131	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.268	0.048	0.634	0.000	0.891	0.000	0.385	0.004
2	0.104	0.348	0.108	0.226	-0.033	0.748	0.005	0.904
3	0.004	0.904	-0.116	0.582	0.081	0.308	-0.074	0.332
4	-0.083	0.174	0.049	0.418	-0.083	0.660	-0.025	0.668
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.880		1.667		0.487		1.144	

Table C-3: Moving average parameter estimates and p -values after $k = 10$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated $\text{PARMA}_4(0, 1)$ data.

Parameter	Season 0		Season 1		Season 2		Season 3	
θ	0.25		0.65		0.90		0.35	
σ	0.90		1.90		0.50		1.20	

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.187	0.026	0.457	0.184	0.879	0.000	0.394	0.378
2	0.215	0.442	-0.108	0.604	-0.299	0.338	-0.061	0.528
3	0.020	0.748	0.573	0.390	-0.143	0.448	0.279	0.242
4	0.343	0.022	0.131	0.368	0.486	0.418	-0.250	0.086
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.736		1.792		0.389		1.229	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.369	0.000	0.690	0.000	0.888	0.000	0.279	0.250
2	0.087	0.674	0.143	0.332	0.070	0.682	-0.112	0.126
3	0.007	0.904	0.460	0.200	0.045	0.726	0.103	0.436
4	0.057	0.604	0.031	0.772	0.315	0.318	-0.017	0.866
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.876		1.570		0.470		1.139	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	0.267	0.000	0.630	0.000	0.892	0.000	0.389	0.004
2	0.094	0.396	0.099	0.272	-0.040	0.696	0.008	0.842
3	0.001	0.968	-0.119	0.568	0.067	0.400	-0.075	0.318
4	-0.085	0.164	0.049	0.424	-0.095	0.610	-0.028	0.632
:	:	:	:	:	:	:	:	:
$\hat{\sigma}$	0.880		1.667		0.487		1.144	

Table C-4: Moving average parameter estimates and p -values after $k = 20$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA₄(1, 1) data.

Parameter	Season 0		Season 1		Season 2		Season 3	
θ	0.25		0.65		0.90		0.35	
ϕ	-0.90		0.50		0.80		0.25	
σ	0.90		1.90		0.50		1.20	

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.794	0.000	1.372	0.000	1.639	0.000	0.257	0.342
2	-0.099	0.780	-0.482	0.082	1.238	0.000	0.652	0.000
3	-0.383	0.004	-0.621	0.250	-0.476	0.200	0.046	0.802
4	-0.058	0.764	-0.232	0.272	-1.366	0.056	-0.602	0.002
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	-0.408		0.764		0.736		-0.142	
$\hat{\phi}$	-0.386		0.607		0.903		0.398	
$\hat{\sigma}$	0.817		1.084		0.420		0.795	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.669	0.000	1.320	0.000	1.685	0.000	0.330	0.180
2	-0.349	0.174	-0.306	0.114	1.217	0.000	0.478	0.000
3	-0.331	0.000	-0.003	0.992	-0.183	0.528	0.292	0.064
4	-0.342	0.020	-0.029	0.810	0.198	0.277	-0.106	0.412
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.390		0.864		0.768		0.047	
$\hat{\phi}$	-1.059		0.457		0.922		0.283	
$\hat{\sigma}$	0.871		1.738		0.440		1.081	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.620	0.000	1.226	0.000	1.698	0.000	0.584	0.000
2	-0.528	0.000	-0.410	0.000	1.043	0.000	0.365	0.000
3	-0.300	0.000	-0.128	0.618	-0.422	0.006	0.330	0.000
4	-0.333	0.000	0.001	0.984	0.012	0.976	-0.078	0.254
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.284		0.565		0.847		0.369	
$\hat{\phi}$	-0.905		0.661		0.850		0.215	
$\hat{\sigma}$	0.860		1.736		0.472		1.171	

Table C-5: Moving average parameter estimates and p -values after $k = 15$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated $\text{PARMA}_4(1, 1)$ data.

Parameter	Season 0		Season 1		Season 2		Season 3	
θ	0.25		0.65		0.90		0.35	
ϕ	-0.90		0.50		0.80		0.25	
σ	0.90		1.90		0.50		1.20	

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.706	0.000	1.260	0.000	1.651	0.000	0.296	0.304
2	-0.152	0.652	-0.560	0.028	1.304	0.000	0.648	0.000
3	-0.429	0.000	-0.731	0.180	-0.405	0.272	0.236	0.234
4	-0.194	0.322	-0.036	0.858	-1.002	0.186	-0.458	0.016
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	-0.193		0.466		0.616		-0.096	
$\hat{\phi}$	-0.513		0.794		1.035		0.393	
$\hat{\sigma}$	0.857		1.218		0.442		0.904	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.673	0.000	1.379	0.000	1.692	0.000	0.325	0.190
2	-0.287	0.266	-0.304	0.114	1.308	0.000	0.455	0.000
3	-0.324	0.000	0.163	0.734	-0.171	0.556	0.343	0.028
4	-0.393	0.008	-0.001	0.992	0.401	0.576	-0.070	0.582
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.211		0.927		0.744		0.056	
$\hat{\phi}$	-0.884		0.452		0.949		0.269	
$\hat{\sigma}$	0.898		1.789		0.456		1.133	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.615	0.000	1.241	0.000	1.693	0.000	0.604	0.000
2	-0.546	0.000	-0.412	0.000	1.063	0.000	0.351	0.000
3	-0.291	0.000	-0.156	0.542	-0.407	0.008	0.322	0.000
4	-0.330	0.000	0.027	0.696	-0.002	0.992	-0.070	0.304
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.289		0.571		0.836		0.396	
$\hat{\phi}$	-0.905		0.670		0.857		0.207	
$\hat{\sigma}$	0.861		1.754		0.476		1.191	

Table C-6: Moving average parameter estimates and p -values after $k = 10$ iterations of the innovations algorithm applied to $N_y = 50, 100, 300$ years of simulated PARMA₄(1, 1) data.

Parameter	Season 0		Season 1		Season 2		Season 3	
θ	0.25		0.65		0.90		0.35	
ϕ	-0.90		0.50		0.80		0.25	
σ	0.90		1.90		0.50		1.20	

Number of Years, $N_y = 50$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.649	0.000	1.412	0.000	1.653	0.000	0.467	0.186
2	-0.190	0.596	-0.473	0.066	1.325	0.002	0.468	0.000
3	-0.310	0.002	-0.541	0.412	-0.527	0.156	0.340	0.126
4	-0.225	0.266	0.096	0.604	-0.913	0.338	-0.251	0.156
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	-0.242		0.683		0.714		0.184	
$\hat{\phi}$	-0.406		0.729		0.938		0.283	
$\hat{\sigma}$	0.880		1.634		0.452		1.133	

Number of Years, $N_y = 100$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.679	0.000	1.413	0.000	1.693	0.000	0.274	0.272
2	-0.247	0.338	-0.328	0.092	1.358	0.000	0.438	0.000
3	-0.320	0.000	0.234	0.632	-0.224	0.448	0.344	0.028
4	-0.398	0.008	-0.003	0.984	0.500	0.496	-0.038	0.764
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.222		0.929		0.732		0.016	
$\hat{\phi}$	-0.900		0.484		0.961		0.259	
$\hat{\sigma}$	0.903		1.848		0.465		1.155	

Number of Years, $N_y = 300$								
Lag	$\hat{\psi}_0(\ell)$		$\hat{\psi}_1(\ell)$		$\hat{\psi}_2(\ell)$		$\hat{\psi}_3(\ell)$	
	$\hat{\psi}_0(\ell)$	p	$\hat{\psi}_1(\ell)$	p	$\hat{\psi}_2(\ell)$	p	$\hat{\psi}_3(\ell)$	p
1	-0.617	0.000	1.252	0.000	1.694	0.000	0.574	0.000
2	-0.519	0.000	-0.402	0.000	1.081	0.000	0.345	0.000
3	-0.287	0.000	-0.111	0.668	-0.403	0.008	0.325	0.000
4	-0.332	0.000	0.023	0.742	0.085	0.826	-0.065	0.332
:	:	:	:	:	:	:	:	:
$\hat{\theta}$	0.287		0.601		0.831		0.370	
$\hat{\phi}$	-0.904		0.651		0.863		0.204	
$\hat{\sigma}$	0.863		1.768		0.481		1.198	

Table C-7: Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 50$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in (\cdot) is the test statistic (5.21).

lag	Coefficient	harmonic m						
		0	1	2	3	4	5	6
1	$\hat{c}_m(1)$	0.360	0.128	0.851*	-0.108	-0.059	-0.067	-0.053
			(2.226)	(14.733)	(-1.877)	(-1.021)	(-1.152)	(-1.306)
	$\hat{s}_m(1)$		0.575*	0.461*	0.024	0.050	0.189	
			(9.960)	(7.976)	(0.419)	(0.857)	(3.269)	
	SE		0.058	0.058	0.058	0.058	0.058	0.041
2	$\hat{c}_m(2)$	-0.029	-0.035	-0.170	-0.021	0.005	-0.005	0.042
			(-0.559)	(-2.742)	(-0.340)	(0.082)	(-0.081)	(0.967)
	$\hat{s}_m(2)$		-0.035	-0.076	-0.185	-0.097	-0.069	
			(-0.558)	(-1.222)	(-2.992)	(-1.569)	(-1.111)	
	SE		0.062	0.062	0.062	0.062	0.062	0.044
3	$\hat{c}_m(3)$	-0.041	-0.012	0.138	0.180	0.111	0.138	0.091
			(-0.200)	(2.221)	(2.901)	(1.796)	(2.224)	(2.075)
	$\hat{s}_m(3)$		-0.129	-0.205*	-0.070	-0.059	0.092	
			(-2.085)	(-3.324)	(-1.122)	(-0.962)	(1.487)	
	SE		0.062	0.062	0.062	0.062	0.062	0.044
4	$\hat{c}_m(4)$	0.000	0.017	0.095	-0.134	-0.213*	-0.218*	-0.108
			(0.274)	(1.526)	(-2.149)	(-3.431)	(-3.516)	(-2.468)
	$\hat{s}_m(4)$		0.037	0.077	0.167	0.092	0.128	
			(0.594)	(1.247)	(2.695)	(1.483)	(2.060)	
	SE		0.062	0.062	0.062	0.062	0.062	0.044
:	:	:	:	:	:	:	:	:

*Fourier coefficients with test statistic ≥ 3.32

Table C-8: Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 100$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in (\cdot) is the test statistic (5.21).

lag	Coefficient	harmonic m						
		0	1	2	3	4	5	6
1	$\hat{c}_m(1)$	0.502	0.261*	0.816*	-0.030	0.000	0.012	0.009
			(6.398)	(19.984)	(-0.742)	(0.004)	(0.303)	(0.308)
	$\hat{s}_m(1)$		0.661*	0.512*	0.018	0.014	0.002	
			(16.193)	(12.554)	(0.435)	(0.349)	(0.055)	
	SE		0.041	0.041	0.041	0.041	0.041	0.029
2	$\hat{c}_m(2)$	0.052	0.037	-0.046	0.013	0.003	-0.031	-0.028
			(0.830)	(-1.024)	(0.292)	(0.061)	(-0.701)	(-0.896)
	$\hat{s}_m(2)$		0.041	0.013	-0.122	-0.085	-0.053	
			(0.903)	(0.295)	(-2.723)	(-1.903)	(-1.191)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
3	$\hat{c}_m(3)$	0.000	0.088	0.177*	0.193*	0.118	0.044	0.060
			(1.967)	(3.944)	(4.300)	(2.631)	(0.973)	(1.896)
	$\hat{s}_m(3)$		-0.107	-0.122	-0.130	-0.071	-0.026	
			(-2.380)	(-2.725)	(-2.907)	(-1.573)	(-0.574)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
4	$\hat{c}_m(4)$	0.117	0.236*	0.105	0.025	-0.063	0.0008	-0.018
			(5.254)	(2.334)	(0.549)	(-1.412)	(0.180)	(-0.562)
	$\hat{s}_m(4)$		0.122	0.155*	0.057	0.100	0.071	
			(2.704)	(3.438)	(1.278)	(2.223)	(1.586)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
:	:	:	:	:	:	:	:	:

*Fourier coefficients with test statistic ≥ 3.32

Table C-9: Discrete Fourier transform of moving average parameter estimates $\hat{\psi}_i(\ell)$ at season i and lag $\ell = 1, \dots, 4$, and standard errors (SE), after $k = 15$ iterations of the innovations algorithm applied to simulated $N_y = 300$ years of simulated $\text{PARMA}_{12}(0, 1)$ data. Note that the value in $(.)$ is the test statistic (5.21).

lag	Coefficient	harmonic m						
		0	1	2	3	4	5	6
1	$\hat{c}_m(1)$	0.477	0.292*	0.786*	-0.013	-0.006	0.003	0.025
			(12.386)	(33.328)	(-0.557)	(-0.252)	(0.145)	(1.491)
	$\hat{s}_m(1)$		0.723*	0.537*	0.018	-0.014	-0.008	
			(30.693)	(22.793)	(0.781)	(-0.576)	(-0.353)	
	SE		0.041	0.041	0.041	0.041	0.041	0.029
2	$\hat{c}_m(2)$	0.027	0.069	0.029	0.041	-0.050	-0.026	-0.018
			(2.700)	(1.122)	(1.590)	(-1.941)	(-1.209)	(-1.006)
	$\hat{s}_m(2)$		0.032	0.016	-0.053	-0.026	0.002	
			(1.245)	(0.645)	(-2.702)	(-1.003)	(0.085)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
3	$\hat{c}_m(3)$	0.057	0.095*	0.087*	0.013	0.047	0.051	0.051
			(3.708)	(3.384)	(0.486)	(1.831)	(1.978)	(2.825)
	$\hat{s}_m(3)$		0.003	-0.010	0.012	-0.048	-0.004	
			(0.106)	(-0.378)	(0.460)	(-1.885)	(-0.142)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
4	$\hat{c}_m(4)$	0.053	0.061	0.054	0.001	0.006	-0.015	-0.027
			(2.364)	(2.118)	(0.049)	(0.230)	(-0.567)	(-1.500)
	$\hat{s}_m(4)$		0.064	0.102*	0.023	0.053	0.022	
			(2.498)	(3.946)	(0.880)	(2.060)	(0.876)	
	SE		0.045	0.045	0.045	0.045	0.045	0.032
:	:	:	:	:	:	:	:	:

*Fourier coefficients with test statistic ≥ 3.32

Table C-10: Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated PARMA₁₂(1, 1) data ($N_y = 100$ years). Note that the value in (.) is the test statistic (5.34) and (5.38).

Parameter	Statistic	harmonic m						
		0	1	2	3	4	5	6
$\hat{\theta}_t$	\hat{c}_{am}	0.438	0.282*	0.220	-0.070	-0.076	-0.185	-0.122
			(3.930)	(3.058)	(-0.969)	(-1.056)	(-2.565)	(-2.407)
	\hat{s}_{am}		0.256*	0.231	-0.043	-0.090	-0.271*	
			(3.554)	(3.216)	(-0.596)	(-1.252)	(-3.767)	
	SE		0.072	0.072	0.072	0.072	0.072	0.051
$\hat{\phi}_t$	\hat{c}_{bm}	0.245	0.063	0.452*	0.010	0.033	0.204	0.145
			(0.769)	(5.643)	(0.119)	(0.405)	(2.472)	(2.478)
	\hat{s}_{bm}		0.466*	-0.056	0.071	0.088	0.213	
			(5.640)	(-0.677)	(0.857)	(1.065)	(2.580)	
	SE		0.083	0.083	0.083	0.083	0.083	0.058

*Fourier coefficients with test statistic ≥ 3.32

Table C-11: Discrete Fourier transform of model parameters estimates and standard errors (SE) for simulated PARMA₁₂(1, 1) data ($N_y = 300$ years). Note that the value in (.) is the test statistic (5.34) and (5.38).

Parameter	Statistic	harmonic m						
		0	1	2	3	4	5	6
$\hat{\theta}_t$	\hat{c}_{am}	0.356	0.140*	0.216*	0.013	-0.062	-0.041	-0.032
			(3.388)	(5.232)	(0.316)	(-1.496)	(-0.999)	(-1.108)
	\hat{s}_{am}		0.430*	0.302*	0.058	-0.004	-0.069	
			(10.449)	(7.339)	(1.410)	(-0.106)	(-1.670)	
	SE		0.041	0.041	0.041	0.041	0.041	0.029
$\hat{\phi}_t$	\hat{c}_{bm}	0.332	0.251*	0.470*	-0.047	0.039	0.057	0.031
			(5.283)	(9.905)	(-0.982)	(0.827)	(1.210)	(0.928)
	\hat{s}_{bm}		0.340*	-0.140	-0.027	-0.007	0.036	
			(7.163)	(-2.952)	(-0.572)	(-0.137)	(0.750)	
	SE		0.083	0.083	0.083	0.083	0.083	0.058

*Fourier coefficients with test statistic ≥ 3.32

Table C-12: Some statistical distributions

Distribution	pdf and/or cdf	Range	Moments
Uniform	$f_X(x) = \frac{1}{b-a}$	$-\infty < a < b < \infty$	$\mu_x = \frac{a+b}{2}, \sigma_x^2 = \frac{(b-a)^2}{12}$
Normal	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right]$	$-\infty < x < \infty$	μ_x, σ_x^2
Lognormal	$f_X(x) = \frac{1}{x\sqrt{2\pi\sigma_y^2}} \exp\left[-\frac{(y-\mu_y^2)}{2\sigma_y^2}\right]$ $y = \ln x$	$x > 0$	$\mu_y = \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right)$ $\sigma_y^2 = \mu_y^2[\exp\sigma_y^2 - 1]$ $CS_x = 3CV_x + CV_x^3$
2-parameter Exponential	$f_X(x) = \beta \exp[-\beta(x - \xi)]$ $F_X(x) = 1 - \exp[-\beta(x - \xi)]$	$x \geq \xi; \beta > 0$	$\mu_x = \xi + \frac{1}{\beta}; \sigma_x^2 = \frac{1}{\beta^2}$ $CS_x = 2$
3-parameter Gamma (Pearson Type 3)	$f_X(x) = \frac{\beta^\alpha (x-\xi)^{\alpha-1} \exp(-\beta(x-\xi))}{\Gamma(\alpha)}$ $\Gamma(\alpha+1) = \alpha\Gamma(\alpha), \Gamma(1/2) = \pi^{1/2}$ $\Gamma(n) = (n-1)!$ for any integer $n \geq 1$	$x \geq \xi; \alpha, \beta > 0$	$\mu_x = \xi + \frac{\alpha}{\beta}$ $\sigma_y^2 = \frac{\alpha}{\beta^2}$ $CS_x = \frac{2}{\sqrt{\alpha}}$
GEV	$F_X(x) = \exp\left\{-\left[1 - \left(\frac{\kappa(x-\xi)}{\alpha}\right)\right]^{1/\kappa}\right\}$	$(\sigma_x^2 \text{ exists for } \kappa > -0.5)$ when $\kappa > 0, x < \left(\xi + \frac{\alpha}{\kappa}\right)$ when $\kappa < 0, x > \left(\xi + \frac{\alpha}{\kappa}\right)$	$\mu_x = \xi + \left(\frac{\alpha}{\kappa}\right) [1 - \Gamma(1 + \kappa)]$ $\sigma_x^2 = \left(\frac{\alpha}{\kappa}\right)^2 \{\Gamma(1 + 2\kappa) - [\Gamma(1 + \kappa)]^2\}$
Weibull	$f_X = \left(\frac{k}{\alpha}\right) \left(\frac{x}{\alpha}\right)^{k-1} \exp\left[-\left(\frac{x}{\alpha}\right)^k\right]$ $F_X = 1 - \exp\left[-(x/\alpha)^k\right]$	$x > 0; \alpha, k > 0$	$\mu_x = \alpha\Gamma\left(1 + \frac{1}{k}\right)$ $\sigma_x^2 = \alpha^2 \left\{ \Gamma\left(1 + \frac{2}{k}\right) - \Gamma\left(1 + \frac{1}{k}\right)^2 \right\}$
Generalized Pareto	$f_X(x) = \left(\frac{1}{\alpha}\right) \left[1 - \kappa \frac{(x-\xi)}{\alpha}\right]^{1/\kappa-1}$ $F_X(x) = 1 - \left[1 - \kappa \frac{(x-\xi)}{\alpha}\right]^{1/\kappa}$	for $\kappa < 0, \xi \leq x < \infty$ for $\kappa > 0, \xi \leq x \leq \xi + \frac{\alpha}{\kappa}$ (CS_x exists for $\kappa > -0.33$)	$\mu_x = \xi + \frac{\alpha}{(1+\kappa)}$ $\sigma_x^2 = \frac{\alpha^2}{[(1+\kappa)^2(1+2\kappa)]}$ $CS_x = \frac{2(1-\kappa)(1+2\kappa)^{1/2}}{(1+3\kappa)}$

where Mean: $\mu_x = EX$, Variance: $\sigma_x^2 = E(X - \mu_x)^2$, Coefficient of Variation: $CV_x = \frac{\mu_x}{\sigma_x}$, and Coefficient of Skewness: $CS_x = \frac{E(X - \mu_x)^3}{\sigma_x^3}$