Winter 2008

A new perspective: Atlantic herring (Clupea harengus) as a case study for time series analysis and historical data

Emily Klein
University of New Hampshire, Durham

A NEW PERSPECTIVE: ATLANTIC HERRING (*CLUPEA HARENGUS*) AS A CASE STUDY FOR TIME SERIES ANALYSIS AND HISTORICAL DATA

BY

EMILY KLEIN

BS, University of California, San Diego, 2003

THESIS

Submitted to the University of New Hampshire

in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

In

Natural Resources: Environmental Conservation

December 2008

This thesis has been examined and approved.

Thesis Director, Andrew A. Rosenberg
Professor of Natural Resources Policy and Management
and Earth, Oceans, and Space

Marryellen M. Lutcavage
Research Associate Professor of Zoology

Andrew B. Cooper
Associate Professor of Resource and Environmental Management
Simon Fraser University

Sept. 8, 2008
Date

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my committee members, Andy Rosenberg, Andy Cooper, and Molly Lutcavage. Andy Rosenberg has been a fantastic advisor, and I know few others would have pushed me intellectually the way that he has. Although he readily provided guidance and support when needed, he always asked questions that forced me to think just a bit more and never gave the answer when I could come up with it myself. I must also thank him greatly for the freedom I was given. He allowed me to pursue questions and research I found interesting, developing my own project and following my own interests. For these reasons, I would not have grown as a scientist the way I have in the past two years with someone else as my advisor.

I thank Andy Cooper for his patience and ability to put up with a barrage of e-mails and questions, even while on vacation. I believe there are very few people who have Andy's ability to explain complex math and fisheries science in such a simple and straightforward manner. Working with Andy has made learning time series analysis far more of a pleasure than it would have been without him, despite the distance. For his accessibility, patience, and guidance, I am thankful.

Thank you so much to Molly and the Large Pelagics Research Center. Without their generosity and support, none of this research would have been possible. Molly provided a new insight and was always enthusiastic about connecting work on herring to the greater ecosystem and higher level predators.

Her lab was also very helpful over the course of my work, whether through discussions of herring and the ecosystem in general or statistical methods. I specifically thank Ben Galuardi and Walt Golet for their time and advice.

I am deeply grateful to Karen Alexander and William Leavenworth of the History of Marine Animal Populations (HMAP) group at UNH. They have spent countless hours helping me find sources, locate contacts and archives, and shift through substantial text for the valuable information. In addition, numerous hours were spent editing and reviewing my work, and their editing ability has been without doubt significant in my success thus far.

Over the course of gathering the information necessary for this work, I am very indebted to numerous individuals across New England and in Canada who provided not only their time, but opened the doors of their archives and libraries to me without hesitation. The individuals at St. Andrews Biological Station were significant in this, specifically Rob Stephenson and Jack Fife. In addition to spending many hours in discussions with me, they took the time to acquaint me with their facilities through a tour guided by Jack, and provided all the necessary resources and access I could need. In addition to Rob and Jack, I would also like to thank Mike Power at SABS, as well as Shale Rosen at the Gulf of Maine Research Institute (GMRI), David Libby and Matt Cieri at Maine DMR, Jeff Kaelin of Research Associates, Al West of Connors Brothers, Jackie Riley of NMFS, Peter Steenstra of Craig Brook Hatchery, and Vaughn Anthony, Ted Ames, and Mary Beth Tooley for their time, advice, and support. Work here could not have continued without their time, generosity, and accommodation.

Last, but certainly not least, I thank my friends and family. To my parents, Marilyn and Rick Klein, I would not be the person I am today or have come this far without their unconditional love and support. I do not know how they did it, but I can only hope to be half the parents that they are. I am so happy to make them proud and can never be able to thank them enough. To my friends in New Hampshire and elsewhere, thank you for the unconditional friendship, stress relief, and always being there when things got rough. Special thank you's to my friends on the phone, Kristan Lunan, Janna Jellerson, and Lee Robinson, and to those close, Jessica Cherry and Jessie Knapp, who all saw me through some very difficult times. Miss Cherry - what would I have done without dinners at your place and basil-lime gimlets.

Thank you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS...iii

LIST OF TABLES...viii

LIST OF FIGURES...x

ABSTRACT...xii

CHAPTER PAGE

INTRODUCTION..1

Background Information..5

General Study Area..12

General Methods...13

Overview..15

I. CONSTRUCTION OF HISTORICAL TIME SERIES & TIMELINE................16

Introduction...16

Methods...17

Results...21

Discussion..32

II. ARIMA MODELING OF TIME SERIES, BACKCASTING FOR COMPLETE TIME SERIES...36

Introduction...36

Methods...48

Results 51

Discussion 72

III. INTERVENTION ANALYSIS 78

Introduction 78

Methods 83

Results 84

Discussion 89

IV. CROSS CORRELATION 93

Introduction 93

Methods 96

Results 97

Discussion 98

CONCLUSION 101

LIST OF REFERENCES 110

APPENDICES 121

APPENDIX A: LIST OF SOURCES BY LOCATION AND YEARS COVERED, INCLUDING A SUMMARY OF CONTENT 122

APPENDIX B: LIST OF INDIVIDUALS INTERVIEWED 126

APPENDIX C: SUMMARY OF CONVERSION FACTORS 128

APPENDIX D: QUALITATIVE TIMELINE 134

APPENDIX E: SCRIPT FOR CHAPTER II 143

APPENDIX F: BACKTRANSFORMATION OF MODELS 148

APPENDIX G: SCRIPT FOR CHAPTER III 150

APPENDIX H: SCRIPT FOR CHAPTERIV 159

LIST OF TABLES

Table 1. Summary of expected influences on herring landings and the data needed to define them as variables....19

Table 2. Reported Canadian herring products....23

Table 3: Reported Canadian herring product units....23

Table 4. Changes in herring product units over time....24

Table 5: Comparison between visually significant fluctuations in the Maine herring fishery time series and the qualitative time line....29

Table 6: Comparison between visually significant fluctuations in the Canadian herring fishery time series and the qualitative time line....30

Table 7. ACF and PACF behavior for hypothesizing possible ARIMA parameters....43

Table 8. ACF and PACF behavior for SARIMA model parameter estimation....47

Table 9. Results from goodness of fit tests for the Maine herring data....55

Table 10. Parameter estimation for the ARIMA(1,1,0) model of Maine herring, 1937-2006....57

Table 11. Results from goodness of fit tests for the Canadian herring data....58

Table 12. Parameter estimation for the ARIMA(1,1,1) model of the Canadian herring data....59

Table 13. Results from goodness of fit tests for the SABS SST data....62

Table 14. Results from goodness of fit tests for the SABS salinity....64

Table 15. Results from goodness of fit tests for the annual SABS SST models....70
-

Table 16. Parameter estimation for the ARIMA(1,1,1) model of the raw SABS SST data....71

Table 17. Results from goodness of fit tests for the SABS salinity data....71

Table 18. Parameter estimation for the ARIMA(1,0,0) model of the SABS salinity data..71

Table 19. Summary of models for transformed data (ln and differenced Maine herring pounds, ln and differenced Canadian herring pounds, differenced SABS SST, and raw SABS salinity)..72

Table 20. Interventions detected in the Maine herring data..............................85

Table 21. Interventions detected in the Canadian herring data................86 – 88

Table 22. Interventions detected with lower sensitivity in the Canadian herring data..89

Table 23. Summary of results from cross-correlation analysis.........................98

LIST OF FIGURES

Figure 1. Canadian herring products over time........25

Figure 2. Time plots of Maine (top) and Canadian (bottom) herring pounds aggregated by year........27

Figure 3. Comparison of select events with the Maine (top) and Canadian (bottom) herring time series........31

Figure 4. Time plot of the Canadian herring time series generated in R statistical software........52

Figure 5. Time plot of the Canadian herring time series generated in R statistical software........52

Figure 6. An example of nonstationary behavior: the ACF and Partial ACF of the raw Maine herring data, 1937 – 200........52

Figure 7. Natual log-transformed and differenced data for the Maine herring, 1937-2006........53

Figure 8. Natual log-transformed and differenced data for the Canadian herring, 1952-2006........53

Figure 9. Time plot of the SABS SST data by month, 1983 – 2006........54

Figure 10. Time plot of the SABS salinity data by month, 1983 – 2006........54

Figure 11. ACF and Partial ACF of the ln-transformed and differenced Maine data........55

Figure 12. Results of residual diagnostic tests for the Maine herring data. At left are the histogram and Q-Q plot of the residuals, at right the time plot, ACF, and p-values for the Ljung-Box-Pierce statistic........56

Figure 13. ACF and Partial ACF of the ln-transformed & differenced Canadian herring data........57

Figure 14. Results of residual diagnostic tests for the Canada herring data. Left: histogram and Q-Q plot of the residuals. Right: time plot, ACF, and p-values for the Ljung-Box-Pierce statistic........59

Figure 15. ACF and Partial ACF of the seasonally differenced SABS SST data (left) and salinity data (right) by month, 1983 – 2006. 60

Figure 16. Results of residual diagnostic tests for ARIMA(1,0,1)x(1,0,1), top, ARIMA(1,0,1)x (0,1,2), middle, and ARIMA(1,0,1)x(1,1,2) of the raw SABS SST data. At left are the residual histograms and Q-Q plots, at right the time plot, ACF, and Ljung-Box-Pierce statistic results........................63

Figure 17. Results of residual diagnostic tests for ARIMA(1,0,1)x(0,1,1) of the raw SABS salinity data. At left are the histogram and Q-Q plot of the residuals, at right the time plot, ACF, and p-values for the Ljung-Box-Pierce statistic.. 65

Figure 18. Completed Maine herring time series, reported and predicted values, 1871-2007..66

Figure 19. Completed Canadian herring time series, reported and predicted values, 1871-2007..66

Figure 20. Completed SABS SST monthly time series, reported and predicted values, 1924-2006..67

Figure 21. Completed SABS salinity monthly series, reported and predicted values, 1924-2006..67

Figure 22. Completed SABS SST annual time series, reported and predicted values, 1924-2006..68

Figure 23. Completed SABS salinity annual time series, reported and predicted values, 1924-2006..68

Figure 24. ACF and PACF for the SABS SST (left) and SABS salinity (right) annual time series. 69

Figure 25. ACF and PACF of the stationary SABS SST time series..................70

ABSTRACT

A NEW PERSPECTIVE: ATLANTIC HERRING (*CLUPEA HARENGUS*) AS A CASE STUDY FOR TIME SERIES ANALYSIS AND HISTORICAL DATA

by

Emily Klein

University of New Hampshire, December, 2008

This thesis endeavors to develop methods for the historical analysis of a specific species and location to begin understanding fishery patterns and change over time. The main goal was to develop statistical methods to address historical data and provide long-term information on fishery trends and potential relationships between the fishery and outside influences. The Atlantic herring (*Clupea harengus*) fishery was investigated for underlying patterns and the possible impact of outside variables and events from 1870 to 2007.

In the Gulf of Maine, Atlantic herring (*Clupea harengus*) provide critical forage for many economically valuable species, while supporting a major New England fishery. Extensive research and stock assessments conducted on herring since the 1960s have focused on recent patterns of distribution, abundance, and other fishery characteristics. This work has often neglected longer-term patterns or changes and the long history of anthropogenic influence and exploitation. Further, the current management strategy for herring may be insufficient and herring ecology is not fully understood. Specific questions remain on stock structure and the viability of inshore populations, in addition to the

possibly major changes in herring abundance and distribution suggested by historical documents. Due to these questions and their ecological and economic importance, herring are an interesting case study for the investigation of historical data and the application of time series analysis (TSA). Here, TSA was used to explore long-term herring fishery data and the possible influence of anthropogenic events and natural drivers from 1871 to the present (2007).

Historical information on Atlantic herring and oceanographic features was compiled from many sources across New England and in St. Andrews Bay, Canada. For herring, the information was aggregated into a time series by total pounds per year for Maine and the Canadian Bay of Fundy. In addition, a time series was built for sea surface temperature (SST) and surface salinity at St. Andrews Biological Station (SABS) in Canada. Finally, a timeline constructed from the qualitative historical text summarized potentially influential socioeconomic and industry events by year. An initial visual comparison explored possible correlation between fluctuations in the herring time series and events in the time line. Viable events were found to explain many of the visually identified fluctuations.

Once time series were constructed, TSA was used to model the underlying patterns of the herring fishery and oceanographic data. More specifically, auto-regressive-integrated-moving-average (ARIMA) models were applied. These models were then used to interpolate the missing years for complete time series, and ARIMA models were run again on these complete data sets. The final model for the Maine herring fishery was an ARIMA(1,1,0), meaning that the pounds in one year was explained, at least in part, by pounds the year before. For Canada, the model was an ARIMA(0,1,1), indicating that the pounds were more explained by the conservation of noise, or error, from the year previously.

The models developed were then used to begin examining the impact of the events from the qualitative timeline and oceanographic features (SST and salinity) on the fishery time series. Intervention analysis detected outliers, called interventions, representing years of unexpected change in the herring time series. These years were compared to the qualitative time line to determine a possible explanatory event. Such events were speculated for the majority of interventions found. Finally, cross-correlation analysis compared the herring time series with the SABS SST and salinity time series for possible cause-and-effect relationships. The analysis found no significant relationships between the series.

This study demonstrated the potential of TSA and historical data, including the qualitative literature, to better understand fisheries over the long term. TSA is a useful tool for applying historical data to study ecosystems in their entirety, from historical fisheries to today, rather than isolated in time or context. Results can broaden the temporal and ecosystem perspective in which fishery statistics are examined, and methodologies can be refined and expanded in the future. However, as used here, TSA addresses only catch statistics, not abundance or other population parameters. These methods should be used in conjunction with traditional statistical approaches and to inform stock assessment.

INTRODUCTION

With an estimated 76% of fisheries stocks fully exploited, over-exploited or depleted (FAO 2005), it is clear that current management is not meeting the needs of many fisheries worldwide (Masood 1997, Batstone and Sharp 2003). Recent studies suggest much of this decline is a failure in fisheries science to provide sufficient and appropriate information to management, comprehensively understand fishery ecology, or adequately predict future scenarios (Walters and Maguire 1996, Masood 1997, Rose 1997, Batstone and Sharp 2003). There is a growing body of literature that suggests the lack of information and understanding may be the result of a limited temporal perspective (Pauly 1995, Jackson 1997, Pauly et al. 2000, Jackson 2001, Jackson et al. 2001, Sáenz-Arroyo et al. 2005, Lotze et al. 2006, Worm et al. 2006). The time frame of current science and management is often too brief to encompass the lifespan of many species, or address the effects of long-term climatic and oceanic cycles (Jackson 2001, Jackson et al. 2001). As a result, we do not have a clear understanding of ocean baselines, long-term changes, or ecosystem interactions, which reduces our ability to manage our oceans effectively (Pauly 1995, Jackson 1997, Jackson et al. 2001, Sáenz-Arroyo et al. 2005, Worm et al. 2006). This evolving area of study argues that an increased awareness of marine ecology history is necessary to supply additional insight and effectively address the current situation (Jackson 1997, Jackson et al. 2001, Smith and Link 2005).

Historical records can reveal ecosystem processes, population patterns, and the prolonged effects of the human community and environmental change. The subsequent long-term view is necessary to appreciate and more successfully manage our marine resources.

The analysis of historical marine ecology is a rapidly growing field. Past research has addressed historic populations of single species (Baumgartner et al. 1992, Rogers-Bennett et. al. 2002, Rosenberg et al. 2005), but analysis has primarily focused on large-scale, often global, trends (Pauly 1995, Pauly 2000, Jackson et al. 2001, Lotze et al. 2006, Worm et al. 2006). Few studies have concentrated on individual species and geographic locations, making the application of historical work to management and research difficult. Approaches to historical analysis must be more specific and directed at answering well-defined questions for particular ecosystems and fisheries. Results from such research can connect historical work directly to management and science and begin to provide information currently lacking in fisheries science.

Atlantic herring (*Clupea harengus*) are an interesting case study for developing methods for historical fishery analysis. In the Gulf of Maine (GOM), these small pelagics provide forage for numerous other species and support a commercially critical New England fishery. Herring have been heavily exploited for well over 150 years, providing an extensive historical record for analysis. The fishery is particularly interesting because the current management strategy may be inadequate and herring ecology is not fully understood. Herring abundance and distribution variability and complex stock structure in the Gulf have resulted in correspondingly variable catches and distinct management challenges (Tupper et al. 1998). At present, GOM herring are managed in the U.S. as a single stock. However, historical documents indicate a much more complex stock structure, suggesting more localized stock and extensive inshore populations (Moore 1898). Although current assessment lists herring as possibly underutilized (TRAC 2006), some fishermen are concerned that inshore populations are fully exploited or overfished (Stevenson et al. 1997, Plante 2006, Libby 2007). Finally, contemporary research and management utilize information only as far back as the 1960s. They rarely address long-term fishery patterns or major changes in herring populations indicated by historical documents (Moore 1898).

Historical analysis may shed light on what management can expect from GOM herring and whether the current management approach and fishery prognosis are accurate. To date, little has been done to assess the fishery over the course of intense human exploitation, and provide an appropriate baseline for evaluating the modern fishery. Although an extensive historical record for herring exists, current management and science is operating without directly utilizing the information available in these resources. Understanding Atlantic herring over the long term may have broad application for other species in fisheries science.

This thesis endeavors to develop methods for historical analysis for a specific fishery and location to provide information and advice to current management. Here, the Atlantic herring fishery is investigated over an extended time period (1870-present) for underlying long-term patterns and the possible influence of outside variables and events. This approach includes quantitative data as well as qualitative information. Qualitative data sources are valuable to understanding fisheries over time, but can be difficult to address analytically. In consequence, they are frequently ignored in statistical methods. The methods presented provide a means of including this information.

The overall goal is to provide information regarding long-term fishery patterns and potential relationships between the fishery and outside influences. Several objectives follow. Objective 1 is the development of databases of historical quantitative and qualitative information on the Atlantic herring fishery and possible influential variables (salinity, sea surface temperature, socioeconomic and industry events). This also includes the construction of time series and a qualitative time line from these databases. Objective 2 is to model these series using time series analysis, which will be described later. These models provide the underlying patterns necessary for further time series investigation, which are the focus of the final two objectives. Objective 3 uses intervention analysis to look for possible impacts of socioeconomic and industry event on the fishery. The final objective investigates correlations between the fishery and oceanographic features.

The techniques developed by this research are meant to be repeatable. Long-term analysis of many species is possible given the lengthy fishery records that are available in historical records, but often overlooked. Methodologies used in this study and the subsequent application of results to fisheries management can be replicated for other fisheries. In this way, the work can begin providing more information across time to deepen our awareness of change and ecosystem dynamics over time.

Background Information

Atlantic herring (*Clupea harengus*) are one of the most commercially and ecologically important species in the North Atlantic Ocean ("Sardines" 1945, Day 1951, Tupper et al. 1998). A migratory and pelagic fish, they are found on both coasts of the North Atlantic, and herring schools are well-known currently for annual variability in abundance and distribution (Scattergood and Tibbo 1959, Tupper et al. 1998, Reid et al. 1999). On the western seaboard of the Atlantic, herring are commonly found from Greenland and Labrador to Cape Cod and Block Island, with winter populations periodically venturing as far south as Cape Hatteras and South Carolina (Munroe 2002). In the Gulf of Maine, herring are one of the most common species and can be found along the entire coast as well as on offshore banks (Earll 1887, Bigelow and Schroeder 1953, Munroe 2002).

General Ecology

Herring are a key forage species in the GOM during all life history stages (Munroe 2002). Researchers have observed numerous other fish species feeding on herring eggs and on egg beds, in addition to invertebrates such as moon snails, hermit crabs, and starfish (McKenzie, 1964, Caddy and Iles 1973, Messieh et al. 1985). Various additional predators on adult, juvenile, and larval herring include cod, haddock, hake, skates, pollock, mackerel, and tuna, among many others (Tupper et al. 1998, Reid et al. 1999, Munroe 2002). Predators are not limited to marine fishes, several species of seabirds, whales, seals, and dolphins also feed on herring (Tupper et al. 1998, Munroe 2002, Stevenson and Scott 2005). Interestingly, a predator of herring cited as being one of the most important by historical documents is squid (Moore 1898), a species that is infrequently mentioned in the literature (Bigelow and Schroeder 1953, Reid et al. 1999, Stevenson and Scott 2005) and not listed as significant in any current documents.

According to current studies, both juvenile and adult herring feed predominantly on copepods (Tupper et al. 1998). Adult herring also ingest euphasiid shrimp, amphipods, mysids, and northern sand lance (Link and Almeida 2000). Historical accounts, on the other hand, record adult herring as feeding equally on copepods and euphasiid shrimp, with shrimp being preferred (Moore 1898). Juvenile and larval herring, unable to ingest the larger shrimps, historically fed primarily on copepods (Moore 1898, Bigelow and Schroeder 1953). Herring feed predominantly at dawn and dusk, or during the night, with increased activity on moonlit nights (Tupper et al. 1998). They are opportunistic feeders and will vary their feeding habits between filter feeding and actively chasing specific prey depending on light and prey availability (Moore 1898, Johnson 1939, Bigelow and Schoeder 1953, Stevenson and Scott 2005).

Physical and Environmental Effects

Physical and environmental factors influencing the distribution of herring include temperature, salinity, currents, and depth (Ridgway 1975). Recent research has shown temperature to be the most influential of environmental impacts, affecting both growth and larval survival rates (Lough and Grosslein 1975, Lough et al. 1980, Graham et al. 1990), as well as feeding ecology and spawning behavior (Graham et al. 1972, Berenbeim and Sigaev 1977, Haegele and Schweigert 1985). However, herring generally exhibit a wide tolerance for varying temperatures (Stickney 1967). They also appear quite tolerant of salinities, although preference for both salinity and temperature may vary seasonally (Stickney 1967). The stage most vulnerable to environmental conditions are herring eggs, developing normally only at temperatures between 8 – 13 degrees Celsius and completely intolerant of salinities below 20ppt (Tupper et al. 1998). As for depth, herring generally inhabit relatively shallower shelf waters, usually less than 100m, although adults undertake seasonal migrations to depths of 200m (Ridgway 1975).

The Historical Atlantic Herring Fishery

Historically, Atlantic herring was a fixed-gear inshore fishery, with grounds "practically continuous" from the Bay of Fundy to Cape Cod (Moore 1898). During the later half of the 1800s, the majority of the fishery focused on the use of weirs for canned or smoked and salted herring, although some other methods were used, including gill nets and torching (Moore 1898). The sardine industry, centered in Eastport and Lubec, Maine, also began and flourished during this time (Hall 1898).

The Gulf of Maine weir fishery officially arose with the introduction of large brush weirs around 1820. Initially, weirs were not particularly successful, but the effectiveness of this method grew rapidly as fishermen became more adept in construction and placement. The fishery was successfully operating by 1828, and weirs had superseded the use of all other gear in the herring fishery (Hall 1898). The weir fishery continued to develop with the expansion of the sardine industry and increasing demand for product (Hall 1898, Moore 1898).

By 1896 it was generally maintained that the fishery was catching more herring than ever before (Hall 1898, Moore 1898). In spite of these assertions, there were also numerous documented claims of a decrease in Atlantic herring and failed weirs by fishermen during the later half of the 1800s. Reports of decreases can be found as early as 1850, and continue through the 1890s. Explanations for these decreases vary by location and fishermen, and include the use of gill nets breaking up schools and weirs capturing too many juvenile herring. Other fishermen also claimed that weirs kept spawning aggregations from reaching their spawning grounds, or that pollution and refuse from local industry or noise pollution from foghorns and steamboats caused herring to avoid certain areas. Moore (1898) concluded that these claims of decline were exaggerated and that no significant decrease in the fishery had occurred. He also declared that there were no practices at the time that would significantly affect the fishery in the future.

Despite this general conclusion, Moore (1898) gave no explanation for the historical loss of certain localized stocks. These included the Quoddy River herring and a population of winter herring that had previously supported a profitable fishery in Maine and the Bay of Fundy. Some local fishermen claimed anthropogenic reasons, including overfishing, for the loss of these stocks. Such conclusions were again dismissed by Moore (1898), although he provided no alternative for their disappearance. Neither stock has returned to any great degree. Both were visually distinct from other populations due to physical characteristics (Quoddy herring) or behavior (winter herring) (Moore 1898). For these reasons, their disappearance was quite conspicuous, and it is conceivable that additional yet less discernible localized stocks may have been in decline or completely lost over time with much less notice.

By the end of the 1800s, numerous technological advances had been made, primarily in the sardine canning industry (Earll and Smith 1887). Purse and haul seining was gaining in popularity, despite laws against them passed under pressure from the weir fishery (Webber 1921). However, by the end of the century, prices were low, competition was high, and the market was overcrowded (Pike 2000, Gilman 2001). Syndicates initiated to regulate competition and pricing controlled almost all of the industry in 1900, but these failed in approximately 1903 (Pike 2000, Gilman 2001). World War I saw another boom in the sardine industry with the demand for cheap food and embargoes on sardine imports from Europe (Davis 1950).

The end of the war saw a slump in the fishery, and it did not recover to previous levels during the 1920s (Davis 1950, Pike 2000). The economic collapse of the U.S. and the Great Depression maintained these low levels, taking a particular toll on the sardine industry (Davis 1950, Pike 2000). Many plants were forced to close, towns went bankrupt, and production fell (Davis 1950). The fishery did not recover until the end of the Great Depression and World War II in 1941 (Davis 1950). Again, demand was high for this time period, and fell again after the war ended (Pike 2000).

Current Fishery Trends

The herring fishery increased dramatically again with the development of new fisheries in the 1960s (Anthony and Waring, 1980). Fishing on Georges Bank began in 1961, the Nova Scotia adult purse seine fishery in 1964-65, and the western adult herring GOM-Jeffreys Ledge fishery in 1967 (Anthony and Waring, 1980). The distant-water fleets of the international fishery placed intense pressure on the Georges Bank stocks during this time, and landings peaked in 1968 (Anthony and Waring, 1980, Stevenson et al., 1997). Otter trawls and purse seines were in heavy use by the last half of the 1960s and early 1970s (NEFMC 1999). The offshore Georges Bank fishery officially collapsed in 1977, and no spawning was observed until 1984 (Anthony and Waring 1980, Stephenson and Kornfield 1990, Townsend 1992, Overholtz and Friedland 2002). As a result, the focus returned to state waters and fixed gear ("Atlantic herring" 2008).

In order to rebuild the western GOM stock, a federal fishery management plan (FMP) was developed in 1976 that included a 200 mile limit for foreign vessels (Stevenson et al. 1997). In 1982, herring was deemed a prohibited species for foreign fleets within the U.S. Exclusive Economic Zone (EEZ), requiring foreign fleets to discard all herring as bycatch ("Atlantic herring" 2008). The resulting inshore shift after the collapse of Georges Bank increased pressure on these stocks and the nearshore fixed gear fishery failed in the 1980s (Stevenson et al. 1997, Tupper et al. 1998). The Georges Bank herring began to recover by the mid-1980s (Stevenson et al. 1997), and efforts were made to shift fishing pressure offshore again to federal waters (NEFMC 1999). As a result, the use of mobile gear and landings increased during this time and through the 1990s (NEFMC 1999). By 1994, mid-water trawling for herring in both the U.S. and Canada had begun ("Atlantic herring" 2008).

Recent trends over the past decade have included an increased shift to mobile gear, a reduced availability of inshore herring to fixed gear, and a dominance of the fishery by single and paired mid-water trawlers. Current assessments of herring on Georges Bank indicate that stocks have recovered (Plante 2006, TRAC 2006), and that the Gulf of Maine fishery may be "underutilized" (Stevenson and Scott 2005). Despite these conclusions, there is concern that remaining nearshore stocks are under heavy exploitation and may be presently overfished (Stevenson et al. 1997, Stevenson and Scott 2005, Plante 2006, Libby 2007).

General Study Area

Walter A. Rich wrote in 1929: "A very striking and peculiar body of water is this Gulf of Maine, markedly different from any other... on the coast line of the eastern United States" (Rich, 1929). The Gulf, a body of water covering 90,700 square kilometers along the eastern seaboard, is unique in bathymetry, water, tides, climate, and coastline. In addition, it has historically supported an incredible array of marine species and an ecosystem so prolific that fishers once believed "no other fishing area equaling [the Gulf]...in productivity exists anywhere else in the world" (Rich, 1929). This abundance has consequently been exploited to varying degrees by people for hundreds of years, and continues to support a plethora of economically vital fisheries today. The scientific, economic, and social values of the GOM it an ecologically important and intriguing area for study. Historical data sets have been maintained and scientific research carried out in the Gulf since well before the turn of the century, providing the historical information necessary for analysis.

This work addresses the Maine GOM and the Canadian Bay of Fundy herring fisheries, from 1870 to 2007. These areas have the most consistent reports and are the most significant for the herring fishery. The 1870s mark the establishment of the sardine industry and the beginning of extensive exploitation of herring. Consequently, they are also the beginning of more accurate and consistent fishery records.

General Methods

Time series analysis (TSA) utilizes statistical models to represent processes over a period of time. These models are based on observations, a sample, of the process taken at regular intervals (Hartmann et al. 1980). The goals of TSA are to examine the correlated and time-related behavior of these observations, use this correlation and behavior to model the series as a function of its own past history, and apply this model to forecast future behavior given current conditions (Parzen 1961, Box and Jenkins 1976, Hare 1997).

Time series analysis is widely used in many fields and sophisticated methods have been developed in economics, business, and the social and behavioral sciences (Hartmann et al. 1980, Chen and Tiao 1990, Hare 1997). In the biological sciences, it has wide application in understanding population fluctuations and forecasting, and may reveal new insight that cannot be detected by traditional statistical analysis alone (Turchin and Taylor 1992, Ellner and Turchin 1995, Kim et al. 1997). TSA can identify significant patterns in ecological data, including long-term fishery statistics. However, these methods have not been as extensively applied to ecology and fisheries, although its use has expanded in recent years (Hare 1997).

Time series analysis (TSA) is appropriate because of its applicability to fisheries data and possibly substantial benefits over traditional fisheries science approaches (Jensen 1976, Hare 1997, Park 1998). It requires less information that may be more reliable that that required for other analyses, and is flexible to address numerous interactions (Jensen 1976, Park 1998). Many of the traditional approaches used in fisheries science require catch and effort data, whereas TSA only requires landings (Jensen 1976). Further, many current procedures are misapplied to catch, and TSA offers a more statistically appropriate alternative (Hare 1997). Although time series analysis is not a new field, applying these methods to fisheries has not been explored extensively. TSA may not have been applied as readily in the past due to the need for many observations over time to be robust (Velicer and Colby 2005). However, the long-term view of historical records provides these observations and opens the door for the application of TSA. It should be kept in mind that TSA as applied to catch does not necessarily answer the same questions as traditional fishery science methods. When using only catch data, TSA cannot make any conclusions regarding fish populations, unlike other methods which strive to do just that.

The application of TSA to fisheries data is particularly pertinent (Hare 1997, Park 1998). TSA has advantages over traditional mathematical approaches to fisheries statistics, particularly for forecasting. It requires only historical time series of data, such as catch over time (Jensen 1976). Traditional methods often require derived variables and additional information, such as effort, which can be less available and accurate than catch (Jensen 1976). Methods of time series analysis offer major benefits for addressing the questions here by incorporating temporal structure that other approaches do not (Park 1998).

Overview

This thesis is divided into four chapters. Chapter One involves the construction of historical databases and time series necessary for analysis. In Chapters 2, ARIMA methods were used to model the time series for herring in the Gulf of Maine and two oceanographic variables: sea surface temperature and salinity. These models are developed in order to perform intervention and cross-correlation analyses. Chapter 3 uses intervention analysis to compare the herring landings to events in the qualitative literature. Cross-correlation is addressed in Chapter 4, where it is used to investigate possible relationships between herring landings and the oceanographic variables.

CHAPTER I

CONSTRUCTION OF HISTORICAL TIME SERIES & TIMELINE

Introduction

Identifying and accessing historical data can be challenging and time-consuming. Historical documents contain a great deal of information, but not all is relevant and it can be difficult to access and organize. Extracting useful data frequently requires delving through numerous archives and substantial texts, at times reminiscent of the proverbial search for a needle. However, further analysis cannot continue without this initial lengthy process.

Historical information is both quantitative, in the form of tables and statistics, and qualitative, as descriptive text. Quantitative data can be explored using many analytical methods from ecology and statistics, but the use of qualitative or anecdotal records in current ecological studies is less common. However, it can be amenable and informative to research, especially historical analysis. Printed and manuscript texts contain pertinent qualitative content, and incorporating it can broaden understanding of ecosystems over time. These records can provide additional industry, political, social, economic, or environmental information. For fisheries science, it may place fisheries in a broader human context, as opposed to analyzing quantitative data in isolation.

Additionally, anecdotal text may inform statistical methodologies. Understanding industrial, political, economical, or social influences can help determine the accuracy of tabular data and statistical conclusions. For example, no fishery reports were available for the Great Depression, and the backcasted values (Chapter Two) for this time predict relatively consistent landings. However, the reading of anecdotal accounts regarding the effect of the Depression on the herring industry calls into question these predictions. Thus, placing a fishery in an historic human context with descriptive text can help confirm or question analytical conclusions. It can also help inform or develop analysis and determine what to expect from results (Facchini et al. 2007).

This chapter concerns how the historical herring and oceanographic data for this thesis was acquired and organized into databases and the resultant time series. The goals are to provide appropriately organized information for future analysis, and begin including qualitative information in the analyses. Numerous historical documents on Gulf of Maine fisheries (Appendix A) provided the information, which resulted in two quantitative time series and one qualitative time line. These data sets were necessary for the additional statistical approaches of Chapters 2-4.

Methods

During the summer of 2007, data sources were identified and records acquired in Maine, Massachusetts, and Canada. Interviews with individuals in herring management, science, and industry aided in identifying additional sources and how historical analysis might benefit current management and science. A list of these individuals is in Appendix B and a summary of sources and records is in Appendix A. Once acquired, both quantitative and qualitative information was prepared for further analysis. Databases and time series of quantitative statistics were organized for plotting and the mathematical approaches of Chapters 2-4. Time plots of both the time series and the time line provided the preliminary visual comparison.

Data Acquisition

There were four general areas chosen as possible sources for influences acting on the herring fishery. Initially, these general areas were broken down into more specific components to determine the types of data needed to express them as variables in the analyses (Table 1). Previously mentioned interviews with herring management, science, and industry representatives helped define and identify additional resources for information. However, not all aspects could be addressed, because information was either not available or was unattainable within the time constraints of a master's thesis. This thesis focuses on the Maine GOM and Canadian Bay of Fundy herring fisheries, specific oceanographic features (sea surface temperature and salinity), and identifiable socioeconomic and industry events from the qualitative literature.

	Specific components	Data needed
Area 1: Environment	• Temperature • Salinity • Currents • Weather patterns • Wind	• Long-term oceanographic data, data on weather patterns, nutrients • Knowledge of effects on herring
Area 2: Market & Demand	• Demand/Consumer preference • Price: what's valuable? • Factories: strikes, fires, etc.	• # Factories, information on factories over time • Preferences, prices • Recessions, etc. • Descriptions of market
Area 3: Fishery	• Changing effort • Changing technology • Movement of fishery over time & why	• Effort over time • Technological changes and their effects • Changing grounds, etc
Area 4: Social	• Changing laws • Changing population patterns • Additional events	• Laws over time • Social information on where people lived and why • Additional events

Table 1. Summary of expected influences on herring landings and the data needed to define them as variables.

Data Sources

Historical information came primarily from governmental and industrial sources in Maine, Massachusetts, and New Brunswick, Canada. Appendix A has a summary of locations and the primary sources available. Information existed as both tables and text, which was categorized as either quantitative (fishery statistics) or qualitative (descriptive). The spatial and temporal detail of these data varied over time and by source and location.

Time Series – Quantitative Information

After acquisition, databases organized quantitative statistics from tables in the historical reports. These databases were classified by source and varied by time period, frequency of reporting, length of reporting, product units and product type. Maine data were consistent as landings in pounds or metric tons but were spatially and temporally disaggregate. Historical Canadian herring data were reported by product, and these product types and their units varied in time (tables 1C-D.2, fig. 1A). In order to combine such incongruent data sets, a common time scale and unit would be needed. Summing by year temporally aggregated the data, as not all reports were at any finer temporal scale. Different reported herring products were aggregated through a common unit, in this case by weight in pounds because many landings were already in pounds. Combining data via a common weight unit required conversion factors for those products not in pounds. The qualitative literature determined these factors (Appendix C).

Once constructed, time series were plotted over time. These time plots can guide the time series analysis described in Chapter 2. The existence of trends, particularly periodic behavior, can help define approaches used. The time plot was also compared to the qualitative timeline discussed below.

<u>Timeline – Qualitative Information</u>

A wealth of information regarding the herring fishery, industry, and socioeconomic and political atmosphere is available in the qualitative text. Such information can be valuable for understanding a fishery over time, but does not lend itself easily to quantitative analysis. To incorporate this information into the work here, it was summarized into a timeline of events that may have had a significant effect on herring landings.

Once complete, the timeline was compared to time plots of herring landings. This comparison identified possible relationships between events and landings variations. Preliminary comparison involved plotting only particularly important events against the herring time series. Relatively large anomalies in the landings were then noted and compared to the qualitative timeline in its entirety. Lagged relationships, where fishery effects appeared more than one year after a related event, were also considered, but are more difficult to identify as the actual time lag is unknown.

Results

Data Acquisition

Quantitative historical information was available on herring landings and products in addition to information on vessels, gear, etc. Historical oceanographic information regarding weather, rainfall, sea temperature, and salinity also exists. This thesis included herring and sardines, sea surface temperature (SST), and salinity. Recent information (1960 – 2007) was already in digital form, but historical records (late 1800s – 1960) required transcription from paper reports. See Appendix A for a summary of these reports. Records were quite consistent and only small periods or single years were missing. Some intervals contained additional spatial or temporal detail.

Qualitative texts contained explanations of herring fishery practices, descriptions of market forces (including consumer preference and overall demand), industry information (number of factories, cannery fires and worker strikes, etc.), information on developing fishery technology, gear changes, and accounts of weather, etc., as well as fishermen's interviews. Accounts of additional social aspects such as wars, international relations, economic markets, etc., were also available. Using this information from extensive and reliable resources, a fairly continuous timeline regarding market, industry, and technology, was possible.

Time Series – Quantitative Herring Fishery Information

As mentioned previously, herring and sardine fishery reports varied in time, frequency, and by product and unit reported. Many reports were yearly and in pounds landed, but some were by month or by herring product prepared. A common time interval (year) and unit (pound weight) aggregated the data to construct a complete time series for analysis. Historical herring and sardine pounds were combined as well, because current landings report them collectively.

Data was easily summed by year, but translating the various herring products to pounds proved more challenging. Canadian data reported 26 different measures of herring products, 14 in the tables and another 12 in the qualitative text, while the United States reports were consistent in pounds or metric ton. Additional difficulty arose because products were not prepared consistently over time (see figure 1). While eight different herring products may be reported for several years, only four or five may be reported for the next decade. Reasons for this disparity are unknown, but likely tied to market forces such as price and consumer preference. Units for products also varied with time, some products being reported in several different units over many years. A list of herring products and units is in tables 2, 3 and 4. This made combining data sets via a common denominator difficult. Fortunately, there is an abundance of information in the qualitative literature on factors to convert products to pounds, although it is not always readily available.

In statistical tables	**In text**
Herrings	Round herring
Salted or Pickled (used interchangeably)	Gibbed herring
Smoked	Split herring
Smoked and Kippered	Hard or red herring
Kippered	Bloater herring
Kippered in Cans	Kippered herring
Kippered/Boneless	Length-wise
Skinned/Boneless	Medium-scaled
Large Canned	No. 1
Home Consumption	Tucktails
Fresh or Frozen	"Brook-trout"/"Sea-trout"/"Ocean-Trout"
Canned	"Mustards"
As Bait	"Oils"
As Fertilizer	"Herring mackerel"/"Blueback mackerel"

Table 2. Reported Canadian herring products.

Barrel	Can
Barrel - Fresh herring	Can - Quarter (4.5"x3"x1")
Barrel - Herring oil	Can - Half (4.5"x3.5"x2")
Barrel - Pickled herring	Can - Sardines
Barrel - Russian sardines	Case
Basket	CWT
Box	Herring - Bloated (100)
Box - "Length-wise" (12"x6.5"x2.75")	Herring - Round
Box - "Med-scaled" (12"x6.5"x2.75")	Herring - Gibbed
Box - "No. 1" (12"x6.5"x2.75")	Herring - Split
Box - Pickled herring	Hogshead
Box - Smoked herring	Keg - Russian Sardines
Box – Bloaters	Pail - Russian sardines
Bushel (US standard)	Sardine - Russian

Table 3: Reported Canadian herring product units.

Year	Product	Previous Unit	New Unit
1871	Herrings	first appearance in record	brls
1872	Herrings, smoked	first appearance in record	boxes
1880	Sardines	first appearance in record	hhs/hhds
1881	Herrings, frozen	first appearance in record	per 100/ per 100 lbs
1883/ 84	Herrings, home consumption	first appearance in record	lbs
	Herrings, pickled	first appearance in record	brls
1892	Herrings, smoked	boxes	lbs
	Herrings, salted	first appearance in record	brls
	Sardines, canned	first appearance in record	cans
1893	Herring, fresh or frozen	per 100/per 100 lbs	lbs
	Sardines, preserved in cans	first appearance in record	cans
1894	Sardines, canned	cans	lbs
	Sardines	hhs/hhds	brls
1895	Sardines, canned	lbs	cans
1898	Sardines "in oil"	first appearance in record	
	Herrings, kippered in cans	first appearance in record	cans
	Herrings, kippered	first appearance in record	lbs
1906	Herrings, smoked & kippered	first appearance in record	lbs
1907	Herrings, boneless & kippered	first appearance in record	lbs
1908	Herrings, skinned & boneless	first appearance in record	lbs
1910	Sardines, canned	cans	cases
	Herrings (landed)	first appearance in record	cwt
	Herrings, smoked	lbs	cwt
	Herrings, as bait	first appearance in record	brls
	Herrings, as fertilizer	first appearance in record	brls
	Herrings, used fresh	first appearance in record	cwt
	Herrings, smoked & kippered	lbs	cwt
1911	Herrings, canned	first appearance in record	cases
1912	Sardines, fresh & salted	first appearance in record	brls
	Herring, canned, kippered	cans	cases
1913	Sardines, fresh & salted	first appearance in record	cwts
1914	Sardines, fresh & salted	cwts	brls
	Herring	brls	cwt

Table 4. Changes in herring product units over time

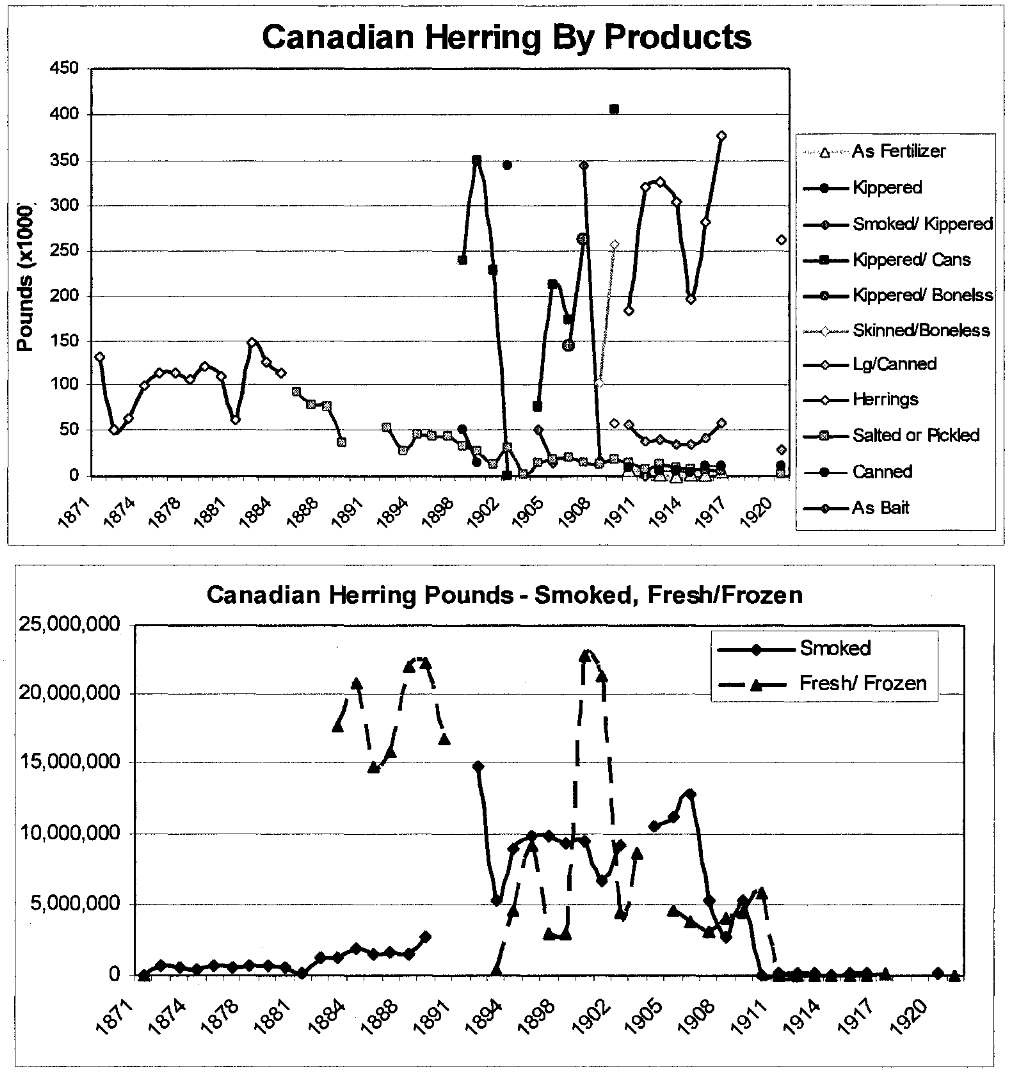
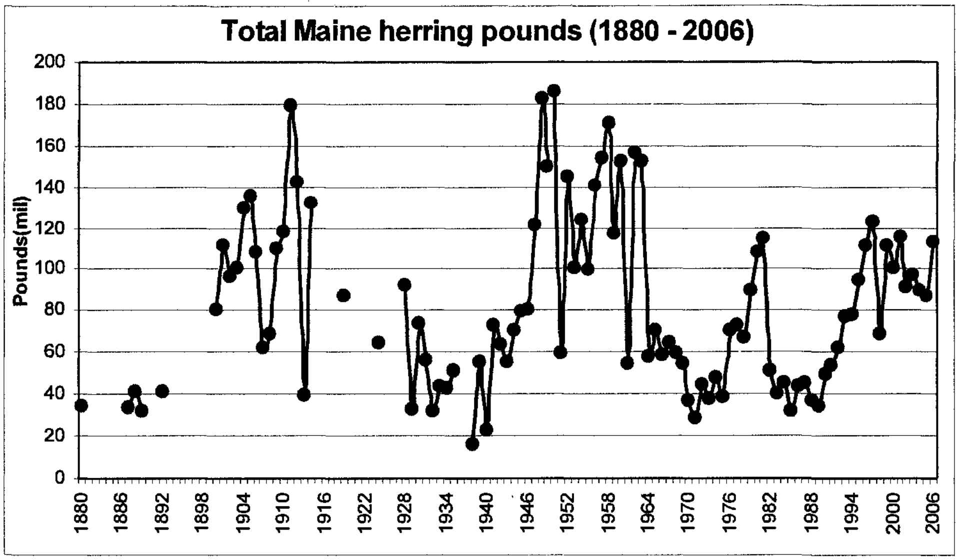


Figure 1. Canadian herring products over time. Top is all products except for smoked and fresh or frozen herring, which are plotted on bottom.

Conversion factors translated products into either prepared fish pounds or fresh fish pounds (Appendix C). Fresh fish pounds were chosen because they allowed ready combination with current landings. For a few products ("as bait," "as fertilizer," "skinned/boneless," and all kippered products aside from those in cans), conversion to prepared fish was the only factor established, and was used instead. Although not ideal, using prepared pounds in these few instances was considered conservative. Prepared fish weigh less than fresh fish as they are generally cleaned, gutted, dried, etc. Furthermore, descriptions in the literature were clear that conversion resulted in pounds of fish separate from pounds of additional ingredients in prepared products. Therefore, the use of prepared pounds was acceptable when necessary. No conversion factors could be found for two products (canned in cases and kippered cans in cases), and they were subsequently dropped from the analysis. Products without a fresh fish conversion or a conversion factor at all amounted to a small proportion of the total pounds and were deemed negligible to the overall analysis (5.8% total, 5.7% prepared fish conversion only and 0.1% no conversion).

The time series for both Canadian and Maine herring data were successfully combined via a common time interval and weight unit. The use of conversion factors to aggregate products by pounds was extensive in the Canadian data, but much less so for Maine. The final Canadian herring time series includes 1871 – 2007, and Maine covers 1880 – 2007. Both have missing years that are addressed in the next chapter. All final series were plotted over time (fig. 2). No apparent trends were visually identified to inform the time series analysis in chapter two. Finally, time plots were compared to the qualitative timeline, discussed shortly.



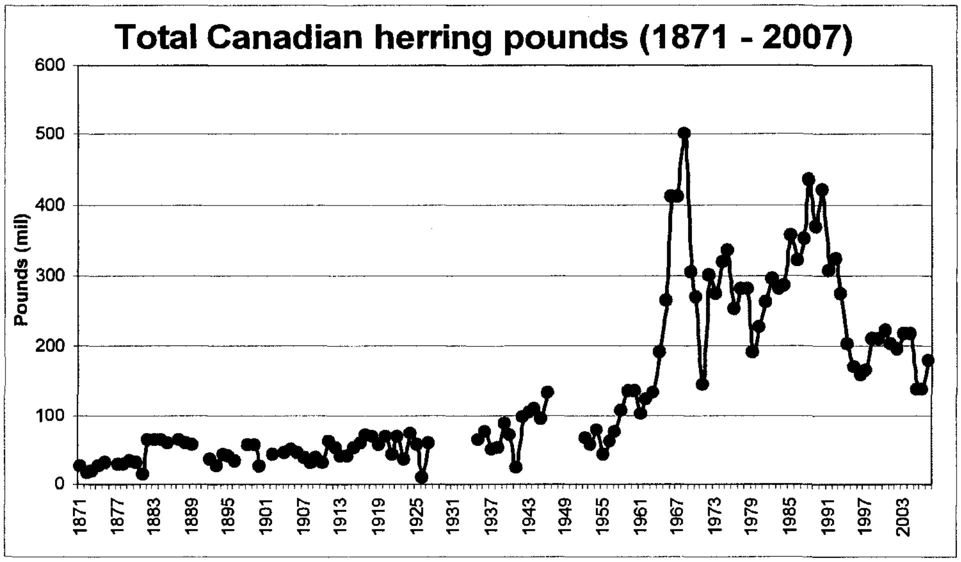


Figure 2. Time plots of Maine (top) and Canadian (bottom) herring pounds aggregated by year.

Time Series – Quantitative Oceanographic Information (SST and Salinity)

Oceanographic data exists primarily as numerical tables, both in digital and paper form. This research used sea surface temperature (SST) and salinity from St. Andrew's Biological Station (SABS) in New Brunswick, Canada. These datasets were already in digital form and reported by month from 1924 to 2007. Both data sets included missing monthly values which made accurate annual aggregation impossible. These missing months had to be estimated first (Chapter 2). Additional oceanographic data exists in digital form and in paper texts that would require transcribing, but the use of this information was outside the scope of this thesis.

Timeline – Qualitative Information

In addition to the time series constructed, a time line of qualitative information was developed (Appendix D). This information was restricted to significant social, political, or fishery-related events and ranged from changes in fishery gear to world wars. Preliminary comparisons of select events from this time line to the fishery time series yielded little apparent correlation when plotted (fig. 3). However, evaluations of relatively large fluctuations in the Maine and Canadian fisheries against the time line in its entirety produced a possible explanatory or contributing event for all major fluctuations (table 5 and 6). This was more so for Canada than for Maine. Most significantly, the World Wars, the Great Depression, and the impact of the offshore foreign fleet and subsequent collapse of the Georges Bank fishery could be most easily hypothesized as influencing herring landings in Maine and Canada.

Year	Increase/ Decrease	Qualitative timeline
1902-3	Decrease	Syndicates fail after attempting to regulate an overcrowded market
1905	Increase	Technological advances
1906-8	Decrease	
1909-11	Increase	Lagged increase (?) - Technological advances, railroad come to Eastport
1912-13	Decrease	Sanitation legislation expanded and enforced
1915	Increase	
1916-1927	No reporting	WWI 1914-1918 increases demand, slump in industry at end with no recovery
1928-29	Decrease	Great Depression 1929-33
1942-48	Increase	WWII and end of Great Depression in 1941, boom in industry is maintained until 1948 according to documents
1951	Decrease	
1953-8	Increase	
1961	Decrease	
1962-3	Increase	Foreign fleets begin offshore fishery in 1961
1964-75	Decrease	
1976-81	Increase	Intense offshore fleet, collapses in 1977 Magnuson Fishery Conservation & Management Act passed in 1976
1982	Decrease	Herring placed on prohibited species list – no-take for foreign vessels within US EEZ, Georges Bank begins to rebuild in mid-80s
1992	Increase	
1998	Decrease	Lagged correlation (?) - Mid-water trawling begins by both US and Canada in 1994

Table 5: Comparison between visually significant fluctuations in the Maine herring fishery time series and the qualitative time line.

Year	Increase/ Decrease	Qualitative timeline
1882	Increase	Lagged increase (?) due to Lubec sardine industry and seining beginning in 1880
1890	Minor decrease	Smoked herring industry important at this time – at low point according to historical text (no trade w/south), was at high in 1889
1897/98	Minor increase	1897 reported as poor year (some sardine factories closed early), improvement reported for 1898
1900	Decrease	Syndicates attempted in US to regulate industry, competition up and market overcrowded for sardines
1914-7	Gradual increase	World War I – increased demand (esp. sardines)
1924-6	Decrease	Slump due to end of war – no recovery, depression in industry begins in the early 1920s
1929-33	No reporting	Great depression, predicted values could inaccurate
1939-46	General increase	World War II – increased demand (esp. sardines)
1947-52	Decrease	Unsure when decrease happened due to missing values. Poor run of herring (1948) and decline in business described
1955-1968	Increase	Offshore foreign fleet develops in early 1960s
1968-71	Decrease	Herring heavily fished by otter trawls and purse seines (1969-72) and by the foreign fleet on Georges Bank
1979	Decrease	Lagged decrease (?) - offshore herring in the GOM crashes in 1977
1980-90	Increase	Herring on prohibited species list – no-take for foreign fleets within US EEZ, Georges Bank begins to rebuild. Fishery focused on inshore
1990-6	Decrease	ASMFC adopts new FMP to address growth of herring resource (1994)
1997-8	Increase	Lagged increase (?) - Mid-water trawling begins in 1994
2005-6	Decrease	ASMFC & NEFMC develop new amendments (2003)
2007	Increase	Amendment 1 to Herring FMP (2006) – limited entry for vessels

Table 6: Comparison between visually significant fluctuations in the Canadian herring fishery time series and the qualitative time line.

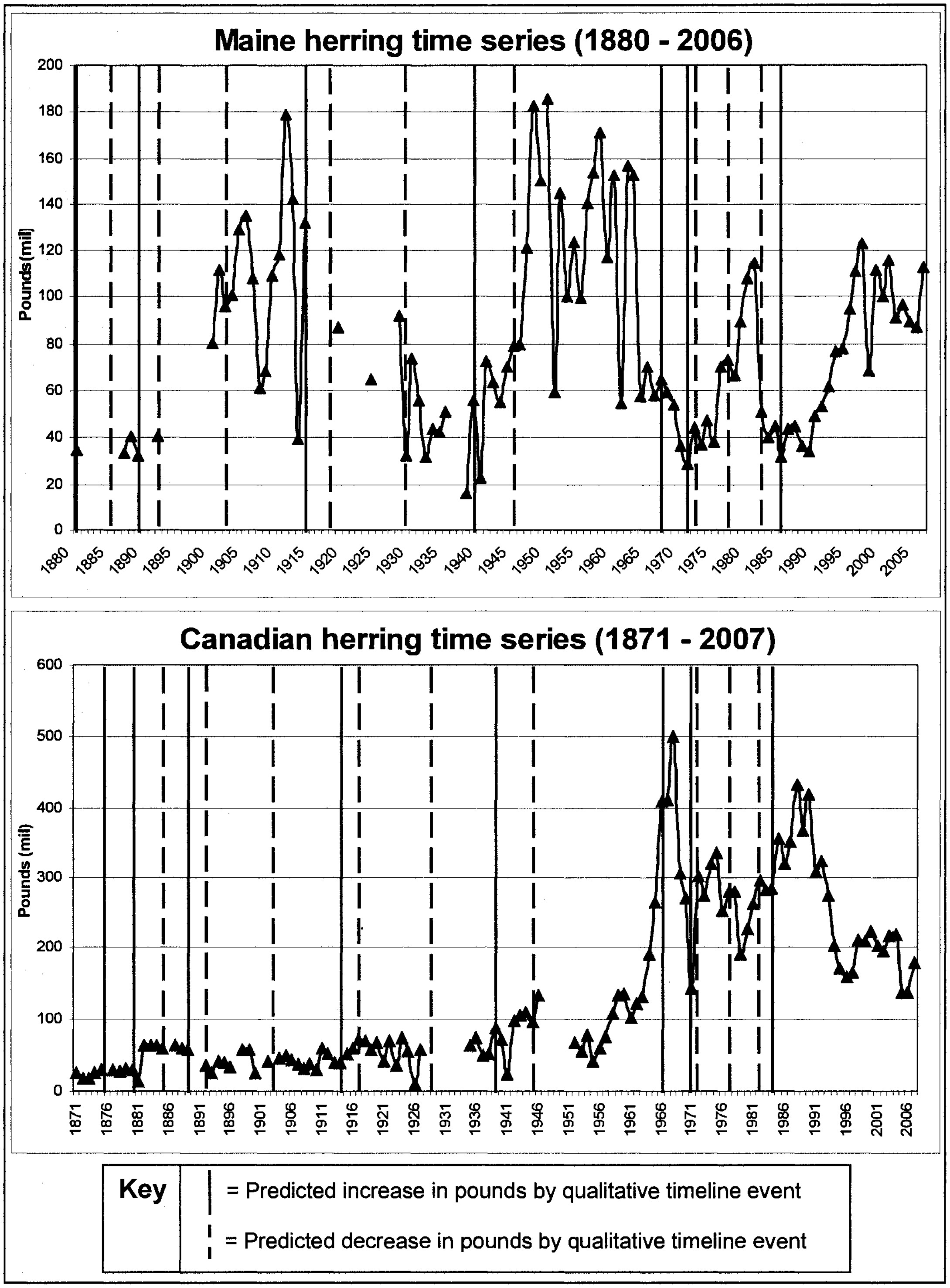


Figure 3. Comparison of select events with the Maine (top) and Canadian (bottom) herring time series.

Discussion

The research here makes clear the value of the historical and qualitative literature. Without anecdotal text, the combination of incongruent data sets of fishery statistics would be extremely difficult, if not virtually impossible. Factors for converting various reported herring products into a common unit of measurement (e.g. weight) are necessary, and are found in historical text. Furthermore, these accounts provide valuable descriptions of herring products, allowing full understanding of the statistics themselves. A comprehensive reading of the accompanying texts is therefore central to understanding historical fisheries and must not be overlooked. The inclusion of additional information via a qualitative timeline, discussed later, further emphasizes this point.

Due to the extended time period and varied data sources, uncertainty was induced into the derived time series. This was exacerbated by the conversion of various products into pounds for several reasons. First, it is not possible to assess the accuracy of conversion factors, especially given the extended time period of this work. In addition, there is little information regarding the time frame for which a factor was appropriate. On occasion, anecdotal information updated factors or confirmed their validity, but this was not always the case. Nonetheless, information available is extensive for improving or more extensively confirming these conversion factors in the future. However, due to the amount of searching required and information available, it was difficult to be thorough in acquiring conversion factors for this extended period and variety of products at this time.

An additional source of uncertainly was discrepancy between reports, primarily in the Maine herring data. There were periods of overlap among various sources for the U.S., and total annual pounds reported did not always agree. These differences sometimes amounted to more than 500,000 pounds. To be consistent, reported landings from Dow (1951) and the Maine Department of Marine Resources were used for the majority of the overlap because the numbers in these sources had the most agreement across all reported landings.

The uncertainty described above is in addition to that inherent in fisheries information, such as misreporting or errors in reporting or transcribing. However, of concern here are relative changes and patterns in the fishery over time, not exact landings for a certain year. Such uncertainty may be less significant in this particular analysis as long as such errors are consistent over time and the underlying patterns are preserved.

Despite these areas of uncertainty, the main goal was to determine long-term patterns that are preserved through time. Assuming that such patterns are conserved, even if exact landings for a certain year is inexact, is a reasonable assumption for this analysis. In addition, the intervention analysis of Chapter 3 is designed to pick up inconsistencies and unexpected changes in the data. These can then be evaluated against changes in reporting, data sources, and the use of conversion factors. This is discussed further in Chapter 3.

Once data is combined and the final time series built, time plots should be visually examined for evident patterns, such as cyclic behavior or obvious level shifts (fig. 2). If patterns exist, they can help guide further analysis. Cycles can lead to and help define frequency-domain time series analysis, and obvious changes can provide insight into intervention analysis. Although cyclic behavior has been seen in other small pelagic fisheries over extended time periods (Baumgartner et al. 1992, Klyashtorin 1998), the fact that neither series exhibited visually identifiable patterns in the time plot is not surprising. A wide range of variables influenced both fisheries over this long time period, presumably effecting change at different times and for different reasons. Further in-depth analysis is required to begin to unravel these processes. Additional statistical approaches used here are the focus of Chapters 2-4.

Visual examination can also provide initial hypotheses of correlation between series and events. However, this was strictly a preliminary and subjective investigation that gave no definitive results. The hypothesized influential events listed in tables 5 and 6 are purely speculative about possible correlations between events and fishery fluctuations. Such preliminary speculation can still provide additional insight, aid in connecting the fishery to a broader human context, and guide further, more objective approaches. Initial comparisons help identify particular years to address for general intervention analysis, narrowing the number of years to include. Although outside the scope of this work, comparisons can also promote the inclusion of qualitative information via hypothesis testing. For example, a management act in year X could be hypothesized as having a significant effect on herring landings the following year. Such hypotheses conjectured from visual assessments can then be analytically tested using methods such as intervention analysis. Utilizing this approach in conjunction with time series analysis can result in more rigorous conclusions because the hypothesis can be tested as being in addition, and not as an alternative, to underlying patterns.

In general, visually comparing the timeline to the fishery time series can identify plausible correlations other than fishing pressure or environment. Further analysis can then be carried out via objective statistical methods, such as the intervention analysis in Chapter 3 and the multivariate analysis of Chapter 4. A fishery can thus begin to be understood not only in terms of landings or environmental influence, but also the greater human system. In addition, pertinent qualitative information can be included to inform analysis and enhance overall comprehension.

CHAPTER II

ARIMA MODELING OF TIME SERIES, BACKCASTING FOR COMPLETE TIME SERIES

Introduction

Time series data is common and cover a wide range of disciplines and information (Parzen 1961, Cryer and Chan 2008). It is usually observations made or measurements taken at regularly spaced intervals (StatSoft 2003, Shumway and Stoffer 2006), and the result of a consistent underlying mechanism and random (white) noise, or error (Chatfield 1977, Hartmann et al. 1980, StatSoft 2003). Compared to traditional statistical data, time series data is many observations of one subject or process through time, as opposed to taking one observation on many subjects or processes at one point in time (Velicer and Colby 2005). This encourages the investigation of change over time and the identification of underlying process patterns (Velicer and Colby 2005). To do this, the application of TSA is limited by the number of observations required, the minimum recommended being from 20 to 50 values. The application of time series data also does not accommodate missing values in a data set. Both of these constraints can make acquiring complete data sets of enough observations difficult (Velicer and Colby 2005).

Applying traditional statistics to time series data can lead to erroneous conclusions. Conventional methods assume that adjacent observations are independent and identically distributed (Shumway and Stoffer 2006), and that the series is not dependent on time (Chatfield 1997). In time series data, these assumptions are violated (Box and Tiao 1975, Box and Jenkins 1976, Shumway and Stoffer 2006). Positive interdependency between observations decreases the apparent variability in the data and increases the probability of a Type I error, whereas negative dependency increases variability and Type II errors (Velicer and Colby, 2005). TSA accounts for this inherent interdependency among values, thus leading to accurate statistical conclusions (Box and Jenkins 1976, Velicer and Colby, 2005, Shumway and Stoffer 2006).

There are frequency-domain and time-domain TSA. The focus here is on the time-domain, which describes current values as being dependent, or correlated, on past observations, as opposed to the frequency-domain, which utilizes periodic sinusoidal patterns to describe fluctuations (Shumway and Stoffer 2006). Dependencies or correlations between observations are described in terms of lags. A lag is the time interval between units, therefore a lag=1 correlation is when the observation is correlated, or dependent on, the process at one time interval in the past (Hartmann et al. 1980). Correlations of lag = 2 are therefore correlated on the observations one unit and two units of time in the past, and so on (Hartmann et al. 1980).

Time series observations can be visualized as one realization of a stochastic process that could have generated many time series (Hartmann et al.

1980, McDowall et al. 1980). Often, only one realization is available and used for analysis (Hare 1997). As such, modeling of the process itself must be developed from the parameters and autocorrelation structure of the single realization (McDowall et al. 1980, Hare 1997). As the underlying model of a time series is constructed using the mean, variance, and correlation structure, there must be constancy in these parameters through time (Hare 1997). A time series with constant parameters is said to be "stationary."

Stationary time series, or the idea of "stationarity", is critical for the application of TSA. In terms of correlation structure, stationarity also means that the impacts of past observations or errors decrease quickly in time (Hartmann et al. 1980). Strictly stationary series vary consistently about a constant mean, and have a mean, variance, and covariance that are not dependent on time (Box and Jenkins 1976, Jensen 1976, Hartmann et al. 1980). Strict stationarity is often too strong for most real-world time series, and assessing it is difficult (Shumway and Stoffer 2006). Weak stationarity, which refers to only the mean and variance, is an acceptable alternative (Shumway and Stoffer 2006). Stated simply, the mean and variance of a weakly stationary series do not change over time and their relationships are based on relative and not absolute position in time (Hare 1997). In a weakly stationary series, other observations collected at different points in time would result in the same correlation structure between values (Hare 1997). Here, the term 'stationary' refers to weakly stationary for simplicity.

In reality, raw time series data are rarely stationary (Hartmann et al. 1980, McDowall et al. 1980, Shumway and Stoffer 2006). Real data often exhibit fluctuations around different means and changing variance, although the general underlying behavior of the series may be relatively stable over time (Box and Jenkins 1976). Methods for handling nonstationary data, including models as described below, are common in TSA (Box and Jenkins 1976).

Simple ways to model time series data include the autoregressive model and the moving-average model (Box and Jenkins 1976, Hartmann et al. 1980). In autoregressive models, the observation at time *t* can be expressed in terms of the previous observations and an error term (Box and Jenkins 1976, Shumway and Stoffer 2006). "Autoregressive" refers to the fact that the current value is regressed, or dependent, on the previous values in the same series (Box and Jenkins 1976). The linear model of an autoregressive time series is,

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + w_t$$

where ϕ is the impact of past observations, x_t denotes the past observation at time *t*, and w_t is a white noise error process at time *t* (Shumway and Stoffer 2006). Autoregressive models are of order *p*, which denotes the number of autoregressive parameters, or ϕ s, in the model (Box and Jenkins 1976). For example, autoregressive modeling of fisheries data would express the catch at time *t* as a function of the catch at *t-1*, etc. (Jensen 1976).

Moving average models occur when the current observation is dependent on aggregations of past shock, or error, in the series, not past values (Box and Jenkins 1976). A basic form of a moving average model is,

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \ldots + \theta_q w_{t-q}$$

with θ_q denoting the impact of past error terms, w_{t-q} (Shumway and Stoffer 2006). These models are of order q, indicating the order, or number, of moving average parameters (θs) (Shumway and Stoffer 2006).

ARIMA Models

Combining both autoregressive and moving average models into a mixed autoregressive-moving average model can achieve greater flexibility when describing time series data (Box and Jenkins 1976). A common form of these is the *autoregressive-integrated-moving average* (ARIMA) model. ARIMA models were initially developed and popularized by Box and Jenkins in 1970 (Shumway and Stoffer 2006). They are more adaptable than autoregressive or moving-average models alone because they include three structural parameters: an autoregressive element, a moving average element, and an integration, or differencing, element (Box and Jenkins 1976). To summarize these aspects, the models are designated as ARIMA (p,d,q) models. The order of the autoregressive component is represented by parameter p, the order of integration by d, and the order of the moving average component by parameter p (Box and Jenkins 1976, Velicer and Colby, 2005). In practice, most models of real data rarely have p, d, or q values greater than 2 (Box and Jenkins 1976).

The differencing or integration parameter, d, in ARIMA models addresses nonstationary series (McDowell et al. 1980). It is commonly applied to series with stochastic behavior to remove trends and stabilize the mean (Hare 1997). Differencing, denoted as the integrating factor $\nabla^d x_t$, subtracts the first value from the second, the second from the third, and so on, hence $\nabla^d x = x_t - x_{t+1}$ (McDowall et al. 1980, Hare 1997). The *d*'th difference stabilizes a nonstationary mean (Box and Jenkins 1976). Nonstationary data may also require variance stabilization, which can be achieved with square root, natural log, or Box-Cox power transformations (Shumway and Stoffer 2006).

ARIMA models entail several assumptions. The first is a stationary series with parameters that are also stationary and not dependent on time. To achieve this, the parameters must be inside the unit circle, i.e. they must be a fraction between -1 and 1 (McDowall et al.1980, Velicer and Colby 2005). This maintains the parameter within the "bounds of stationarity" and its impact will decrease quickly with time (Box and Jenkins 1976, McDowall et al. 1980). In more detail, parameters of one equate into past values or shocks with the same weight regardless of position in time (McDowall et al. 1980). They therefore do not diminish, result in behavior that is "perfectly predictable," and the series is thus nonstationary (Velicer and Colby 2005). Parameter values greater than one or less than negative one are past values or shocks that become increasingly important as time passes, also resulting in a nonstationary process (McDowall et al. 1980). Finally, the correlation structure is also stationary and not dependent on time. Such structure is the same if observations were taken at different intervals or time points and the series is invertible (Shumway and Stoffer 2006).

Additional assumptions refer to the error, or shock, values of a series. These must have a zero mean, i.e. mean $(a_t) = 0$, a constant variance, be independent with a covariance structure of zero (covariance$[a_t a_{t+k}] = 0$), and must be normally distributed (a_t ~ N) (McDowall et al. 1980). In sum, the error terms must be normal and independently and identically distributed (iid) (McDowall et al. 1980, Shumway and Stoffer 2006).

Building ARIMA (p,d,q) models

Building ARIMA models involve several basic steps: 1) plotting the data to look for obvious patterns, 2) investigating stationarity of the series and possibly transforming the data, 3) identifying possible ARIMA parameters, 4) goodness of fit and diagnostics, 5) model choice and parameter estimation (Box and Tiao 1975, Box and Jenkins 1976, Shumway and Stoffer 2006). Although ARIMA models can be flexible to address many additional analyses, such as those in Chapters 2 -4 of this thesis, the underlying model building process is the same.

Initial plotting involves visual inspection for anomalies and possible patterns, and nonstationary behavior. Indications of a nonstationary series include inconstant variance, obvious trends, or changing means. To achieve stationarity, the data may be transformed to stabilize both the mean, using linear regression residuals or differencing, and variance, using natural log, square-root, and Box-Cox power transformations (Shumway and Stoffer 2006). Autocorrelation functions (ACF) and partial autocorrelation functions (PACF) further aid in the detection of nonstationarity. Slow decay in a plot of the ACF versus lag signifies that the dependence among the observations is not decaying to zero quickly enough (Chen and Tiao 1990, Shumway and Stoffer 2006). This suggests that differencing is necessary and the process is deemed nonstationary

(Hare 1997, Shumway and Stoffer 2006). Visual inspection of plotted data, once transformed, can help determine if the transformations are sufficient.

When the data is stationary, the ACF and PACF plots (referred to here as ACF and PACF) suggest possible ARIMA models based on remaining significant correlation between and within lags. The ACF looks at between lag structure, the PACF within lag (Shumway and Stoffer 2006). This structure provides evidence for the *p* (AR) and *q* (MA) orders of the ARIMA model, allowing for hypotheses about possible model choice (Hartmann et al. 1980). Table 7 summarizes the behavior used for this evaluation. If the process is simply white noise, no significant lags should be evident in the ACF or PACF (Hartmann et al. 1980).

	AR(*p*)	**MA(*q*)**	**ARMA(*p,q*)**
ACF	Exponential decay (tails off)	Cuts off after lag *q*	Exponential decay (tails off)
PACF	Cuts off after lag *p*	Exponential decay (tails off)	Exponential decay (tails off)

Table 7. ACF and PACF behavior for hypothesizing possible ARIMA parameters.

The behavior of the ACF and PACF may not always be clear. In addition, many models may be similar in nature and not significantly different, or better, fits for the data (Hare 1997, Shumway and Stoffer 2006). For these reasons, precision about model choice is not imperative at this time in the ARIMA model process (Shumway and Stoffer 2006). Several different models can be hypothesized, fit to the data, and compared. Goodness of fit tests help determine final model choice. There are several available that can be applied as part of the model diagnostics.

The first, Akaike's Information Criterion (AIC), developed by Hirotsugu Akaike, ranks models by balancing the error of the fit versus the number of parameters in each model (Akaike 1974, Shumway and Stoffer 2006). The model with the lowest AIC score is the best fit, although values must be several points different to be significantly different statistically. Alternatives include the Bias Corrected AIC (AICc) developed by Sugiura (1978) and Hurvich and Tsai (1989), which corrects AIC for smaller sample distributions, and Schwartz's Information Criterion (SIC) (Shumway and Stoffer 2006). AICc is best for small sample sizes, especially with a relatively large number of model parameters. It converges to AIC with larger samples and can therefore be employed to all sample sizes (Burnham and Anderson 2004). SIC, also known as the Bayesian Information Criterion (BIC) or Schwartz's Bayesian Criterion (SBC), utilizes a correction term based on Bayesian statistics (Georgakarakos et al.2006, Shumway and Stoffer 2006). It performs well at large sample sizes and often chooses models of smaller order than AIC or AICc (Hare 1997, Shumway and Stoffer 2006).

Once goodness of fit tests establish the best model choice, additional diagnostics determine if the model meets ARIMA analysis assumptions. These include standardized residuals that are marginally normally distributed, uncorrelated and iid with mean = 0 and variance =1 (Box and Jenkins 1976, Shumway and Stoffer 2006). To investigate if the residuals meet these assumptions, visual inspections and tests are performed. Histograms, normal probability plots, or Q-Q plots can also reveal serious departures from a normal distribution. A time plot of the standardized residuals should show no obvious pattern or numerous significant outliers, as this may suggest residual correlation (Shumway and Stoffer 2006). Further tests for correlation include an ACF and a Ljung-Box-Pierce Q-statistic of the residuals tests. The Q-statistic addresses correlation cumulatively. Correlation can be considerable collectively even when individual residual autocorrelations are small (Shumway and Stoffer 2006). This will be illustrated by significance in a plot of Q-statistic.

After a model is chosen based on goodness of fit and diagnostics, model parameters can be estimated using statistical software. Here, R (R Development Core Team 2008) and S-Plus FinMetrics (Insightful Corporation 2007) statistical software packages were used. The fitted model can then be used for additional analysis and forecasting future values (Box and Jenkins 1976, Shumway and Stoffer 2006).

Estimated parameters denote how the series behavior changes from one time point to the next. For stationary data, these parameters will be inside the unit circle and they are comparable to a correlation coefficient. A parameter equal to zero means there is no dependency in the data. Positive parameters indicate that the behavior at time $t + 1$ will be in the same direction as at time t, and negative parameters indicate the behavior at $t + 1$ will be in the opposite direction (Velicer and Colby 2005).

SARIMA(p,d,q)x(P,D,Q) models

Cyclic tendencies or seasonal patterns are common in time series data (Cyrer and Chan 2008). Such patterns can induce another level of correlation at reoccurring seasonal lags and requires additional statistics for appropriate modeling (Shumway and Stoffer 2006, Cryer and Chan 2008). Stochastic seasonal, or SARIMA, models, work well for these series (Cryer and Chan 2008). The fundamental difference between ARIMA and SARIMA models is the periodicity of the data. Seasonal time series data exhibits cyclic behavior of a basic time interval: period s (Box and Jenkins 1976). Yearly seasonal data, such as the oceanographic data here, often reflects monthly periodicity, i.e. $s = 12$ (Box and Jenkins 1976, Tiao 1983). Simply stated, for $s = 12$, observations in one month are correlated on the observation 12 months prior (i.e. January data is correlated with, or similar to, the January a year ago). Mixed SARIMA models exhibit seasonal correlation and the non-seasonal operators of ARIMA models, i.e. lags close in time are also correlated (Shumway and Stoffer 2006). All models addressed here are mixed models. Such data is common in physical, biological, and economic systems (Chatfield 1977, Shumway and Stoffer 2006).

SARIMA models expand ARIMA theory by allowing for the identification and incorporation of seasonal lags. Mixed SARIMA models are designated ARIMA(p,d,q) x (P,D,Q), with p,d,q the non-seasonal ARIMA structure and P,D,Q the seasonal aspects. Both are hypothesized from the correlation structure of ACFs and PACFs. Correlation behavior at multiples of the seasonal lag (s) suggests orders for $P,D,$ and Q. Table 8 gives the behavior used to make conjectures about the orders of P,D and Q,. Behavior between seasonal lags indicates p,d,q parameters (Table 7) (Shumway and Stoffer 2006).

	$AR(P)_s$	$MA(Q)_s$	ARMA (P, Q)
ACF	Tails off at lags of multiple *s*	Cuts off after lag *Qs*	Tails off at lags of multiple *s*
PACF	Cuts off after lag *s*	Tails off at lags of multiple *Qs*	Tails off at lags of multiple *s*

Table 8. ACF and PACF behavior for SARIMA model parameter estimation.

SARIMA modeling approaches are very similar to ARIMA techniques. Goodness of fit tests and diagnostics are the same, and ACFs and PACFs are used to evaluate stationarity and postulate model structural parameters. However, nonstationarity for mixed seasonal data results from correlated lags close in time and correlation at multiples of *s* (Box and Jenkins 1976, Shumway and Stoffer 2006). Both correlation behaviors are indicated by plots of the ACF and PACF (Tables 7 & 8). If either exists, differencing at lag = 1 and seasonal differencing at the seasonal lag = *s* induces stationarity (Shumway and Stoffer 2006). Seasonal differencing subtracts the observation at *s* lags in the future: $\nabla^s x = x_t - x_{t+s}$ (Shumway and Stoffer 2006).

Backcasting

TSA requires many observations to be robust. Obtaining numerous observations over time often results in missing values in the time series (Velicer and Colby 2005). However, time series analysis itself does not allow for missing values. If they exist, observations must be interpolated from the existing data. Methods for doing so include simply deleting the missing value ("deletion"), using the mean of a series, using the mean of adjacent values, and maximum likelihood (Velicer and Colby 2005). Another method is backcasting, which utilizes the forecasting ability of TSA in reverse. In backcasting, an ARIMA model is initially fit to a complete series of observations after any missing observations. The data is then reversed and the model forecasts for the number of missing values. The resulting series, including the predicted values, is reversed again, and the predicted numbers provide for missing observations back in time, i.e. "backcasting". ARIMA models require invertible time series (Shumway and Stoffer 2006), making this process feasible. Here, backcasting interpolates the missing values for the herring and oceanographic data.

In this chapter, time series analysis (TSA) modeled the time series of herring landings from Chapter one. ARIMA approaches, described in the General Methods section of the Introduction, fit models to each annual series. Seasonal ARIMA (SARIMA) models fit monthly oceanographic data. The resulting models were used to backcast missing data and complete these time series. Models for the complete series were necessary for further analysis and prediction. Oceanographic data included sea surface temperature (SST) and salinity for St. Andrews Biological Station (SABS) from 1924-2007. All three were recorded by month as seasonal time series and required an additional methodological approach: Seasonal ARIMA (SARIMA) models.

Methods

Time series analysis and ARIMA models were applied to the four time series data sets from Chapter One: Maine herring, Canadian herring, St.

Andrew's Biological Station (SABS) SST, and SABS salinity. ARIMA models fit to the annual Maine and Canadian herring data backcast for missing observations to complete these series. The oceanographic data (SABS SST and salinity) were seasonal in nature and recorded per month. SARIMA models fit this data and backcast for missing monthly observations. Once the data set was complete, oceanographic data was averaged by year for annual aggregation. Final ARIMA models were fit to all the complete annual series. R statistical software applied these methods (R Development Core Team 2008).

Preliminary Data Exploration

To begin, each time series was loaded as a separate list into the R statistical analysis software. Initially, only the longest series of complete observations was used to avoid missing values. The software converted the lists into time series objects, and determined stationarity via visual inspection and sample autocorrelation (ACF) and partial autocorrelation (PACF) functions. When series were deemed nonstationary, natural log-transformations stabilized the variance and differencing stabilized the mean. Seasonal differencing removed seasonal nonstationarity in the oceanographic data.

Model Building – ARIMA (p,d,q)

Results of ACFs and PACFs on the stationary herring data and determined possible ARIMA model forms. All possible models were modeled in R using the arima() command, accounting for any differencing of the raw data.

Goodness of fit tests, including Akaike's Information Criterion (AIC), in addition to biased-corrected AIC (AICc) and Schwarz's Information Criterion (SIC), compared these models for final model choice. Diagnostics run on the final model residuals determined if they met model assumptions of normal distribution and iid. These diagnostics included a histogram and Q-Q plot to investigate marginal normalcy, and a time plot, ACF, and Ljung-Box-Pierce Q-statistic to address residual correlation.

Model Building - Seasonal ARIMA (SARIMA) *(p,d,q)x(P,D,Q)*

Because it was recorded by month, the oceanographic data was loaded into R as a time series with a frequency of 12. Once entered, a visual examination of time plots and ACFs and PACFs determined if the data was stationary. This seasonal data required inspection of both correlation between lags close together and correlation at multiples of the seasonal lag, s = 12.When the data was nonstationary, natural log-transformations stabilized variance and differencing and seasonal differencing stabilized the mean. ACFs and PACFs of the stationary data indicated possible orders for model parameters *p,d,q* and *P,D,Q*. The R software ran the resulting hypothesized models using the arima() command. The model fit tests and diagnostics used for ARIMA models determined model fit and if the model residuals met assumptions.

Backcasting for Complete Time Series

Once a model for each time series was established, it was used to backcast for missing values. The time series were reversed and the data transformed when necessary to achieve stationarity. Previously chosen models were refit to the reversed and stationarity series to ensure validity. Once confirmed, the model "predicted" for missing years on the reversed time series via the R command pred().These results were back-transformed when necessary to remove differencing, seasonal differencing, and natural log-transformations. The series was then reversed and backcasted values replaced missing observations in the original series. Repeating the process iteratively supplied values for all missing years. Once complete, ARIMA models, goodness of fit tests, and diagnostics were run a final time on the series to confirm correct model choice and estimate model parameters. These parameters were back-transformed (i.e. removal of natural log and differencing) for models of the complete herring pounds time series.

Appendix E contains script for the above analysis.

Results

Preliminary Data Exploration

Figures 4 and 5 provide time plots of the raw herring data for Maine and Canada. ACF and PACF plots of both indicated nonstationarity, specifically the slow decay in the ACF (see Fig. 6 for example). Natural log-transformations and differencing induced stationarity in both series (Fig. 7 - 8).

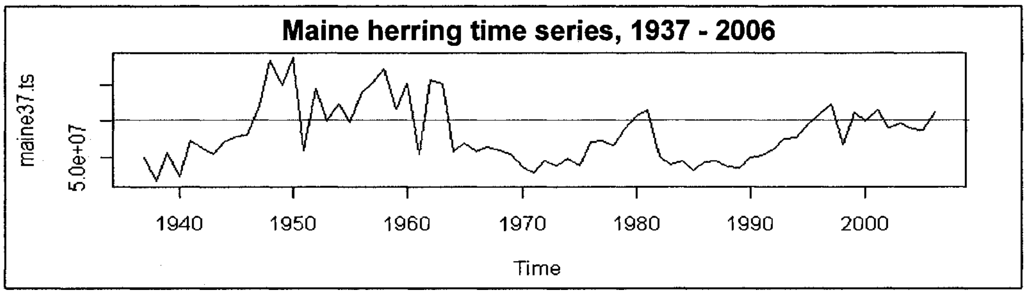


Figure 4. Time plot of the Maine herring time series generated in R statistical software.

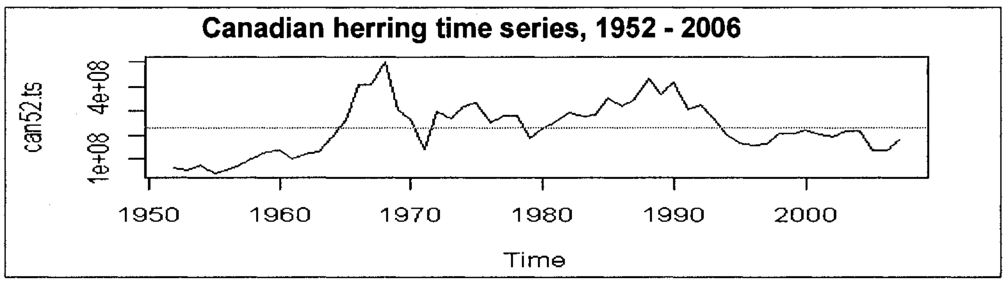


Figure 5. Time plot of the Canadian herring time series generated in R statistical software.

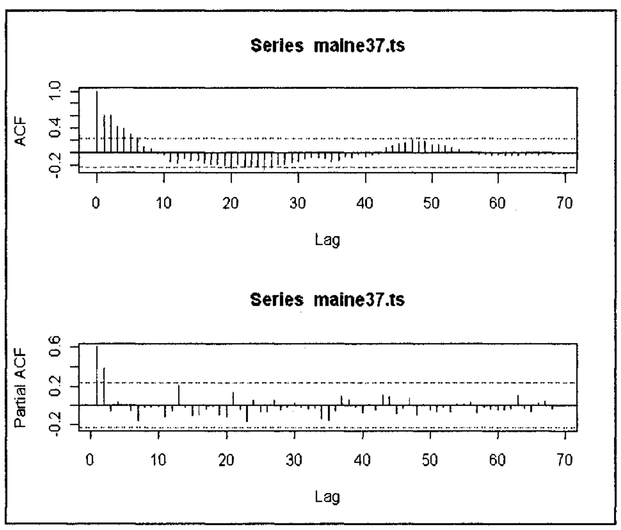


Figure 6. An example of nonstationary behavior: the ACF and Partial ACF of the raw Maine herring data, 1937 – 2006.

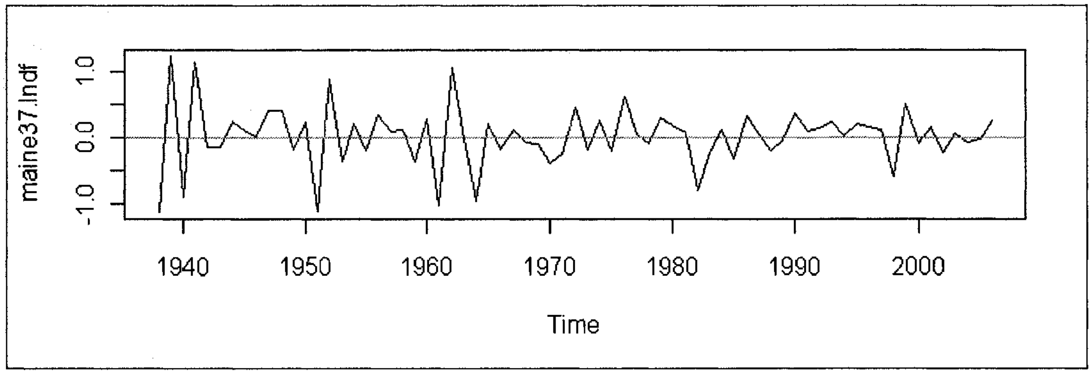


Figure 7. Natural log-transformed and differenced data for the Maine herring, 1937-2006.

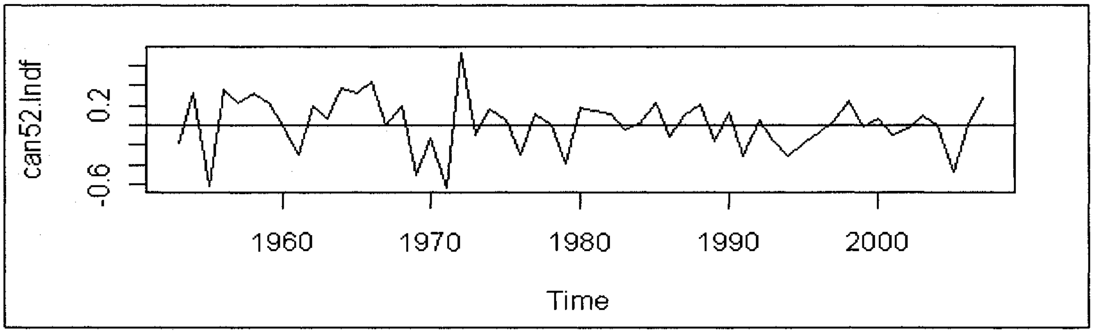


Figure 8. Natural log-transformed and differenced data for the Canadian herring, 1952-2006.

As with the herring data, the oceanographic data was plotted over time (Fig. 9 – 10) and initial ACFs and PACFs determined stationarity. Both exhibited oscillating correlated lags in their ACF, indicating nonstationary seasonal time series and recommending seasonal differencing. The overall ACFs were not slowly decreasing, indicating normal differencing was not needed. Once seasonally differenced, the ACFs and PACFs of both series confirmed stationarity.

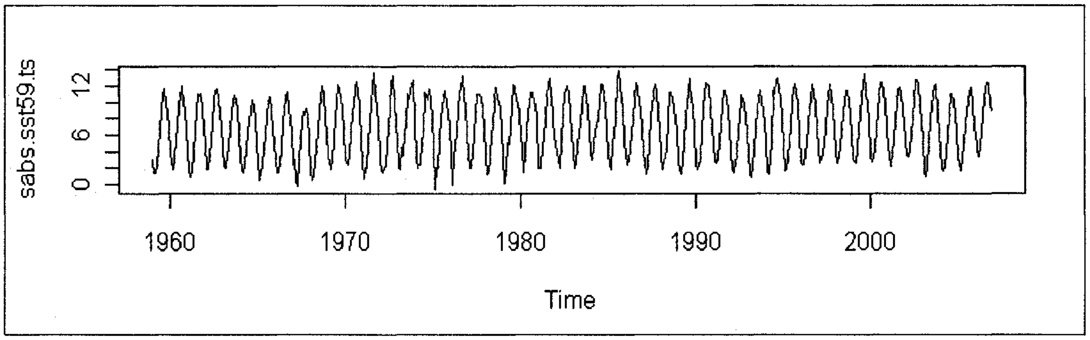


Figure 9. Time plot of the SABS SST data by month, 1983 – 2006.

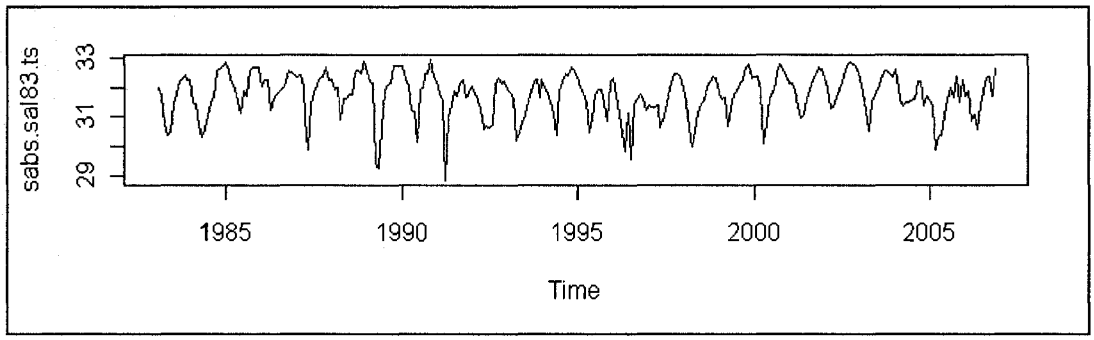


Figure 10. Time plot of the SABS salinity data by month, 1983 – 2006.

Maine herring pounds

Model determination. The ACF for the transformed and stationary data appeared to spike at lag = -1 and the PACF to dampen out exponentially. This led to an initial model hypothesis of a moving average, or ARIMA (0,1,1), for the natural log-transformed data (Fig. 11). As alternatives, an ARIMA(1,1,0) was considered, due to the significant lag at -1 in the PACF, as well as an ARIMA (1,1,1), in case both series were tailing off.

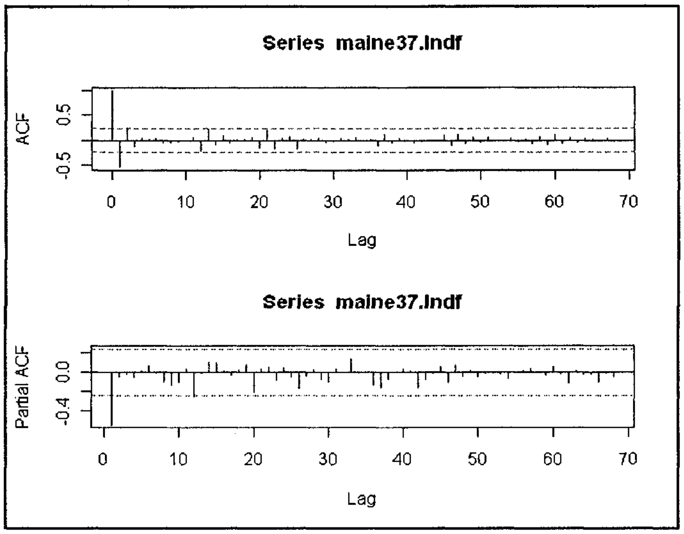


Figure 11. ACF and Partial ACF of the ln-transformed and differenced Maine data.

Goodness of fit & diagnostics. Despite initial conclusions, all goodness of fit tests (AIC, AICc, and SIC) chose the ARIMA(1,1,0) as the best model fit for the ln-transformed Maine data (table 9). The ARIMA(1,1,0) was not a statistically better fit than the ARIMA(1,1,1), but was significantly better than the ARIMA(0,1,1). The ARIMA(1,1,0) was chosen as the final model for parsimony, i.e. fewer parameters, and because it had lower values for all three tests.

	AIC	**AICc**	**SIC**
ARIMA(0,1,1)	73.16	-1.161626	-1.807333
ARIMA(1,1,0)	**67.92**	**-1.239409**	**-1.885116**
ARIMA(1,1,1)	69.89	-1.222361	-1.824123

Table 9. Results from goodness of fit tests for the Maine herring data.

Residual diagnostic tests did not present significant departures from model assumptions, indicating the ARIMA (1,1,0) model appropriate for Maine herring (Fig. 12). The histograms and Q-Q plots of the residuals exhibited a left-skewed distribution, but it was not of serious concern. The time plot did not reveal any obvious pattern and, despite a few outliers, rarely exceeded two standard deviations in magnitude. No significant correlation between residual lags was apparent in either the ACF or the Ljung-Box-Pierce Q-statistic. Finally, models parameters were all within the unit circle (i.e. between -1 and 1). All tests therefore confirmed the model as a good fit to the data and met ARIMA model assumptions. Table 10 summarizes the model and initial estimated parameters.

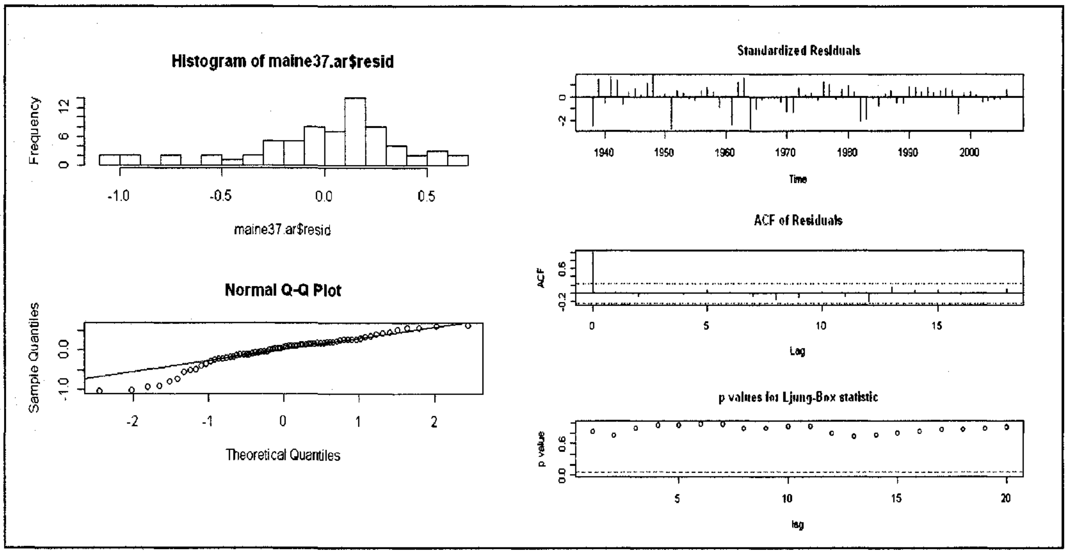


Figure 12. Results of residual diagnostic tests for the Maine herring data. At left are the histogram and Q-Q plot of the residuals, at right the time plot, ACF, and p-values for the Ljung-Box-Pierce statistic.

	AR parameter	s.e.	Intercept	s.e.	σ^2
ARIMA(1,1,0)	-0.5836	0.129	0.0167	0.0289	0.1428

Table 10. Parameter estimation for the ARIMA(1,1,0) model of Maine herring, 1937-2006.

Canadian herring pounds

Model determination. The ACF for the stationary Canadian herring data had no significant lags to suggest possible ARIMA model orders (Fig. 13). Therefore, a white noise process, ARIMA(0,1,0) was run in addition to ARIMA(0,1,1), ARIMA(1,1,0) and ARIMA(1,1,1) for comparison.

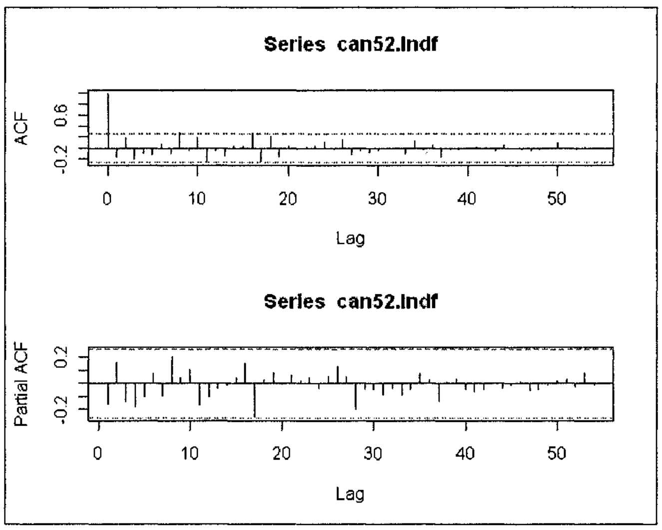


Figure 13. ACF and Partial ACF of the ln-transformed & differenced Canadian herring data.

Goodness of fit & diagnostics. AIC and AICc goodness of fit tests chose the ARIMA(1,1,1) as the best model fit for the natural log-transformed Canadian herring data, but the SIC gave a lower value to the white noise process (Table

11). However, the model output for the ARIMA(1,1,1) estimated parameters too close to the unit circle for both *p* and *q* (*p* = -0.991, *q* = 0.9396, table 2e). The ARIMA(0,1,0) model, being barely a marginally less adequate fit, was used for diagnostics instead. Parameters are given in table 2e.

	AIC	**AICc**	**SIC**
ARIMA(0,1,0)	12.76	-1.640783	**-2.678518**
ARIMA(0,1,1)	13.32	-1.628312	-2.632374
ARIMA(1,1,0)	13.67	-1.621877	-2.62594
ARIMA(1,1,1)	**10.84**	**-1.683285**	-2.655211

Table 11. Results from goodness of fit tests for the Canadian herring data.

Residual diagnostics revealed significantly correlated lags in the Ljung-Box-Pierce plot indicating additional parameters. The ARIMA(0,1,1), being only marginally less significant for fit, was selected (Table 12 summarizes parameters). Diagnostics confirmed the ARIMA(0,1,1) met model assumptions. The histogram and Q-Q plot revealed a slight left-skew in the residuals, but again, it was not a significant concern (Fig. 14). The residual diagnostics confirmed that there were no evident patterns in the time plot and only a few outliers. The ACF and Ljung-Box-Pierce plot revealed no additional correlation in the residuals (fig 2n). Chosen models therefore met model assumptions of normalcy and iid.

	AR (1)	s.e	MA (1)	s.e.	Intercept	σ^2
ARIMA(1,1,1)	-0.9910	0.0201	0.9396	0.0195	0.0195	0.06075
ARIMA(0,1,0)					0.0178	0.06866
ARIMA(0,1,1)			-0.1261	0.1215	0.0177	0.06729

Table 12. Parameter estimation for the ARIMA(1,1,1) model of the Canadian herring data.

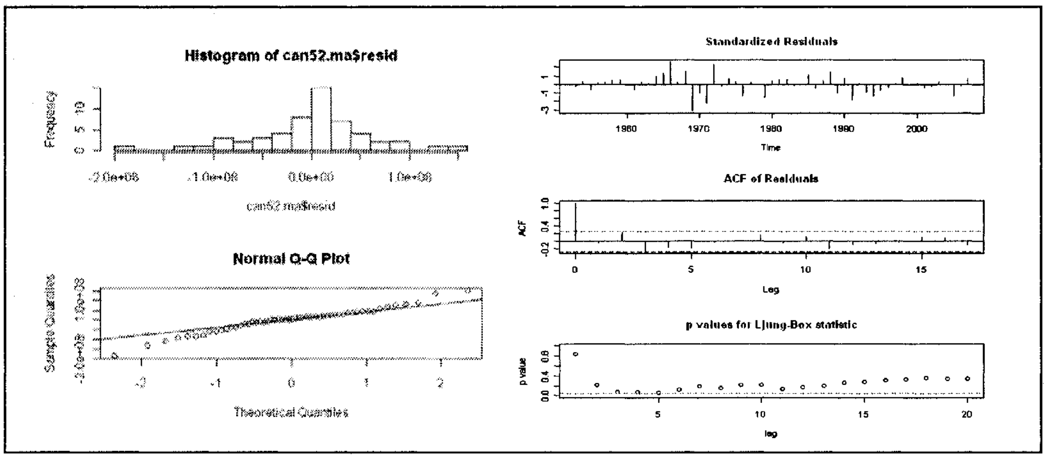


Figure 14. Results of residual diagnostic tests for the Canada herring data. Left: histogram and Q-Q plot of the residuals. Right: time plot, ACF, and p-values for the Ljung-Box-Pierce statistic.

Sea Surface Temperature (SST) and Salinity

Model determination – SARIMA. The ACFs and PACFs of both the SST and salinity stationary data indicated possible model parameter orders (Fig. 15). Between-lag behavior for *p,d,q* parameters indicated either ARIMA(1,0,0) or ARIMA(1,0,1) processes. For *P,D,Q* parameters of the SABS SST data, the correlation behavior at seasonal lags appeared to have a significant lag in the ACF, with the PACF dampening out. The PACF did have significant lags at *ks*, with *k* = *1, 2, 3*. This suggests a seasonal MA process, but alternative models were also included for comparison. Because of the significant lags greater than one, models with orders of 2 for structural seasonal parameters were also considered.

For the salinity data *p,d,q* parameters, between-lag behavior indicated a dampening of the ACF and a spike at lag = 1 for the PACF. This suggested *p,d,q* parameters of (1,0,0) for the salinity data. Significant seasonal lags existed in the salinity series for both the ACF and PACF, with a secondary significant seasonal lag in the PACF. Alternatively, either could have been dampening out. Thus, *P,D,Q* possibilities for the data included (1,1,1), (1,1,0), or (0,1,2). Alternative models were again included for comparison.

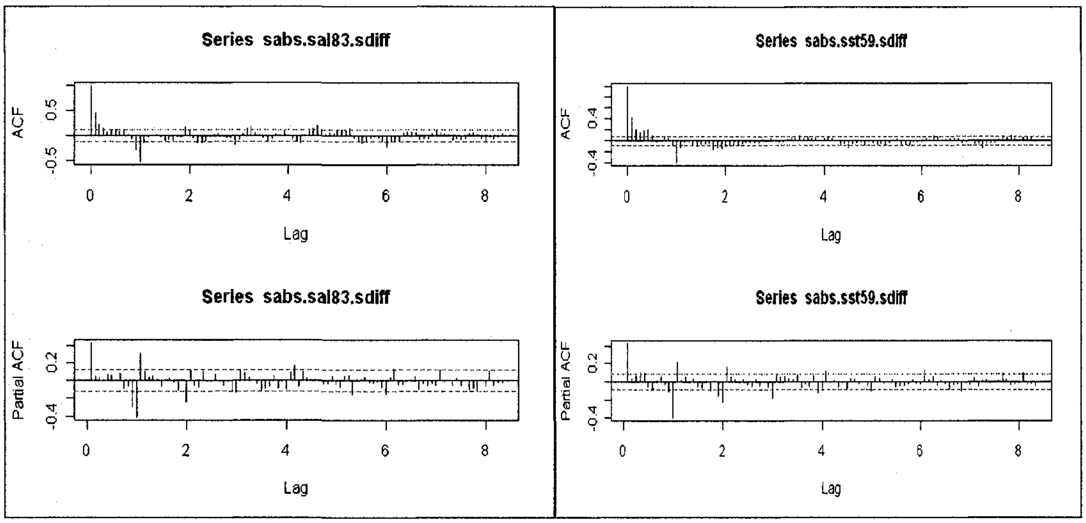


Figure 15. ACF and Partial ACF of the seasonally differenced SABS SST data (left) and salinity data (right) by month, 1983 – 2006.

<u>Goodness of fit & diagnostics.</u> Table 13 summarizes the different model fits for the SST series. AIC values were lowest for ARIMA(1,0,1)x(1,1,1), but

were not significantly different from several others. Based on this, AICc and SIC fit tests were run on ARIMA(1,0,1)x(1,1,1), ARIMA(1,0,1)x(0,1,2), ARIMA(1,0,1)x(1,1,2) and ARIMA(1,0,1)x(2,1,1) for the raw data. The AICc values chose the ARIMA(1,0,1)x(1,1,2) model, but the SIC values agreed with the AIC. Given the number of observations in the series (564), SIC is a good choice for fit evaluation. However, because goodness of fit values did not differ significantly, diagnostics were run on the ARIMA(1,0,1)x(1,1,1), the ARIMA(1,0,1)x(0,1,2), and the ARIMA(1,0,1)x(1,1,2) models of the raw data (Fig. 16). Diagnostics revealed that both models had relatively normally distributed residuals, however, the ARIMA(1,0,1)x(1,1,1) and the ARIMA(1,0,1)x(0,1,2) displayed marginal correlation in the residuals, suggesting additional parameters. Not surprisingly, this correlation was gone in the diagnostics for the ARIMA(1,0,1)x(1,1,2). Because goodness of fit suggested the models were not significantly different, the ARIMA(1,0,1)x(1,1,2) was chosen and used for backcasting.

Model for raw data	**R object name**	**AIC**	**AICc**	**SIC**
ARIMA(1,0,0)x(1,1,0)	sabs.sst59.ar	1570.29		
ARIMA(1,0,0)x(0,1,1)	sabs.sst59.ma	1416.74		
ARIMA(1,0,0)x(1,1,1)	sabs.sst59.arma	1412.36		
ARIMA(1,0,0)x(0,1,2)	sabs.sst59.ma2	1412.77		
ARIMA(1,0,0)x(1,1,2)	sabs.sst59.arma2	1416.37		
ARIMA(1,0,0)x(2,1,0)	sabs.sst59.ar3	1515.46		
ARIMA(1,0,0)x(2,1,1)	sabs.sst59.arma3	1413.84		
ARIMA(1,0,1)x(1,1,0)	sabs.sst59.ar1b	1567.99		
ARIMA(1,0,1)x(0,1,1)	sabs.sst59.ma1b	1410.37		
ARIMA(1,0,1)x(1,1,1)	**sabs.sst59.arma1b**	**1407.6**	0.5755877	**-0.397404**
ARIMA(1,0,1)x(0,1,2)	sabs.sst59.ma2b	**1407.89**	0.5755384	**-0.397453**
ARIMA(1,0,1)x(1,1,2)	sabs.sst59.arma2b	**1409.6**	**0.5717702**	-0.393612
ARIMA(1,0,1)x(2,1,0)	sabs.sst59.ar3b	1511.62		
ARIMA(1,0,1)x(2,1,1)	sabs.sst59.arma3b	**1409.1**	1.597454	-0.385812

Table 13. Results from goodness of fit tests for the SABS SST data.

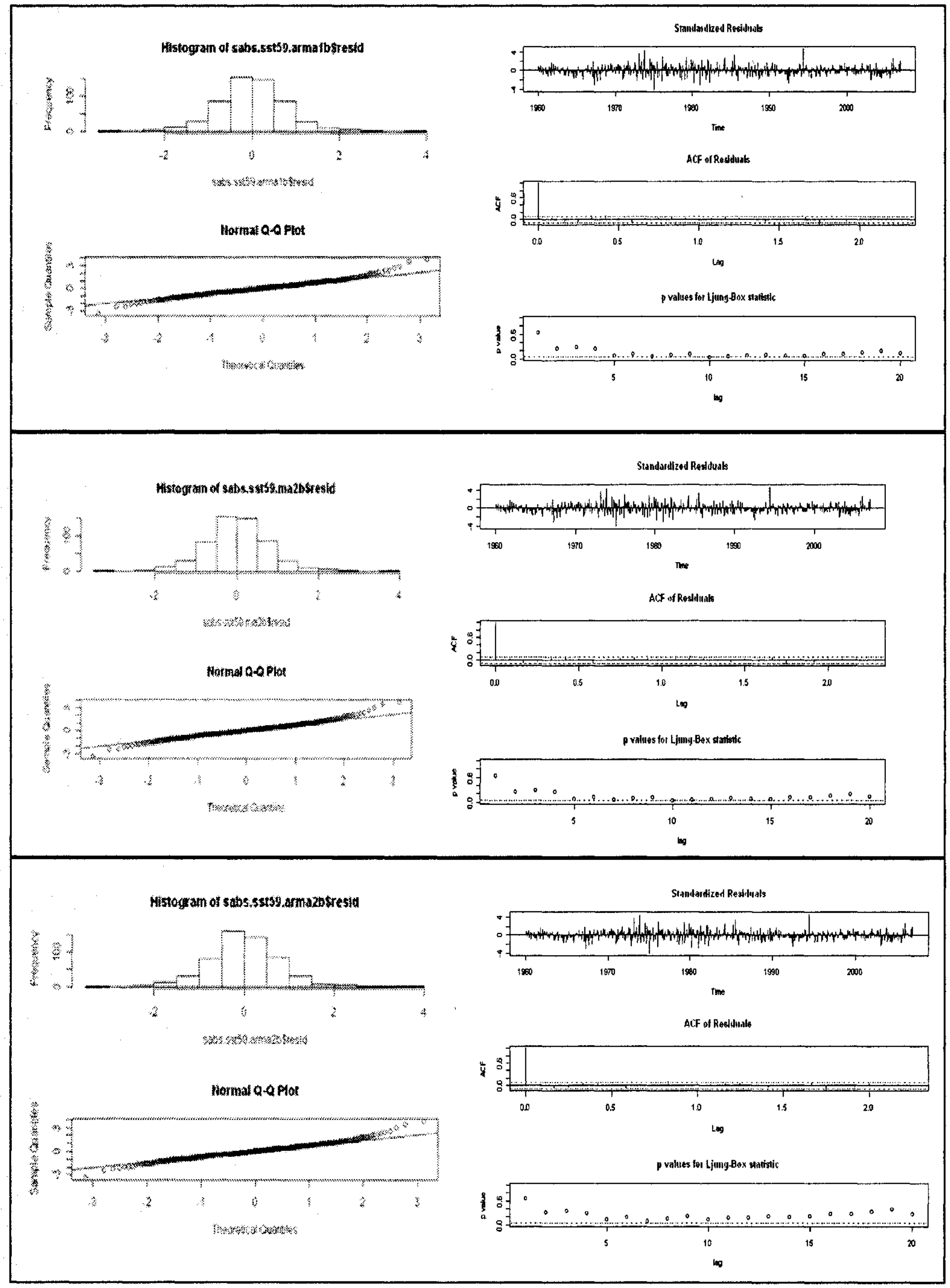


Figure 16. Results of residual diagnostic tests for ARIMA(1,0,1)x(1,0,1), top, ARIMA(1,0,1)x (0,1,2), middle, and ARIMA(1,0,1)x(1,1,2) of the raw SABS SST data. At left are the residual histograms and Q-Q plots, at right the time plot, ACF, and Ljung-Box-Pierce statistic results.

For the raw salinity data, the ARIMA(1,0,1)x(0,1,1) had the lowest AIC values (Table 14). It was not significantly different from alternative models ARIMA(1,0,1)x(1,1,1), ARIMA(1,0,1)x(0,1,2), and ARIMA(1,0,1)x(2,1,1). The SIC confirmed model choice of the ARIMA(1,0,1)x(0,1,1), but the AICc tests gave a lower value for the ARIMA(1,0,1)x(2,0,1). Again, given the number of observations (274), the SIC may be more robust. Taking these results into consideration, the ARIMA(1,0,1)x(0,1,1) was chosen for the monthly SABS salinity data. Diagnostic tests found the residuals to be normal and iid, and there were no residual correlation apparent (Fig. 17). The model was a good fit, met assumptions, and was subsequently used for backcasting.

Model	**R object name**	**AIC**	**AICc**	**SIC**
ARIMA(1,0,0)x(1,0,0)	sabs.sal83.ar	403.54		
ARIMA(1,0,0)x(0,0,1)	sabs.sal83.ma	349.39		
ARIMA(1,0,0)x(1,0,1)	sabs.sal83.arma	351.38		
ARIMA(1,0,0)x(0,0,2)	sabs.sal83.ma2	351.39		
ARIMA(1,0,0)x(1,0,2)	sabs.sal83.arma2	353.14		
ARIMA(1,0,0)x(2,0,0)	sabs.sal83.ar3	385.25		
ARIMA(1,0,0)x(2,0,1)	sabs.sal83.arma3	351.67		
ARIMA(1,0,1)x(1,0,0)	sabs.sal59.ar1b	396.98		
ARIMA(1,0,1)x(0,0,1)	**sabs.sal59.ma1b**	**344.42**	**-0.654156**	**-1.62244**
ARIMA(1,0,1)x(1,0,1)	sabs.sal59.arma1b	**345.71**	-0.646579	-1.60195
ARIMA(1,0,1)x(0,0,2)	sabs.sal59.ma2b	**345.85**	-0.646468	-1.60184
ARIMA(1,0,1)x(1,0,2)	sabs.sal59.arma2b	346.45		
ARIMA(1,0,1)x(2,0,0)	sabs.sal59.ar3b	380.69		
ARIMA(1,0,1)x(2,0,1)	sabs.sal59.arma3b	**345.42**	**-0.695506**	-1.59308

Table 14. Results from goodness of fit tests for the SABS salinity data.

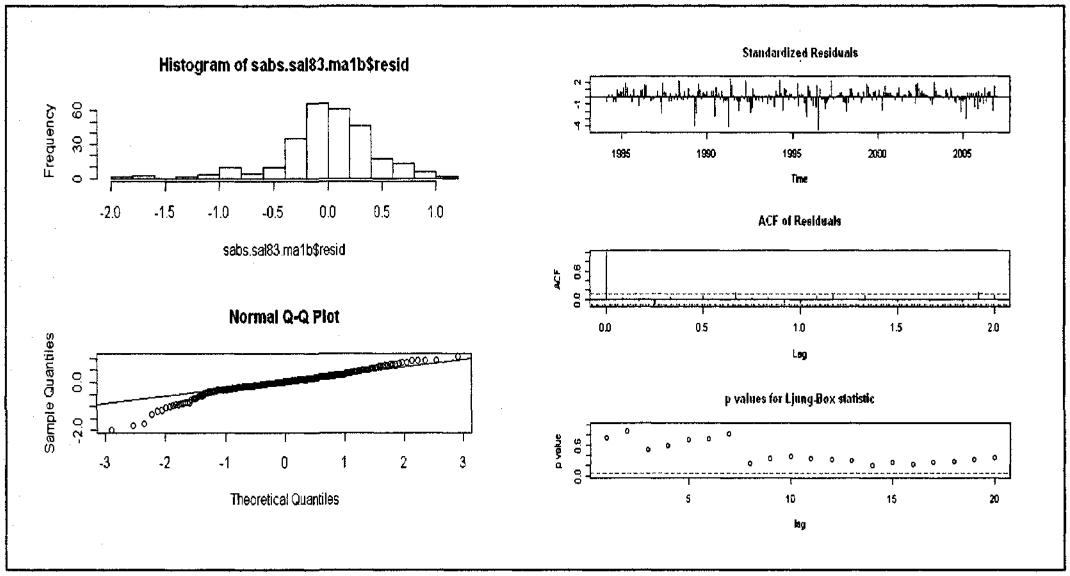


Figure 17. Results of residual diagnostic tests for ARIMA(1,0,1)x(0,1,1) of the raw SABS salinity data. At left are the histogram and Q-Q plot of the residuals, at right the time plot, ACF, and p-values for the Ljung-Box-Pierce statistic.

Backcasting

Backcasting provided all missing years for the Maine herring landings (1881-6, 1890-1, 1893-9, 1915-8, 1920-3, 1925-7, 1936). This analysis completed the Maine herring time series from 1880 through 2006 (Fig. 18). The backcasting approach also completed the Canadian series (1871-2006, Fig. 19) for all missing years (1876, 1886, 1890-1, 1897, 1901, 1903, 1928-34, 1947-51). SARIMA models and analysis on the oceanographic data filled in missing months (Fig. 20 - 21) and allowed for yearly aggregation from 1924 – 2007 (Fig. 2w-x).

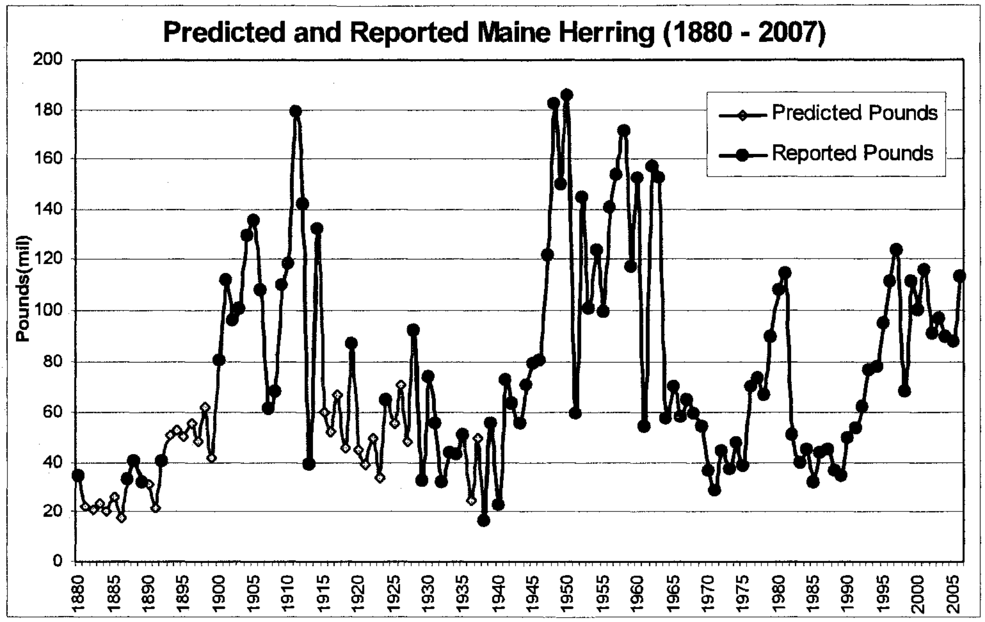


Figure 18. Completed Maine herring time series, reported and predicted values, 1871-2007.

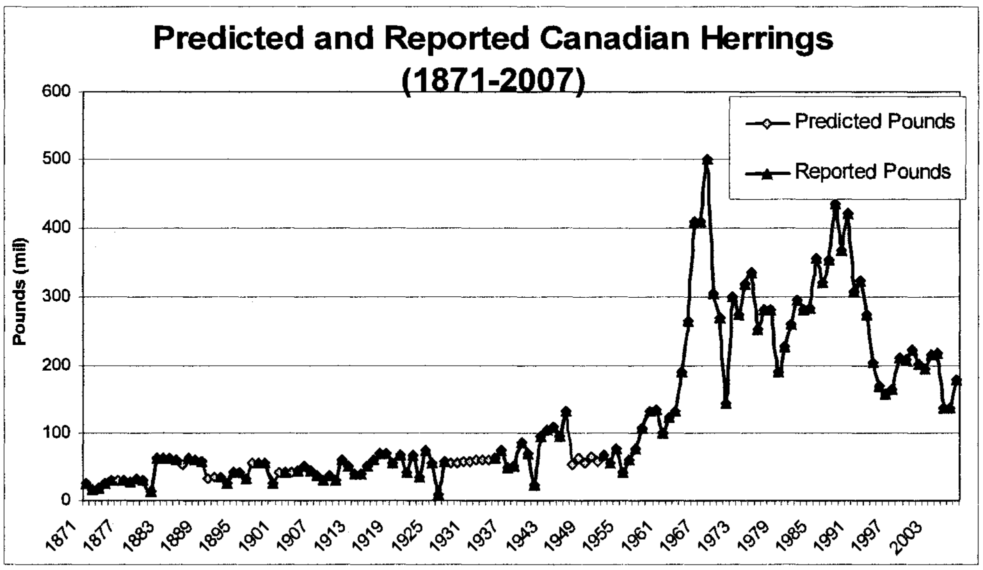


Figure 19. Completed Canadian herring time series, reported and predicted values, 1871-2007.

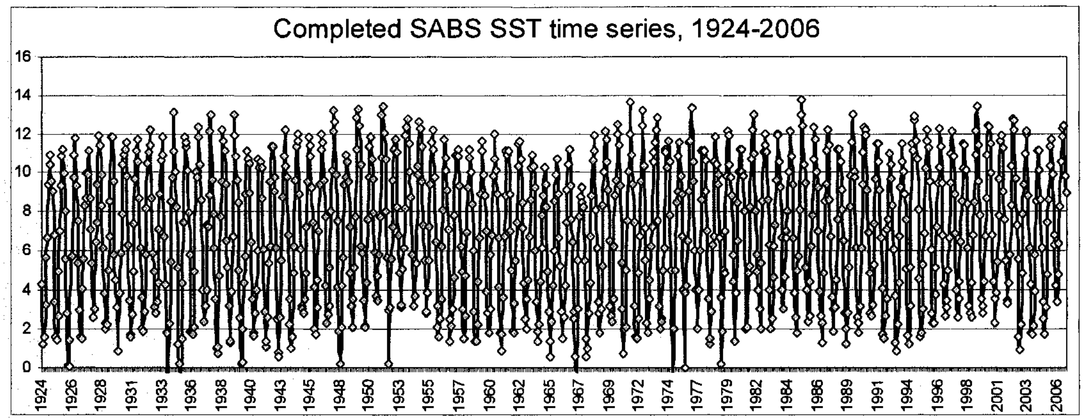


Figure 20. Completed SABS SST monthly time series, reported and predicted values, 1924-2006

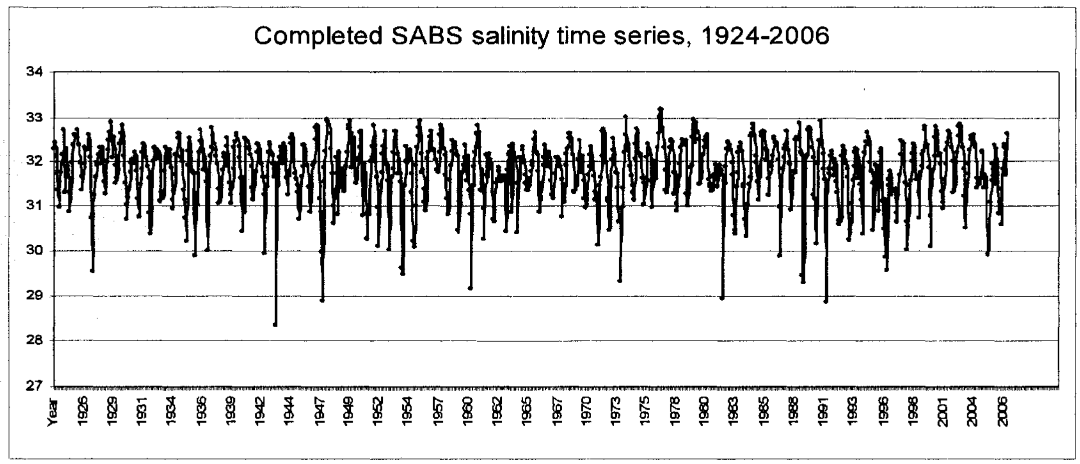


Figure 21. Completed SABS salinity monthly series, reported and predicted values, 1924-2006.

Oceanographic Annual Models

Averaging by year aggregated the oceanographic time series once missing monthly values were backcasted (Fig. 22 - 23). Once annual, an ARIMA model was fit to both series. The ACFs and PACFs of the SABS data sets indicated that the SABS salinity data was stationary and the SABS SST was nonstationary (Fig. 24). Differencing removed trend and attained stationarity in the SABS SST data.

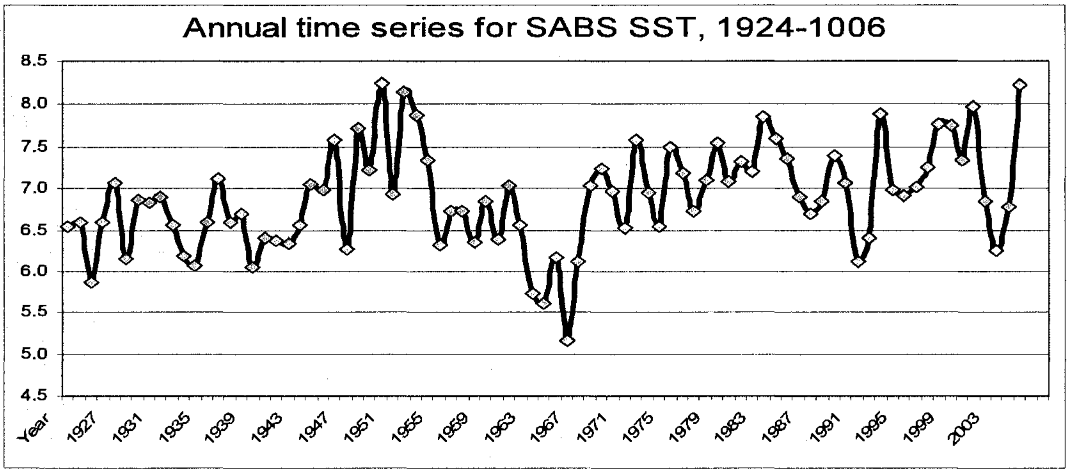


Figure 22. Completed SABS SST annual time series, reported and predicted values, 1924-2006

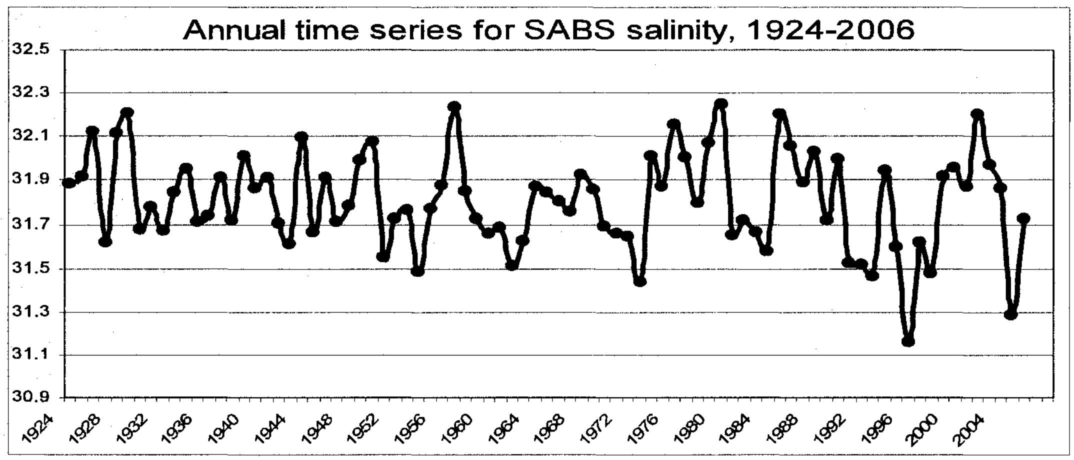


Figure 23. Completed SABS salinity annual time series, reported and predicted values, 1924-2006

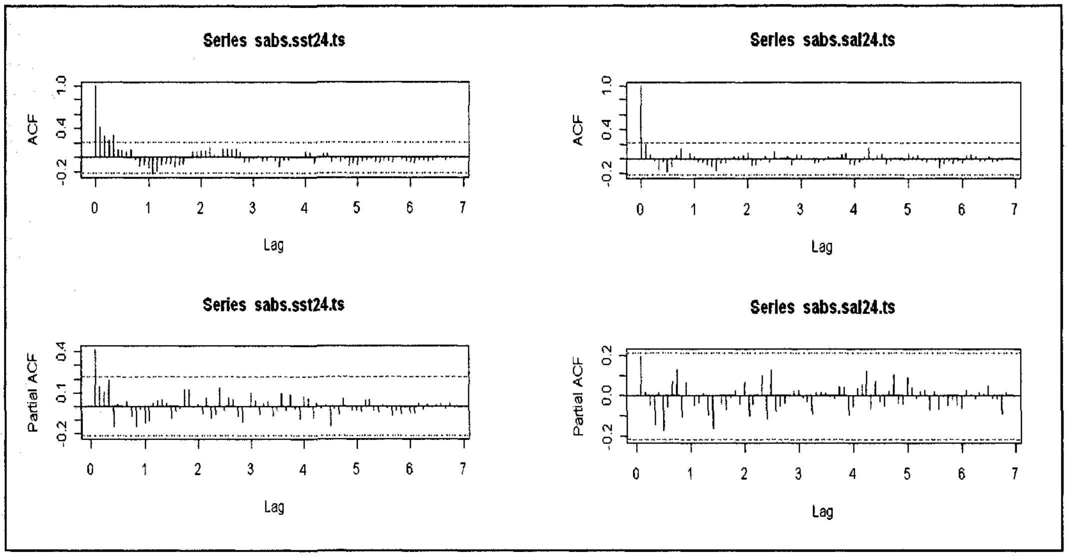


Figure 24. ACF and PACF for the SABS SST (left) and SABS salinity (right) annual time series.

Model determination – ARIMA. The ACF and PACF of the stationary SABS SST appeared to spike at lag = 1 in the ACF and to dampen out in the PACF (Fig. 25), suggesting an ARIMA(0,1,1) model for the undifferenced data. ARIMA(1,1,0) and an ARIMA(1,1,1) were also modeled for comparison. All three models in addition to a white noise process (ARIMA[0,0,0]) were run for the SABS salinity data as the ACF and PACF did not display any significant lags (Fig. 24).

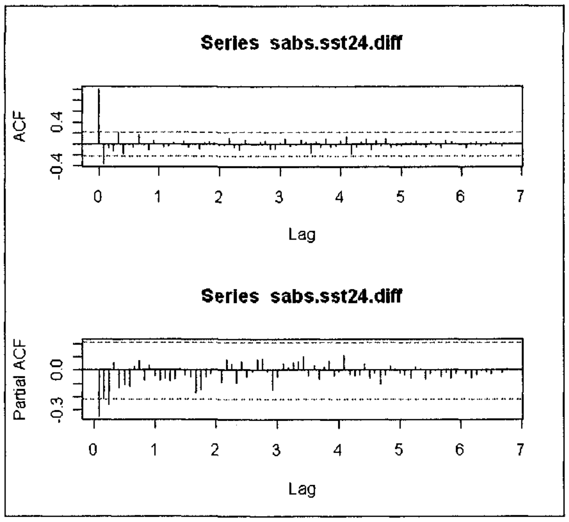


Figure 25. ACF and PACF of the stationary SABS SST time series

<u>Goodness of fit & diagnostics.</u> AIC goodness of fit tests of the SABS SST data indicated the ARIMA(1,1,0) and ARIMA(1,1,1) models were not significantly different fits, although the ARIMA(1,1,1) had a lower AIC value (Table 15). Both were statistically better than the originally hypothesized ARIMA(0,1,1). The AICc fit test agreed, but the SIC chose the ARIMA(1,1,0). An ARIMA(1,1,0) process was therefore selected for low goodness of fit values and parsimony. Table 16 provides estimated model parameters from R output.

	AIC	**AICc**	**SIC**
ARIMA(0,1,1)	**142.05**	-0.1512913	**-1.148052**
ARIMA(1,1,0)	146.91	-0.09158072	-1.088342
ARIMA(1,1,1)	**141.51**	**-0.1569078**	-1.126379

Table 15. Results from goodness of fit tests for the annual SABS SST models.

	AR (1)	s.e	Intercept	s.e.	σ^2	Log lik
ARIMA(1,1,0)	0.4404	0.1013	6.8987	0.1066	0.3008	-68.03

Table 16. Parameter estimation for the ARIMA(1,1,1) model of the raw SABS SST data.

The AIC values for the SABS salinity data indicated an ARIMA(1,0,1), but none of the other models were significantly different, or only marginally so (Table 17). Thus, according to AIC, all three models are sufficient fits to the data. AICc tests gave lower values to the ARIMA(1,0,0), yet SIC did so for the ARIMA(0,0,0) (Table 17). For parsimony, the white noise model was adopted for the SABS salinity time series. Residual diagnostics revealed both models for SST and salinity meet assumptions. Table 18 summarizes parameter estimates.

	AIC	AICc	SIC
ARIMA(0,0,0)	**-14.6**	-2.037286	**-3.061977**
ARIMA(1,0,0)	-15.75	**-2.050386**	-3.047147
ARIMA(0,0,1)	-15.53	-2.047635	-3.044396
ARIMA(1,0,1)	**-13.77**	-2.024689	-2.99416

Table 17. Results from goodness of fit tests for the SABS salinity data.

	Intercept	s.e.	σ ^2	Log lik
ARIMA(0,0,0)	31.804	0.0288	0.04503	10.88

Table 18. Parameter estimation for the ARIMA(1,0,0) model of the SABS salinity data.

Final parameter estimation & model. Once complete annual time series were available for all data sets, the statistical software estimated model parameters. Table 19 summarizes the final models for all four time series. The models are for the natural log-transformed and differenced series, and cannot be directly applied to the catch data. For final models of the herring pounds, back-transforming removed these functions (Appendix F), and the final models for herring are in Table 19.

Time series	**Model** *(for transformed data)*	σ^2	**Log-likelihood**
Maine herring pounds	$y_i = 0.010_{(0.022)} - 0.5592_{(0.0736)}\, y_{i-1} + \hat{w}_t$	0.1472	-58.25
Canadian herring pounds	$y_i = 0.0148_{(0.0139)} - 0.5480_{(0.0808)}\, \hat{w}_{t-1} + \hat{w}_t$	0.1271	-52.91
SABS SST	$y_i = 6.8987_{(0.1066)} + 0.4404_{(0.1013)}\, y_{i-1} + \hat{w}_t$	0.3008	-68.03
SABS salinity	$y_i = 31.804_{(0.0288)} + \hat{w}_t$	0.0452	10.76

Maine herring pounds	$x_t = x_{t-1} * [e^\wedge(0.01 - 0.5592(\ln(\frac{x_{t-1}}{x_t}))\,]$
Canadian herring pounds	$x_t = x_{t-1} * [e^\wedge\,(0.0148 - 0.5480 * w_{t-1})\,]$ *where* $w_{t-1} = y_{i-1(obs)} - y_{i-1(pred)}$ *and* $y_{i-1(obs)} = \ln(\frac{x_{t-1}}{x_t})$

Table 19. Summary of models for transformed data (ln and differenced Maine herring pounds, ln and differenced Canadian herring pounds, differenced SABS SST, and raw SABS salinity) and models for the herring annual pounds data (bottom).

Discussion

Time series analysis is an effective tool for identifying consistent patterns in fishery statistics. Exploring and modeling these patterns provides the basis for a variety of further analyses. Intervention analysis (Chapter 3) assesses significant outliers and level shifts that can test the response of a fishery to specific external events. These impacts can include management action, gear and technology advances, market and demand changes, or, more broadly, wars, economic recessions, etc. Cross-correlation analysis (Chapter 4) explores relationships between multiple time series, such as the effect of various oceanographic features on fishery data. Cross-correlation can also test relationships between different fisheries, presenting insight across species. It is these analyses that can begin to tease apart various pressures acting on fisheries over time. Fisheries can then be understood in a broader context that includes environmental fluctuations, the human community, and the greater ecosystem.

Models developed through ARIMA approaches can also be used to forecast future conditions, such as potential catch. TSA forecasting can be advantageous over traditional catch prediction methods as it requires only historical fishery statistics (Jensen 1976). Other methods necessitate additional and often more derived variables, such as effort, which can be less accurate and available than catch data (Jensen 1976). Using ARIMA models may therefore provide more accurate predictions of catch using more precise and accessible information. Additional approaches, such as intervention and cross-correlation analysis, can improve this ability through the development of more in-depth models.

In this chapter, analysis provided underlying ARIMA models for herring catch in both Maine and Canada, and for several oceanographic time series. These models were developed for two reasons. First, they are necessary for the intervention analysis in Chapter 3 and the cross-correlation in Chapter 4. Second, the prediction ability of TSA and the developed ARIMA models interpolate missing observations through backcasting. These values provided complete, long-term time series for all five data sets. TSA could therefore be run on many more observations, capturing more behavior over time and providing more robust conclusions.

The models themselves encourage some exploration of the correlation structure inherent in the fishery over time. The Maine herring data resulted in an ARIMA(1,1,0) model, which would indicate that herring catch in one year can be explained, at least in part, by the catch the year before. The negative coefficient on the autoregressive term indicates that if the observation in one year is above the mean, the following year it will be below (Box and Jenkins 1976, Velicer and Colby 2005, Shumway and Stoffer 2006). Removing the natural log function and differencing revealed even more information regarding the underlying pattern of this fishery. Now, pounds in one year are explained by the change in pounds over the past two years. This means that, if there was a great increase in pounds of fish caught between year *t-2* and year *t-1*, then a much smaller catch can be expect in year *t* than was seen in year *t-1.* The opposite (a very large catch in year *t*) is predicted if there is a great decrease between years *t-2* and *t-1.* If, however, catch is relatively constant, or if there are only small changes in pounds between years *t-2* and *t-1*, catch will continue to be relatively constant in year *t*. In sum, a large difference between the catch two years ago and last year predicts a similarly large difference in the opposite direction between this year and last, but relatively constant catches predict that constancy will continue. According to this model, the Maine pounds are not decreasing or increasing over time, but strive for consistency.

Such a pattern is not surprising for a small pelagic fish with a well developed fishery. This may reflect the ability of herring to respond quickly to impacts (Anthony and Waring 1980). Heavy fishing quickly shows a decrease in pounds landed as the fish responds to the intense exploitation, and a release of fishing pressure results in a rapid rebound. As a small pelagic that matures relatively quickly and for which research has shown to respond quickly to pressure (Anthony and Waring 1980), such a pattern over time makes intuitive sense. Moreover, the fishery itself is well developed. No new grounds are currently being explored nor is technology changing at a constant rate over time. Therefore, there are no reasons to assume a long-term increase. The effects of developments in terms of grounds, gear, etc., appear to be restricted to distinct and relatively short time periods. This may also reflect the fact that herring are fully exploited at this time – significant increases in catch due to changes in gear in one year result in equally significant declines in the catch as opposed to consistent increases over time.

In contrast, consider a fishery for which the coefficient is positive. This would signify that the pounds are in the same direction as the year previously, or that an increase in pounds between years *t-2* and *t-1* predicts an increase for year *t* as well. A decrease between *t-2* and *t-1* would likewise indicate a decrease for this year. In this case, the fishery would be consistently increasing or decreasing over time. An increasing fishery could be imagined as one that is moving from a small scale, inshore industry to offshore banks and with increasing technology over time. Thus, catch is expanding and the population is perhaps not yet fully exploited. If, on the other hand, the fishery is decreasing over time, the fishery may be heavily exploited and the species cannot respond rapidly enough to pressure, i.e. a species with low fecundity.

For Canada, the ARIMA(0,1,1) choice indicates that the catch is not explained by past catch, but instead more by the random error values the year previously, the term w_t. This term is the error in the natural log-transformed and differenced Canadian herring data, i.e. the difference between the observed and the predicted for the transformed data. To be clear, this is not the difference between directly observed and predicted pounds, but the difference once the data has been log-transformed and differenced. Overall, the influence of the past Canadian herring pounds is important for predicting future Canadian pounds, but it is also influenced by the error factor, w_t.

It is certainly interesting that the analysis here determined such an intuitive model for the herring fishery over the long term. Additional application of TSA for other species would determine if similar underlying processes are found for other exploited small pelagics with similar life histories. Applying these methods to other species to determine models for increasing or decreasing fisheries, as described above, is also of significant interest. Results do indicate the need for further analysis and the possible potential of TSA for long-term fishery data.

Further conclusions cannot be drawn from the models themselves. Fisheries are impacted by numerous and convoluted drivers that change in time, and further analysis is required to begin defining them. The persistence of the herring models is intriguing, given the long time period and incredibly wide array of influences presumably acting on the fishery during this time. However, it is important to keep in mind that the data has been transformed to remove changing means and variances, and that the models interpolated missing values. Nonetheless, the specific ARIMA parameters were maintained throughout and conclusions were statistically significant. Finally, despite the ability to derive additional conclusions from these models, they were necessary for backcasting to complete the series. In addition, the underlying models are required for further analyses and prediction. Two such additional methods are carried out in Chapters 3 and 4, although prediction is left for future research.

CHAPTER III

INTERVENTION ANALYSIS

Introduction

Most time series analysis methods assume that observations are generated from a consistent pattern structure (Chen and Tiao 1990). In reality, time series data may not behave this way, as actual processes are under pressure from many external influences (Chen and Tiao 1990, Zivot 2006). Unexpected changes in time series resulting from these influences are well documented (Scheffer et al. 2001, Zivot 2006). Correction for these effects maintains robust statistical methods and accurate conclusions and predictions (Chen and Tiao 1990). However, time series can have such variations and still be considered stationary (Hare 1997). Therefore, it is possible to run ARIMA models and acquire significant results without accounting for these unexpected impacts. It is when the models are applied more widely to additional analysis or prediction that these impacts can become a concern (McDowell et al. 1980).

Unexpected changes in time series data manifest as outliers or level shifts in the data, cumulatively referred to here as interventions (Chen and Tiao 1990). Very simply, an intervention can either change the direction of the series or alter the series level by changing the parameters, i.e. mean or variance by some amount (Glass 1972). Some authors discuss interventions as “transfer functions,” because they transfer the level, slope, or both of a series from one state to another (Box and Jenkins 1976, Hartmann et al. 1980). This literature also addresses the series exhibiting the change, or “transfer,” together with the affecting series. For example, Box and Jenkins (1976) discuss transfer functions in terms of an input series that causes the intervention, and an output series where change due to the intervention is exhibited. Other authors discuss only the series displaying the interventions and describe methods to understand the interventions independently of the series or event that caused them (Box and Tiao 1975, Chen and Tiao 1990). Interventions are, however, the result of a cause and effect relationship, even if the only aspect under consideration is when and how the effect is felt.

Interventions are revealed in time series data in variety of ways. The literature generally defines their behavior by how quickly the effect is expressed (onset) and how long it is sustained (duration) (Hartmann et al. 1980, McDowall et al. 1980, Hare 1997). The intervention can appear suddenly or gradually, and can be continuous or temporary (Hartmann et al. 1980, McDowall et al.1980, Hare 1997). Box and Jenkins (1976) discuss the behavior in detail, and in terms of a *step* response, where the impact is sustained, or an *impulse* response, where the impact is temporary.

Methodology exists to detect unexpected changes (interventions) within an ARIMA model pattern. This approach is commonly referred to as intervention analysis, but is also known as transfer function modeling (Box and Jenkins 1976), interrupted time series analysis or impact assessment (Hartmann et al.

1980, McDowall et al. 1980). Researchers have applied it widely in many fields, especially business, economics, law, and the behavioral and social sciences. Intervention analysis can determine if events such as marketing campaigns, laws, industry improvements (Box and Jenkins 1976), clinical interventions, or experimental manipulations (Hartmann et al. 1980) cause significant changes in a time series. It has been less widely applied to the biological sciences (Murtaugh 2000).

Assumptions of intervention analysis are similar to more basic TSA approaches. It requires many observations, at least 40 are recommended, and a stationary series (Box and Jenkins 1976, Hartmann et al. 1980, Chen and Tiao 1990). If the time series is nonstationary, large deviations resulting from random behavior may appear as outliers when no intervention exists (Chen and Tiao 1990). Inducing stationarity and properly fitting a TSA model distinguishes between such random changes and actual interventions (Chen and Tiao 1990). Therefore, intervention analysis is run after the data is stationary and has been fit to a model (Hartmann et al. 1980, McDowall et al. 1980).

Once the series is stationary, a visual examination of the data plotted over time can help determine if intervention analysis is necessary beyond an ARIMA fit. The plot reveals possible significant outliers and where they occur in time. A model is then fit to the series, and the analysis uses this model to investigate possible interventions. This can be thought of as a secondary model:

$$Y_t = f(I_t) + N_t$$

where N_t signifies the ARIMA fit and $f(I_t)$ represents the intervention. The null hypothesis is that $f(I_t)$ does not have a statistically significant impact on the series. The null is rejected when $f(I_t)$ increases the overall explanatory power of the model (McDowall et al. 1980).

In this chapter, intervention analysis investigates whether socioeconomic and fishery events were reflected in Maine and Canada herring fishery production. Intervention analysis expands the TSA methods of Chapter Two to begin assessing external fishery drivers, in this case possible socioeconomic and industry events. This approach also incorporated the qualitative timeline (Chapter One) into quantitative analysis. Intervention analysis identifies significant shifts in the herring landings, but the qualitative literature provides possible explanations for such abrupt and unexpected change. Therefore significant social, industrial, political, etc., and other historical events can be compared to the landings in a meaningful way.

The S-Plus FinMetrics (Insightful Corporation 2007) module includes an intervention analysis tool for time series data. This software identifies the type and location of interventions in data secondary to the general ARIMA(p,d,q) model. It detects three types of behavior: 1) additive outliers, 2) innovation outliers, and 3) level shifts. Additive outliers (AO) are impacts restricted to a specific time period. Innovation outliers (IO) are not restricted to a time period and can have an effect on subsequent observations. Finally, level shifts (LS) change parameters of the model to a new state, although the underlying behavior remains the same and the new parameters are consistent (Zivot 2006).

The intervention analysis in FinMetrics builds on the idea of regression ARIMA, or REGARIMA, models, which combine regression techniques and ARIMA models. REGARIMA models expand regression analysis to tackle cause and effect in time series data. They use ARIMA methods to consider serially correlated errors specific to time series data when testing if one system, the input, influences a second system, the output. However, ARIMA and REGARIMA models may not be robust for data with outliers or level shifts. FinMetrics addresses this concern through robust change ARIMA models. These models can handle data with outliers and level shifts. They give a more accurate estimate of model parameters and more rigorous model fits than ARIMA or REGARIMA models alone when such interventions exist. Robust change ARIMAs pinpoint where interventions occur and can clean these interventions from the data to provide more accurate forecasting (Zivot 2006).

The FinMetrics approach is similar to procedures developed in Chang et al. (1988), Tsay (1988), and those used by the U.S. Census Bureau (Zivot 2006). Chang et al. (1988) applied likelihood ratio criteria to detect and distinguish between innovational and additive outliers. Tsay (1988) applied least squares and residual variance ratios techniques on univariate data to detect outliers, level shifts, and variance changes. These procedures are fairly simple and are widely applicable to various data (Tsay 1988). FinMetrics combines both approaches to evaluate innovative and additive outliers and level shifts. The main difference with FinMetrics is the use of robust change models and innovation residuals. These residuals are based on filtered estimates of model parameters using log-likelihood, not classical maximum likelihood estimates (Zivot 2006). Bianco et al. (2001) developed these estimates, called τ-filtered estimates, for REGARIMA models.

The FinMetrics module detects an outlier by computing a test statistic and a critical value. When the test statistic is greater than the critical value, an outlier is detected (Zivot 2006). The time and type of the outlier are determined where the double maximum of the test statistic is attained (Zivot 2006). Critical values are usually dependent on the number of observations in the data set, but can be arbitrarily established (Zivot 2006). They are similar to a constant used in Chang et al. (1988). These authors recommend a value of 3 for high sensitivity and less than 200 observations, 3.5 for medium and 200-500 observations, and 4 for low sensitivity in outlier detection and greater than 500 observations (Chang et al. 1988). The FinMetrics module uses these for the default critical values, but they can be changed in the command script (Zivot 2006).

Methods

Landings data for Maine and Canada were loaded into S-Plus and converted into time series objects. Next, natural log transformations and differencing induced stationarity when necessary in the time series prior to the analysis. Once the data was stationary, the arima.mle() command determined an ARIMA(p,d,q) model. AIC goodness of fit tests determined final model choice and this was compared with the results from R.

Once the appropriate ARIMA models was confirmed, the arima.rob() command in the FinMetrics module initiated intervention analysis on both stationary herring time series using the model (*p, q*) orders from ARIMA fits. The command runs robust change ARIMA models to determine additive and innovative outliers and level shifts. The results provided year and type of outlier for each time series, and were compared to the qualitative timeline for possible explanations of outliers found.

Appendix F contains all script.

Results

Both Maine and Canadian herring pounds were nonstationary according to ACF and PACF results. To induce stationarity, natural log transformations stabilized the variance and differencing the mean. Following these transformations, a second set of ACF and PACF plots confirmed stationarity. Once the data was stationary, S-Plus FinMetrics fit ARIMA models to the data and provided AIC results. These goodness of fit tests chose the ARIMA(1,1,0) for the raw Maine herring data, confirming the results from Chapter Two. AIC chose ARIMA(0,1,1) for the raw Canadian data, also confirming earlier results. Both of these models were used for the robust ARIMA and intervention analysis.

For Maine, the robust ARIMA revealed three interventions, summarized in Table 20. The model determined all detected outliers to be significant (t(126) = 3.184, 4.197, 4.139). The interventions occurred in 1951, 1964, and 1982. Comparisons with the qualitative timeline revealed possible correlations between events and all three interventions. 1951 was the least informative, with references only stating that catches declined during the 1950s (Anthony and Waring 1980). In 1964, the USSR was fishing with otter trawls but diverting attention from herring to other species on Georges Bank, and the Nova Scotia adult purse seine fishery began with subsequently large catches (Anthony and Waring 1980). In 1982, herring were placed on a prohibited species list, rendering their landing by foreign fleets illegal within the U.S. Exclusive Economic Zone (EEZ) ("Atlantic herring" 2008). Also in this year, the National Marine Fisheries Service (NMFS) withdrew the 1978 herring fishery management plan (FMP) ("Atlantic herring" 2008).

Year	Type	Impact	t-value
1951	IO	-1.003	3.184
1964	AO	-0.9661	4.197
1982	AO	-0.9772	4.139

Table 20. Interventions detected in the Maine herring data.

The intervention analysis results for the Canadian data were quite different from Maine. The analysis found 67 interventions over the course of the time series, beginning in 1874. Table 21 summarizes these results and the possible events correlated. To focus on fewer interventions, the analysis was run a second time with a critical value of 4 to reduce sensitivity. This analysis found 8 interventions (Table 22). All interventions were significant (see Tables 21-22 for t-values).

Year	Type	Impact	t-value	Event
1874	AO	0.2852	4.945	First attempt at Russian sardines
1875	IO	0.1483	9.025	Sardine industry begins in Eastport Duty of $4 on imported sardines
1881	IO	-0.7588	4.227	*None*
1882	AO	1.386	6.715	*None*
1886	IO	-0.1484	8.643	Canadian law passed against seining Canneries organize to regulate prices Sardine industry overcrowded Bad fire in Eastport destroys several factories Law prohibiting canning after Dec. 15 in effect Birch-bark for torching decreased locally, fishermen using cotton and kerosene, weir fishermen complaining about oil in water Many improvements since 1880 Frozen herring/winter herring trade, fish abundant
1890	AO	-0.5776	3.906	Frying fish in oil replaced by cooking with steam Solder principle expense for canning industry (4) Smoked herring at low point
1893	IO	-0.3655	3.882	Electricity introduced in factories
1894	IO	0.4433	3.584	*None*
1896	IO	-0.2369	5.29	Reported as very bad season – no reason given
1897	IO	0.4834	3.788	Canning season shortened by 40 days, in addition to other regulations Reported as a bad year, some factories closed early
1900	IO	-0.8382	4.548	Sea Cost Packing gains control of AM Can CO. plant Russian sardine industry at height – then declines Height of Eastport population – declined from here
1901	IO	0.3948	3.836	New fish driers replacing reel ovens
1906	IO	-0.1541	7.877	American Can Co regains control of Sea Coast ~1905-1910: Boats equipped with water-tight tanks
1907	IO	-0.1966	7.033	French labels no longer used, Pure Food & Drug Act goes into effect Child labor laws enacted, but not heavily enforced N. Lubec American Can Co plant burns Sea Coast Canning builds new canning plant 1907-08 in Eastport All cans are 2-piece drawn cans from American Can

Table 21. Interventions detected in the Canadian herring data.

Year	Type	Impact	t-value	Event
1908	IO	-0.2321	5.474	Flaking machines experimented with and in use Railroad comes to Eastport – competes with Boston steamship line
1909	IO	0.1541	7.543	*None*
1910	IO	-0.2784	5.296	*None*
1911	IO	0.6637	4.01	*None*
1913	AO	-0.2982	4.399	Practically no Russian sardines produced by this time 1913-16: Some efforts to improve quality – study of packing process Majority of canneries formed assoc to better industry ME sanitary legislation expanded and inspectors more numerous
1915	AO	0.2234	5.864	Great decrease in French sardines End of year embargo on Norwegian sardines Maine enacted laws restricting seining
1917	IO	0.1486	8.281	Punch type machine discarded for 2 spindle Max Ams machines
1919	IO	-0.2302	5.723	*None*
1921	IO	-0.4906	3.732	Dept of Agriculture publishes recommendations All carriers converted to gas by this time Foreign supply limited Fish meal in US – on the rise/encouraged
1922	IO	0.4214	3.771	Catches low
1923	IO	-0.6646	3.908	Catches low
1924	IO	0.6635	4.12	Catches low
1925	IO	-0.2517	5.039	Catches low
1926	IO	-1.755	8.088	Catches low
1927	IO	1.575	7.43	*None*
1936	IO	0.1412	9.442	Power project dropped by government.
1937	AO	-0.4259	3.683	*None*
1939	IO	0.4919	3.636	World War II
1940	IO	-0.1975	6.762	World War II
1941	IO	-1.136	5.85	Official end of Great Depression, US enters WWII
1942	AO	1.289	6.41	World War II
1945	IO	-0.1698	7.605	World War II
1946	IO	0.3	4.205	*None*
1947	IO	-0.9144	4.832	*None*
1953	IO	-0.202	6.335	1950s Catches decline (Anthony and Waring, 1980)
1954	IO	0.2839	5.141	Catches decline
1955	IO	-0.6164	4.044	Catches decline
1956	AO	0.2938	4.624	Catches decline
1957	IO	0.2053	6.159	Catches decline
1958	AO	0.3318	4.049	Catches decline
1959	IO	0.1982	6.495	*None*

Table 21 *cont.* Interventions detected in the Canadian herring data.

Year	Type	Impact	t-value	Event
1961	IO	-0.3192	4.066	Georges Bank fishery begins, intense pressure from USSR (gill nets)
1962	AO	0.1441	9.215	*None*
1964	AO	0.3721	3.793	USSR diverts attention from herring to other species 1964-8 USSR primarily fishes with otter trawls
1965	IO	0.2904	4.776	Nova Scotia adult purse seine fishery beings Great increase in catch of adult herring off NS 1964-8 USSR primarily fishes with otter trawls
1966	AO	0.4338	3.632	1964-8 USSR primarily fishes with otter trawls Poland begins fishing for herring on Georges Bank
1968	IO	0.1671	7.827	Peak of Georges Bank fishery, catch declines after USSR introduces purse seines into the fishery
1969	AO	-0.5228	3.646	*None*
1970	IO	-0.154	8.217	*None*
1971	IO	-0.6704	3.836	German Democratic Rep. intro's mid-water trawling
1972	IO	0.6586	4.204	ICNAF begins management of adult fisheries Change to midwater gear – possibly resulting from ICNAF quotas First national catch quotas, management "begins in earnest"
1976	IO	-0.3125	4.176	U.S. declares 200 mile limit in FMP 1976-78: NMFS regulates foreign fishing via preliminary FMP
1979	IO	-0.4174	3.886	*None*
1985	IO	0.1923	7.513	Georges Bank herring population begins to rebuild
1988	IO	0.1827	7.469	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers)
1989	IO	-0.1819	7.788	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers)
1991	AO	-0.3345	3.908	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers) Improvements in assessment procedures, single stock complex
1993	IO	-0.1941	7.254	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers)
1994	AO	-0.3594	4.011	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers) ASMFC adopts new FMP Mid-water trawling by U.S. and Canada begins
1995	AO	-0.2132	6.144	Collapse of U.S. nearshore fixed gear fishery Mobile gear gains in importance (purse seines, midwater trawlers)
1998	IO	0.219	6.005	*None*
2005	AO	-0.4976	3.569	*None*
2007	AO	0.2391	5.051	*None*

Table 21 *cont.* Interventions detected in the Canadian herring data.

Year	Type	Impact	t-value	Event
1881	IO	-0.7588	4.227	*None*
1882	AO	1.386	6.715	*None*
1900	IO	-0.8382	4.548	Sea Cost Packing gains control of AM Can CO. plant Russian sardine industry at height – then declines Height of Eastport population – declined from here
1926	IO	-1.755	8.088	Catches low
1927	IO	1.575	7.43	Catches low
1941	IO	-1.136	5.85	Official end of Great Depression US enters WWII
1942	AO	1.289	6.41	World War II
1947	IO	-0.9144	4.832	*None*

Table 22. Interventions detected with lower sensitivity in the Canadian herring data.

Discussion

The possible correlations between events and the fishery in Maine are tentative. The negative innovative outlier would corroborate declining catches reported for herring during the 1950s (Anthony and Waring 1980). As an innovative outlier, it would affect herring pounds outside of the initial onset in 1951. In 1964, the negative intervention may be the result of USSR fleets focusing on species other than herring (Anthony and Waring 1980). The analysis determined this intervention as an additive outlier, limiting its impact to that year only. According to records, the USSR concentrated on hake and haddock from 1963 – 65 (Anthony and Waring 1980), therefore an innovative outlier would seem more likely. However, other countries joining the foreign fleet at this time may have removed the negative effect. For 1982, while placing herring on the prohibited species list is a probable cause for decreased herring landings in that year, it is not definitive. It is questionable because the intervention is an additive outlier, limiting its impact to that time period only. However, the event is a long-term management act, and could be speculated as more likely to cause an innovative outlier or level shift. The other event in 1982, removing the FMP, seems more likely to increase landings, not decrease them. The intervention, on the other hand, is negative, meaning an unexpected decrease.

In Canada, the eight interventions detected once sensitivity was decreased were in 1881, 1882, 1900, 1926, 1927, 1941, 1942, and 1947. There are no events in the qualitative timeline to explain interventions in 1881 and 1882. None of the events specific for 1900 would explain the negative outlier in that year. It could be due to the rise and fall of the syndicates over that time period. Prior to 1900, the market was overcrowded, production was falling, and prices were low. Syndicates to control the market were attempted, but failed in 1903. The negative innovative outlier in 1926 could reflect the economic state of the U.S. at this time and the downward trend in the industry following WWI. The positive additive outlier in 1927 would contradict this trend, but it is for that year only. In 1941, the negative innovative outlier completely contradicts the official end of the Great Depression and WWII. These events should have resulted in a positive innovative outlier, as demand increased. This may explain the additive outlier in 1942. The negative innovative outlier in 1947 would reflect the decline in the herring industry following WWII.

A shortcoming of the analysis here is the inability to detect lagged relationships, where the response of the fishery is not seen for a year or longer.

Such results are probable but impossible to test unless the lag length is known. Instead, only speculation about potential lagging responses is possible by looking at events prior to the intervention. The late 1940s saw a slump in the herring industry after WWII, an effect that could explain the outlier in 1951. The 1960s was a time of increasing catches due to heavy offshore exploitation, yet this would result in a positive response, not a negative one. In the early 1980s, the fishery could still be reacting to the collapse in 1977 of the Georges Bank fishery and subsequent management action.

The above conclusions, however, are purely speculative. This analysis can only detect interventions, not determine cause. There are many other possible reasons for unexpected change in the series. One additional effect may be the conversion of products in the Canadian data, especially because the units and products change in time. Misreported products or biases in the conversion factors may be detected as interventions. For example, a new unit may overestimate herring pounds during its use over several years, resulting in a positive innovation outlier. In addition, certain products may be misreported, or not at all, for a period of time, resulting in a negative outlier. Here, once source for interventions in the time series would be the products dropped from the analysis (canned in cases and kippered cans in cases) because no conversion factors could be found for them. However, herrings canned in cases first appear in the statistics in 1910 and end in 1920. There is an unexplained innovation outlier in 1910, but this is positive, not negative. The loss of pounds due to dropping this product would be a negative intervention. There was no outlier detected in 1920. Kippered cans in cases are in the record from 1898 to 1912. The analysis found no interventions for either year.

Although the results from this chapter are indeterminate at best, they do highlight the complex nature of fishery landings over time. Presumably the reasons behind changes in landing are more complicated than comparing possible events to fishery fluctuations. Addressing one cause, such as socioeconomic or fishery-related events, cannot describe the entire picture. However, although it may not reveal direct correlations, this does not mean absence of influence, and additional scrutiny is needed. Such further analysis can include hypothesis testing for specific years, as opposed to investigating the entire series for possible interventions. Testing specific hypotheses about the influence of a particular year may yield different results. Alternatively, the interventions found here may correlate with interventions in environmental series, or changing products, as described above. Finally, using a more in-depth or hierarchical model, perhaps including environmental variables, etc., for this analysis may provide more certain results. One of the advantages of TSA is the flexibility to include new aspects and rerun earlier models.

Despite the absence of more conclusive results, detecting interventions is necessary for future analysis of time series data. This allows outliers to be removed from the series, which can be critical for prediction. Doing so ensures that the interventions, which are not part of the underlying correlation structure, do not influence forecasted values (McDowall et al. 1980). Therefore, the work here is significant in its own right and should not be overlooked.

CHAPTER IV

CROSS CORRELATION

Introduction

In reality, time series processes do not exist in isolation. They are influenced by various drivers and are capable of affecting other systems. Relationships can occur at a single point in time as interventions (Chapter Three) or continuously as unidirectional or bidirectional associations. In unidirectional correlations, one series, the input, affects a second series, the output (Tiao and Tsay 1983, Hare 1997). Bidirectional, or feedback, relationships exist when either system can affect the other (Tiao and Tsay 1983, Hare 1997). Deterministic inputs can cause step, pulse, and sinusoidal changes in the output series (Box and Jenkins 1976). Dealing with input and output series jointly allows more information to be included and can reduce the variances in each series alone (Tiao and Tsay 1983). This ensures more rigorous conclusions and accurate forecasting than modeling series independently (Tiao and Tsay 1983). For these reasons, understanding these cause and effect interactions can be vital in business, economic, or policy decisions (Tiao and Tsay 1983, Chan et al. 2004), as well as for fisheries management (Garcia et al. 2007).

Traditional approaches for investigating relationships are often unsuitable for time series data (Box and Jenkins 1976, Shumway and Stoffer 2006).

Classical regression techniques may not capture the entire pattern structure and are generally incapable of handling correlation in time series data (Shumway and Stoffer 2006). Further, they may be unable to find an output response masked by noise (Box and Jenkins 1976). TSA allows for both correlation and noise, making the identification of dynamic interaction structure more obtainable (Box and Jenkins 1976). However, ARIMA models alone are not sufficient because they do not explicitly address the influence of external variables (Tiao and Tsay 1983). Multivariate, or cross-correlation, time series analysis evaluates two or more series at once for possible interactions between them, addressing both unidirectional and bidirectional relationships. These approaches build on basic TSA, thus retaining the ability to manage time series data, correlation, and noise (Box and Jenkins 1976).

Cross-correlation TSA is useful in determining relationships between fishery statistics and oceanographic features (Kim et al. 1997), information important to fishery science (Beamish and Mahnken 1999). Although management cannot control environmental conditions, a clear appreciation for the interaction between fish species and the marine environment is significant in directing fishery policy (Jonzen et al. 2002). Understanding these relationships can be critical for determining abundance from catch statistics, which is in turn important for management actions, such as setting catch quotas. When oceanographic factors are influential, landings may have an unpredictable relationship with abundance (Kim et al. 1997). Finally, cross-correlation can help clarify lagged relationships between environment and fish recruitment and year class strength (Kim et al. 1997). Multivariate ARIMA models are therefore useful in fishery science, but have only been employed to a limited extent (Hare 1997, Allen et al. 2006, Garcia et al. 2007).

Multivariate TSA is not straight forward; there are concerns with its application that must be attended to. Similar autocorrelation structure within time series can result in falsely significant correlations that are not actual cause and effect patterns (Katz 1988, Box et al. 1994). Further, cross-correlation is an exploratory analysis only (Hare 1997). Significant results from multivariate analysis are simply correlations or hypotheses of possible relationships; they do not intrinsically identify deterministic relationships (Hare 1997, Garcia et al. 2007). These results are important for forecasting, but further conclusions require additional analysis (Hare 1997, Garcia et al. 2007). Both concerns have been addressed in other fields, namely statistical and economics, but much less so in fisheries and oceanography, where these problems are commonly overlooked in the literature (Hare 1997).

Here, multivariate TSA jointly analyzed the herring and oceanographic time series to assess relationships between the herring fishery and chosen oceanographic features (SST and salinity). The S-Plus FinMetrics module provided the tools to do so via a time series linear regression model. The software utilizes ordinary least squares (OLS), based on minimizing the sum of the squared residuals. This method develops the S-Plus linear model to handle time series regression. Assumptions of OLS include non-trending, or stationary, regressors, the absence of endogenous regressors (variables explained within the series itself), and that the error term is serially uncorrelated with a constant variance (Zivot 2006).

Methods

The S-Plus statistical software and the FinMetrics module provided the tools ran cross-correlation analysis on the herring and oceanographic data. The transformations from Chapter Two rendered all series stationary prior to cross-correlation analysis. The FinMetrics command OLS() looked for significant correlation between the herring pounds and SST or salinity. Unilateral relationships were investigated only, with SST and salinity possibly influencing herring pounds and not vice versa. The OLS command accounted for correlated lags (i.e. AR terms) in the dependent (herring) and in the independent (SST) series when those individual models exhibited AR (p) parameters. Finally, the Newey-West correction was applied to accommodate possibly similar autocorrelation structure. This correction specifically addresses possible serial correlation and heteroskedasticity in the random error term (Zivot 2006).

As with basic ARIMA models, correlation techniques use diagnostic checks to determine if results meet model assumptions. Common methods used here are the Durbin-Watson statistic and the Ljung-Box-Pierce Q-statistic discussed in previous chapters. The Durbin-Watson statistic looks for serial correlation based on the estimated residuals. Values range between 0 and 4, with 2 indicating no serial correlation, those less than 2 suggest positive serial correlation, and those greater than 2, negative correlation. Finally, although the errors do not need to be normally distributed, severe departures from normalcy can indicate that some statistic inferences of the model are invalid, especially if the sample size is small. Therefore, the Jacque-Bera test for normalcy is also used (Zivot 2006).

Results

None of the cross-correlation analyses found significant correlations between either herring time series and oceanographic features. Table 23 summarizes these results. Diagnostics revealed non-normal distributions for all residuals; however this was of little consequence. According to the Ljung-Box-Pierce statistic, no autocorrelation existed for the Maine herring analyses, but both Canadian analyses revealed some autocorrelated residuals. This was not important because autocorrelation can cause false significant correlation, and no correlation was found in this analysis. All Durbin-Watson statistics were very close to 2, confirming no significant serial correlation based on the residuals.

All script is available in Appendix G.

Model	p-values	Durbin-Watson	Jacque-Bera, p-value	Ljung-Box, p-value
Maine herring x Salinity	0.2559, 0.2552	2.0393	9.2609, 0.0098	15.4690, 0.6924
Maine herring x SST	0.8166, 0.4609, 0.584	2.0197	10.2788, 0.0059	13.5630, 0.8085
Canadian herring x Salinity	0.5217, 0.5224	2.6707	134.1155, 0	32.4103, 0.0281
Canadian herring x SST	0.806, 0,6917, 0.6212	2.5660	164.9544, 0	36.5829, 0.0089

Table 23. Summary of results from cross-correlation analysis.

Discussion

Although numerous previous studies have shown evidence for relationships between environment and herring (Stickney 1967, Graham et al. 1972, Lough and Grosslein 1975, Ridgway 1975, Berenbeim and Sigaev 1977, Lough et al. 1980, Haegele and Schweigert 1985, Graham et al. 1990, Tupper et al. 1998), results here do not confirm these conclusions. Instead, they may suggest that herring landings are not tightly coupled to SST and salinity, and that additional drivers should be investigated. If SST and salinity are not significant in determining catch, these oceanographic factors may not be important for shaping herring management in the GOM. However, although no significant correlations were found, this does not mean they do not exist. It is important to reiterate that cross-correlation approaches are not definitive and further analysis is recommended.

Finally, this work does not address abundance data or draw any inference to the impact environmental factors have on herring populations - only catch statistics. This fact may explain the absence of correlation found. Environmental data, like SST and salinity, would influence population dynamics and abundance, which are rarely directly related to catch. Therefore, although SST and salinity may correlate significantly to herring populations, this effect may not translate into the catch statistics.

Extensive literature exists on the relationships between environmental variables and marine species. This research has investigated the impact of various oceanographic influences on all aspects of life history. However, the data used for these studies are often time series in nature, and may not have been handled appropriately (Hare 1997). Further, inappropriate statistical approaches may have been applied. This can be a critical mistake, particularly for multivariate analysis. Often, the autocorrelation among the time series themselves can cause spurious correlations (Hare 1997, Shumway and Stoffer 2006, Garcia et al. 2007). It is important to review approaches used and assumptions made to determine if they are in fact significant. Addressing time series data properly is not the only concern associated when investigating relationships between time series variables. Significant results indicate correlations only, not causal relationships, and recommend further analysis (Hare 1997, Garcia et. al 2007). Often, strong conclusions have been drawn without proper evidence (Katz 1988, Box et al. 1994, Hare 1997).

This research does strive to address time series of data and results appropriately when running correlation analysis. Because findings are in conflict with previous studies, they do pose some concern about correlations drawn in the past between herring and environmental effects. They also corroborate concerns raised by other authors for correlation studies in general (Katz 1988, Box et al. 1994, Hare 1997, Garcia et al. 2007). Revisiting past studies to ensure proper handling of time series data, application of analysis, and interpretation of results may be necessary to confirm past conclusions. The fundamental message is that correlation analysis regarding the influence of environment on fish species over time is perhaps more complicated than previously thought. Even the methods used to evaluate it may require more in-depth consideration.

CONCLUSION

Many authors have argued that fisheries science is not currently adequate to address marine resource sustainability (Walters and Maguire 1996, Masood 1997, Rose 1997, Batstone and Sharp 2003). Masood (1997) maintained that fishery science has been too narrowly focused and biologically based, and a new perspective is needed. One new approach has been ecosystem based management (EBM). The case for EBM is well documented and support for its implementation is growing. However, science and management must be sure to address not only the biological components of marine ecosystems, but also in context of the physical environment and human community. Extensive research exists concerning relationships between physical oceanography and fish populations, but this is seldom incorporated into management theory. While the environment cannot be controlled, understanding how fish populations react to it can be important for effective management (Hare 1997, Allen et. al 2006, Georgakarakos et. al 2006, Garcia et. al 2007). In addition, just as stock assessment can benefit from incorporating ecosystem processes (Cardinale and Modin 1999, Jonzen et al. 2002), analyses have also shown that socioeconomic and fishing industry data can inform management decisions (Arnason 1990, Arnason 1993, Batstone and Sharp 2003). Although support for EBM is growing, less research has been directed at combining approaches to address both environmental and human drivers. These approaches, currently studied primarily in isolation, need to be synthesized for a broader perspective. Management can then address fisheries using a more comprehensive context that includes many significant impacts.

However, simply addressing the ecosystem as it is today is not sufficient. Prediction is important in fisheries management, yet ecosystem models can be unreliable in predicting future scenarios due to insufficient data (Ellner and Turchin 1995, Park 1998, Garcia et. al 2007). Not only might management miscalculate fishery policies by addressing isolated contemporary drivers, they also miss the very significant historical context of human exploitation. In a more general sense, understanding intrinsic trends and exogenous drivers acting on natural resources over the long term is inherently important for managing those resources (Ryding et al. 2007). Fisheries science needs not only a broader perspective that includes the ecosystem, environment and human community, but also one that encompasses change over time. Historical analyses potentially provide the depth of knowledge necessary to address these concerns.

Despite clear fishery science and management needs, skeptics doubt that they can be met by current quantitative analysis (Batstone and Sharp 2003, Garcia et. al 2007). Analytical methods particularly focused on EBM are being developed. However, many of them are mechanistically driven and mathematically complex. Parameters necessary for computation, including growth rate, mortality, fecundity, etc., often range widely and are often extensively manipulated from raw data (Garcia et. al 2007). They can be difficult to establish and may change drastically over time. Furthermore, population response to varying environmental conditions is difficult to understand and incorporate into these models (Garcia et. al 2007). Nor do they often include anthropogenic influences other than fishing pressure. While these approaches are important and a valid endeavor, it may be difficult to ascertain how well they represent the ecosystem in reality or how that system has changed over time.

Some answers to the complex questions facing fisheries science and management today may not lie in increasingly complex analytical methodologies or traditional methods alone. Simpler but applicable additional tools may be found in other fields, such as business and economics. Time series analysis, in particular, may meet needs not satisfied by current analytical approaches. It is general, flexible and adaptive, and a reliable tool for describing and predicting fishery dynamics (Stergiou 1990, Stergiou 1991, Yoo and Zhang 1993, Freeman and Kirkwood 1995). TSA is not limited to addressing one aspect of exogenous influence; many variables can be tested and outcomes revealed for a variety of stressors and scenarios. Managers and researches can update the analysis through time to achieve refined results (Garcia et. al 2007). In general, time series analysis may prove advantageous in some aspects over traditional fishery science approaches. It does not require extensive information on various impacts or derived variables to deliver robust conclusions (Jensen 1976, Garcia et. al 2007). Straightforward data sets of direct observations are the only requirement (Jensen 1976).

Historical analysis opens the door for TSA by providing the extended time series required. TSA in turn allows historical fishery data to be more extensively analyzed. It does not require additional variables, such as effort, that can be difficult to clarify in historical records (Jensen 1976). Much useful data is available in historical sources but it is currently untapped. For GOM fisheries, information exists for an extensive list of ecologically and commercially important species. Historical data analyzed with TSA may significantly improve understanding of change over time and expectations for fisheries now and in the future.

Answers to current fisheries questions do not lie in a broader perspective or in the application of new methods alone, but in combination. Many new analytical techniques employ very complicated computer models and incorporate many variables, often highly processed, but often little raw data. This may not always reflect the natural ecosystem or change over time. As an alternative, fisheries science can also develop creative hypotheses based on broader contextual knowledge that can be tested with simpler analyses such as TSA. Developing resourceful hypotheses involves perceiving the system not in terms of how a single species reacts to environmental influence or fishery impacts in isolation, but how species interact simultaneously with one another and their environment, including the human community in a more holistic sense. The point, therefore, is to develop innovative hypotheses based on a solid understanding of possible relationships in an ecosystem. Historical resources provide the data and context necessary to do this.

Once information is accessed and inventive hypotheses developed, models can be built that are not necessarily mathematically complex, but instead represent change in the system as a whole. Simple analysis, such as TSA, can be used to test hypotheses based on raw data, as opposed to developing large models that attempt to describe the system in its entirety, regardless of what raw data exists. The simple approaches of TSA can be very effective, although more complex statistics can be incorporated. Additional approaches are already advanced in the fields of business, policy, economics, and the social and behavioral sciences. These approaches lend themselves easily to fisheries data to address underlying patterns, multivariate correlations, and exogenous drivers. Additional analyses particularly pertinent for further development include space-state and dynamic TSA, Bayesian and spatial statistics, hierarchical models, chaos theory, artificial neural networks, and game theory. Finally, statistical software is also increasingly available, so that writing code for specific math function is unnecessary. Instead, understanding the underlying theory to ensure it answers the question addressed is more important than being able to calculate a solution with paper and pen.

Finally, understanding the marine ecosystem and all its possible influences in full is not necessary in order to address it through TSA. Fisheries science needs only to advance a resourceful hypothesis. The hypothesis does not need to be correct to be tested, only informed enough to be realistic and asking a applicable question. Historical information already provides knowledge useful in constructing hypotheses, and approaches such as TSA already exist to test them. The fundamental point is to think broadly about data and to employ it creatively for holistic hypotheses about how marine systems function, as opposed to developing complex models that attempt to include as many variables as possible.

However, time series analysis addresses patterns over time and, as it is applied here, catch statistics only. Although it may offer benefits for fishery science, this analysis does not address population dynamics. It is important to be clear on this point – TSA is a different approach from stock assessment, which does attempt to understand population dynamics and abundance. This does not negate the importance or usefulness of TSA for fisheries science, only that it should be used in concert with traditional methods to provide further context and knowledge. However, additional information on effort or life history parameters can be incorporated into TSA. Such an approach would allow TSA to begin addressing population dynamics, yet this may be constrained by the necessary number of consistent observations back in time. As mentioned above, this information may be much less readily available in historical texts.

This master's thesis was an initial foray into developing new approaches through a case study of Atlantic herring. Herring are pertinent because of their ecological and commercial importance, possible management questions, and long history of anthropogenic influence and exploitation. The primary goal was to begin broadening the temporal and ecosystem context in which fishery statistics are examined, as well as to explore the use of additional appropriate and flexible methods. Methodologies can be refined and expanded in the future.

For herring specifically, understanding long-term cycles in the GOM ecosystem, including its human components, is valuable for addressing questions in current management, acquiring additional information, and providing tools for more accurate predictions. The investigation of additional exogenous drivers in the analysis is a clear opportunity for refining techniques given here. There are numerous possibilities, including biological indicators (from fish behavior to further environmental influences), fishery aspects (such as fleet dynamics and gear changes), anthropogenic variables (such as land use and pollution), and more in-depth socioeconomic pressures. Finally, the movement of effort and herring populations from inshore waters to offshore grounds is also significant for further examination. This spatial aspect is not limited to this movement, as recent work has stressed the importance of spatial scale to identifying underlying patterns and processes (Rouyer et al. 2008). A more complete understanding of the drivers affecting the herring fishery can help assess the current GOM herring fishery, as well. If the analysis can identify significant modern drivers acting on the fishery, researchers can also determine if contemporary science and management adequately address their importance.

This work encourages additional and more rigorous analysis not only for herring, but the greater GOM system, in addition to an expansion of approaches. Future work, including planned doctoral research, will expand the methods here from a single species (herring) to a broad ecosystem context by including other fisheries. This approach will explore how changes in the abundance and distribution of forage species due to fishing and other factors are expressed in catch across the system. As for the methods themselves, there are considerable possibilities for more involved models and analysis. Here, only basic and preliminary TSA approaches were applied, and all models were linear. However, factors affecting fisheries are not always additive, but may be multiplicative or interact as well (Rouyer et al. 2008). Methods for multiplicative or dynamic TSA are available (Box and Jenkins 1976, Priestly 1978, Ellner and Turchin 1995, Fachinni et al. 2007). Additional quantitative methods that can be informative include spatial and Bayesian statistics, hierarchical models, and scenario-driven game theory. TSA is flexible to accommodate additional model-building and sophisticated techniques have been developed in other fields that may be easily applicable to fisheries science.

Finally, results can be applied to predict potential future scenarios and provide information for modeling potential management schemes. The predictive application of long-term time series analysis can forecast from current trends. Collective use of many of the approaches described here can refine the model's predictive ability and develop hierarchies of alternative scenarios suggesting future benefits, problems, and trade-offs in implementing various management efforts.

In conclusion, this thesis was meant to be an initial look into new methods and expanding the context and temporal view of fisheries science. The main objectives of this work were to explore the application of time series analysis and the importance of historical data, including the qualitative literature. The approaches used and further possible methods described have real potential for new insight into understanding marine systems in their entirety, as opposed to in isolation today or in the past. Time series analysis and historical data should be considered a useful avenue for providing addition context, both in time and across the system, in order to more effectively evaluate the current fishery situation and plan for the future.

LIST OF REFERENCES

Akaike, H. 1974. A new look at statistical model identification. IEEE Transactions on Automatic Control. 19(6): 716-723.

Allen, M., R. Rosell, and D. Evans. 2006. Predicting catches for the Lough Neagh (Northern Ireland) eel fishery based on stock inputs, effort and environmental variables. Fisheries Management & Ecology. 13: 251-260.

Anthony, V. C., and G. T. Waring. 1980. Estimates of herring spawning stock biomass and egg production for the Georges Bank-Gulf of Maine region. NAFO SCR Doc. 80//IX/135: 38p.

Arnason, R. 1990. Minimum Information Management in Fisheries. The Canadian Journal of Economics. 23(3): 630-653.

Arnason, R. 1993. The Icelandic Individual Transferable Quota System: A Descriptive Account. Marine Resource Economics. 8(3): 201-18.

Atlantic herring. 2008. National Marine Fisheries Service, National Oceanic and Atmospheric Administration. http://www.nmfs.noaa.gov/fishwatch/species/atl_herring.htm (30 April 2008).

Austin, H. M., D. A. Evans and B. L. Norcross. 1986. Time series analyses as a means of examining long term biological data sets. Proceedings of IEEE OCEANS; 1986 September 23-25; Washington DC: p946-952.

Batstone, C. J., and B. M. H. Sharp. 2003. Minimum information management systems and ITQ fisheries management. Journal of Environmental Economics and Management. 45:492-504.

Baumgartner, T. R., A. Soutar, and V. Ferreira-Bartrina. 1992. Reconstruction of the history of pacific sardine and northern anchovy populations over the past two millennia from sediments of the Santa Barbara Basin, California. CalCOFI Rep. 33: 24-40.

Beamish R. J. and C. Mahnken. 1999. Taking the next step in fisheries management. In: Ecosystem approaches for fisheries management. Fairbanks, AK: University of Alaska Sea Grant. p1-22.

Berenbeim, D. Ya., and I. K. Sigaev. 1977. On correlation between the water temperature and the spawning times for Georges Bank herring. ICNAF Res. Doc. 77/VI/42: 6p.

Bianco, A., M. Garcia Ben, E. Martinez, and V. Yohai. 2001. Outlier detection in regression models with ARIMA errors using robust estimates. Journal of Forecasting. 20: 565-579.

Bigelow, H. B., and W. C. Schroeder. 1953. Fishes of the Gulf of Maine. U.S. Fish. Wildl. Serv. Fish. Bull. (special issue). 53: 577p.

Bigney, T. B. 2000. The Maine Sardine Industry: The first one-hundred years 1876-1976. In: Warren, J. L., ed. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. p157-158.

Box, G. E. P. and G. C. Tiao 1975. Intervention analysis with application to economic and environmental problems. Journal of the American Statistical Association. 70(349): 70-79.

Box, G. E. P. and G. M. Jenkins. 1976. Time series analysis: forecasting and control. Holden-Day, San Francisco, CA: 573p.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 1994. Time series analysis: forecasting & control. Revised ed. Holden-Day, San Francisco, CA: 575p.

Burnham, K. P. and D. R. Anderson. 2004. Multimodel Inference Understanding AIC and BIC in Model Selection. Sociological Methods & Research. 33(2): 261-304.

Caddy, J.F. and T. D. Iles. 1973. Underwater observations on herring spawning grounds on Georges Bank. Int. Comm. Northwest Atl. Fish Res. Bull. 10: 131-139.

Cardinale, M. and J. Modin 1999. Change in size at maturity of Baltic cod (*Gadus morhua*) during a period of large variations in stock size and environmental conditions. Fish. Res. 41: 285–295.

Chan, W. S., S. H. Cheung, and K. H. Wu. 2004. Multiple forecast with autoregressive time series models: case studies. Mathematics and Computers in Simulation. 64: 421-430.

Chang, I., G. C. Tiao, and C. Chen. 1988. Estimation of time series parameters in the presence of outliers. Technometrics. 30(2): 193-204.

Chatfield, C. 1977. Some recent developments in time-series analysis. Journal of the Royal Statistical Society, Series A (General). 140(4): 492-510.

Chen, C. and G. C. Tiao. 1990. Random level-shift time series models, ARIMA approximations, and level-shift detection. Journal of Business & Economic Statistics. 8(1): 83-97.

Cryer, J. D. and K.-S. Chan. 2008. Time Series Analysis: with Applications in R. Springer, New York, NY: 491p.

Day, L. R. 1951. The herring, alewife, and shad. Fish. Res. Bd. Canada. General Series 19. Circular.

Davis, H. A. 1950. Eastport and the Sardine Industry. In: Maine Sardine Council eds. 1986. Maine Sardine Industry History: An Anthology.

Dow, R. L. 1951. The Sea Herring (*Clupea harengus*). Department of Sea & Shore Fisheries. Washington D.C.: Govt. Printing Office.

Earll, R. E. and H. M. Smith. 1887. The American Sardine Industry in 1886. U.S. Fish Commission. Washington D.C.: Govt. Printing Office.

Ellner, S. and P. Turchin. 1995. Chaos in a noisy world: New methods and evidence from time-series analysis. The American Naturalist, 145(3): 343-375.

Facchini, A., C. Mocenni, N. Marwan, A. Vincino, and E. Tiezzi. 2007. Nonlinear time series analysis of dissolved oxygen in the *Orbetello Lagoon* (Italy). Ecological Modeling. 203: 339-348.

FAO. 2005. Review of the state of world marine fishery resource. FAO Technical Paper 457. Food and Agriculture Organization of the United Nations: Rome, Italy.

Freeman, S. N. and G. P. Kirkwood. 1995. On a structural time series method for estimating stock biomass and recruitment from catch and effort data. Fish. Res. 22: 77-98.

French, H. 2000. The Eastport Waterfront: 1898-1913. In: Warren, J. L., ed. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. p151-155.

Garcia, S. P., L. B. DeLancey, J. S. Almeida, and R. W. Chapman. 2007. Ecoforecasting in real time for commercial fisheries: the Atlantic white shrimp as a case study. Marine Biology. 152: 155-24.

Georgakarakos, S., S. Koustsoubas, and V. Valavanis. 2006. Time series analysis and forecasting techniques applied on loliginid and ommastrephid landings in Greek waters. Fish. Res. 78: 55-71.

Gilman, J. 2001. Canned: A history of the Sardine Industry (Part One & Two). Published by author, Parsons Printing, St. Stephen, N. B. 304p.

Glass, G. V. 1972. Estimating the effects of intervention into a non-stationary time-series. American Education Research Journal. (93): 463-477.

Graham, J. J., S. B.Chenoweth, and C. W. Davis. 1972. Abundance, distribution, movements, and lengths of larval herring along the western coast of the Gulf of Maine. Fishery Bulletin. 70(2): 307-313.

Graham, J. J., D. K. Stevenson, and K. M. Sherman. 1990. Relation between winter temperature and survival of larval Atlantic herring along the Maine coast. Trans. Am. Fish. Soc. 119(4): 730 – 740.

Haegele, C. W. and J. F. Schweigert 1985. Distribution and characteristics of herring spawning grounds and description of spawning behavior. Can. J. Fish. Aquat. Sci. 42: 39 – 55.

Hall, A. 1898. The herring industry of the Passamaquoddy region, Maine. In: U.S. Fish Commission, Report of the Commissioner for the year ending June 30, 1896, Part XXII. Washington, D. C.:Govt. Printing Office: p443 – 487.

Hare, S. R. 1997. Low frequency climate variability and salmon production. [Ph.D. Dissertation], University of Washington, Seattle, WA. Available from: http://www.iphc.washington.edu/Staff/hare/html/diss/phd.html.

Hartmann, D. P., J. M. Gottman, R. R. Jones, W. Gardner, A. E. Kazdin, and R. S. Vaught. 1980. Interrupted time-series analysis and its application to behavioral data. Journal of Applied Behavioral Analysis. 13: 543-559.

Hurvich, C. M. and C.-L. Tsai. 1989. Regression and time-series model selection in small samples. Biometrika. 76(2): 297-307.

Insightful Corporation. 2007. S-PLUS 8 (Version 8.0.4) [Computer Software]. Seattle WA: Insightful Corporation. Available from: http://www.insightful.com.

Jackson, J. B. C. 1997. Reefs since Columbus. Coral Reefs. 16(Suppl): S23-S32.

Jackson, J. B. C. 2001. What was natural in the coastal oceans? PNAS. 98: 5411-18

Jackson, J. B. C, M. X. Kirby, W.H. Berger, K.A. Bjorndal, L.W. Botsford, B.J. Bourque, R.H. Bradbury, R. Cooke, J. Erlandson, J.A. Estes, T.P. Hughes, S. Kidwell, C.B. Lange, H.S. Lenihan, J.M. Pandolfi, C.H. Peterson, R.S. Steneck, M.J. Tegner, and R.R. Warner. 2001. Historical overfishing and the recent collapse of coastal ecosystems. Science. 293: 629-38.

Jensen, A. L. 1976. Time series analysis and forecasting of Atlantic menhaden catch. Chesapeake Science. 17(4): 305-307.

Johnson, W. H. 1939. Feeding of the herring. J. Fish. Res. Bd Canada. 4(5): 392 – 395.

Jonzen, N., M. Cardinale, A. Gardmark, F. Arrhenius, and P. Lundberg. 2002. Risk of collapse in the eastern Baltic cod fishery. Marine Ecology Progress Series. 240: 225-233.

Katz, R. W. 1988. Use of cross correlations in the search for teleconnections. Journal of Climatology. 8: 241-253.

Kim, S., S. Jung, and C. I. Zhang. 1997. The effect of seasonal anomalies of seawater temperature and salinity on the fluctuation in yield of small yellow croaker, *Pseudosciaena polyactis*, in the Yellow Sea. Fisheries Oceanography 6(1): 1-9.

Klyashtorin, L. B. 1998. Long-term climate change and main commercial fish production in the Atlantic and Pacific. Fish. Res. 37: 115-125.

Lawrence, E.M. (2000) A History of Sardine Canning. In: Warren, J. L., ed. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. p131-134.

Libby, G. 2007. Herring fishery needs attention now. Portland Press Herald, November 6, 2007. http://pressherald.mainetoday.com/story.php?id= 145667&ac=PHedi (November 6, 2007).

Link, J. S., and F. P. Almeida. 2000. An overview and history of the food web dynamics program of the Northeast Fisheries Science Center, Woods Hole, Massachusetts. Woods Hole, MA: NOAA Tech. Mem. NMFS-NE-159: 60pp.

Lotze, H. K., H. S. Lenihan, B. J. Bourque, R. H. Bradbury, R. G. Cooke, M. C. Kay, S. M. Kidwell, M. X. Kirby, C. H. Peterson, J. B. C. Jackson. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science. 312: 1806-1809.

Lough, R. G. and M. D. Grosslein. 1975. Winter mortality of Georges Bank herring larvae. ICNAF Res. Doc. 75/113: 39p.

Lough, R. G., G. R. Bolz, M. R. Pennington, and M. D. Grosslein. 1980. Abundance and mortality estimates for sea herring (*Clupea harengus L.*) larvae spawned in the Georges Bank-Nantucket Shoals area, 1971-1978 seasons, in relation to spawning stock and recruitment. NAFO SCR Doc. 80/IX/129: 59p.

Maine Sardine Industry History. 2000. Compiled by J. L. Warren, Brewer Maine: 160p.

Masood, E. 1997. Fisheries science: all at sea when it comes to politics? Nature. 386: 105-106.

McDowall, D., R. McCleary, E. E. Meidinger, and R. A. Hay, Jr. 1980. Interrupted time series analysis. In: J. L. Sullivan, ed. Series: Quantitative Applications in the Social Sciences. Beverly Hills: Sage Publications, Inc. 96p.

McKenzie, R. A. 1964. Observations on herring spawning off southwest Nova Scotia. J. Fish. Res. Bd. Canada. 21(1): 203 – 205.

Messieh, S., R. Pottle, P. MacPherson, and T. Hurlbut. 1985. Spawning and exploitation of Atlantic herring (*Clupea harengus*) at Escuminac in the southwestern Gulf of St. Lawrence, spring 1983. J. Northw. Atl. Fish. Sci. 6(2): 125 – 133.

Moore, H. F. 1898. Observations on the herring and herring fisheries of the northeast coast, with special reference to the vicinity of Passamaquoddy Bay. In: U.S. Fish Commission, Report of the Commissioner for the year ending June 30, 1896, Part XXII. Washington, D. C.: Govt. Printing Office. p387 – 442.

Munroe, T. A. 2002. Atlantic herring. In: Collette, B. B., and G. Klein-MacPhee, eds. Bigelow and Schroeder's Fishes of the Gulf of Maine. Smithsonian Institution Press, Washington and London: 141 – 156.

Murtaugh, P. A. 2000. Paired intervention analysis in ecology. Journal of Agriculture, Biological, and Environmental Statistics. 5(3): 280-292.

NEFMC. 1999. Final Atlantic herring fishery management plan. New England Fishery Management Council, Saugus, MA. Available from: http://www.nefmc.org/herring/index.html

Overholtz, W. J., and K. D. Friedland. 2002. Recovery of the Gulf of Maine-Georges Bank Atlantic herring (*Clupea harengus*) complex: perspectives based on bottom trawl survey data. Fishery Bulletin. 100(3): 593 – 608.

Park, H.-H. 1998. Analysis and prediction of walleye pollock (*Theragra chalcogramma*) landings in Korea by time series analysis. Fish. Res. 38: 1-7.

Parzen, E. 1961. An approach to time series analysis. The Annals of Mathematical Statistics.32 (4): 951-989.

Pauly, D. 1995. Anecdotes and the shifting baselines syndrome of fisheries. TREE. 10 (10), 430.

Pauly, D. 2000. Global change, fisheries and the integrity of marine ecosystems: the future has already begun. In: Pimentel D., L. Westra and R. R. Ross, eds. Ecological integrity: integrating environment, conservation and health. Washington D.C.: Island Press. p227 – 239.

Pauly, D., V. Christensen, R. Froese, and M.L. Palomares. 2000. Fishing down aquatic food webs. American Scientist. 88(1): 46-51.

Pike, M. 2000. A short history of the Sardine Business in Lubec. In: Warren, J. L., ed. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. p143-146.

Plante, J. M. 2006. TRAC reports herring complex in good shape. Commercial Fisheries News. 33(11).

Priestly, M. B. 1978. Non-linear models in time series analysis. The Statistician, 27: 3/4. Partial Proceedings of the 1978 I.O.S. Annual Conference on Time Series Analysis (and Forecasting). Sept – Dec. 1978: 159 – 176.

R Development Core Team. 2008. R: A Language and Environment for Statistical Computing (Version 2.7.0). [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Available from: http://www.R-project.org.

Reid R. N., L. M. Cargnelli, S. J. Griesbach, D. B. Packer, D. L. Johnson, C. A. Zetlin, W. W. Morse, and P. L. Berrien. 1999. Atlantic herring, *Clupea harengus*, life history and habitat characteristics. NOAA Tech. Mem. NMFS-NE-126, Washington D.C.: 48p.

Rich, W. H. 1929. Fishing grounds of the Gulf of Maine. In: Report of the U.S. Commissioner of Fisheries. Washington, D.C.: Govt. Printing Office.

Ridgway, G. J. 1975. A conceptual model of stocks of herring (*Clupea harengus*) in the Gulf of Maine. ICNAGF Res. Doc. 75/100: 17p.

Rogers-Bennett, L., P. L. Haaker, T. O. Huff, and P. K. Dayton. 2002. Estimating baseline abundances of abalone in California for restoration. CalCOFI Reports. 43, 97-111.

Rose, G. A. 1997. The trouble with fisheries science! Reviews in Fish Biology and Fisheries. 7:365-370.

Rosenberg, A. A., W.J. Bolster, K.E. Alexander, W.B. Leavenworth, A.B. Cooper, and M.G. McKenzie. 2005. The history of ocean resources: modeling cod biomass using historical records. Frontiers in Ecology and the Environment. 3(2): 84-90.

Rouyer, T. J.-M. Fromentin, F. Ménard, B. Cazelles, K. Briand, R. Pianet, B. Planque, and N. C. Stenseth. 2008. Complex interplays among population dynamics, environmental forcing, and exploitation in fisheries. PNAS. 105:14, 5420-5425.

Ryding, K. E., J. J. Mullspaugh, and J. R. Skalski. 2007. Using time series to estimate rates of population change from abundance data. Journal of Wildlife Management. 71:1, 202-207.

Sáenz-Arroyo, A., C. M. Roberts, J. Torre, M. Carino-Olevera, and R. R. Enriquez-Andrade. 2005. Rapidly shifting environmental baselines among fishers of the Gulf of California. PNAS – Biology. 272: 1957-1962.

Sala, E., and O. Aburto-Oropeza, G. Paredes, I. Parra, J. C. Barrera, and P. K. Dayton. 2002. A general model for designing networks of marine reserves. Science. 298, 1991-3.

"Sardines." 1945. US Dept. of the Interior: Fish and Wildlife Service. Fishery Leaflet 5. Chicago, Ill: Govt. Printing Office.

Scattergood, L. W. and S. N. Tibbo. 1959. The herring fishery of the Northwest Atlantic. Fish. Res. Bd. Canada Bull. No. 121. 42p.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature. 413: 591-596.

Shumway, R. H. and D. S. Stoffer. 2006. Time Series Analysis and Its Applications: with R examples. New York, NY: Springer. 575p.

Smith, T.D., and J. S. Link. 2005. Autopsy your dead... and living: a proposal for fisheries science, fisheries management and fisheries. Fish and Fisheries. 6:73-87.

StatSoft, Inc. 2006. Time series analysis In: Electronic StatSoft Textbook, Tulsa, OK. http://www.statsoft.com/textbook/sttimser.html (18 July 2007).

Stephenson, R. L., and I. Kornfield. 1990. Reappearance of spawning Atlantic herring (*Clupea harengus harengus*) on Georges Bank: Population resurgence not recolonization. Can. J. Fish. Aq. Sci. 47: 1060-1064.

Stergiou, K. I. 1990. Prediction of the Mullidae fishery in the Eastern Mediterranean 24 months in advance. Fish. Res. 9:1, 67-74.

Stergiou, K. I. 1991. Short-term fisheries forecasting – comparison of smoothing, ARIMA and regression techniques. Journal of Applied Ichthyology. 7:4, 193-204.

Stevenson, D. K., K. Friedland, and M. Armstrong. 1997. Managing the U.S. Atlantic herring coastal stock complex for long-term sustainable yield. ICES C.M. 1997/DD:02, 14p.

Stevenson, D., and M. L. Scott. 2005. Atlantic herring, *Clupea harengus*, life history and habitat characteristics. NOAA Tech. Mem. NMFS-NE-192: 84p.

Stickney, A. P. 1967. Orientation of juvenile Atlantic herring (*Clupea harengus* L.) to temperature and salinity. FAO Internatl. Conf. on Fish Behavior in Relation to Fishing Techniques and Tactics: Bergen, Norway. FAO Fisheries Reports. 2(3): 323-342.

Sugiura, N. 1978. Further analysis of the data by Akaike's information criterion and finite corrections. Communications in Statistics: Theory and Methods. 7(1):13-26.

Tiao, G.C. 1983. Use of statistical methods in the analysis of environmental data. The American Statistician. 37(4): 459-470.

Tiao, G. C. 1993. An introduction to multiple time series analysis. Medical Care. 31(5 Suppl.): YS71-YS74.

Tiao, G. C., and R. S. Tsay. 1983. Multiple time series modeling and extended sample cross-correlations. Journal of Business & Economic Statistics. 1(1): 43-56.

TRAC. 2006. Gulf of Maine-Georges Bank Herring Stock Complex. TRAC Status Report 2006/01.

Tsay, R. S. 1988. Outliers, level shifts, and variance changes in time series. Journal of Forecasting. 7: 1-20.

Toft, J. D. 2000. Sardines: Some historical data on Maine Sardine Industry. In: Warren, J. L., ed. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. p135-141.

Townsend, D. W. 1992. Ecology of larval herring in relation to the oceanography of the Gulf of Maine. Journal of Plankton Research. 14(4): 467-493.

Tupper, M. H., V. C. Anthony, S. B. Chenoweth, and H. A. MacCluen. 1998. Biology and assessment of Gulf of Maine herring stocks. MEDMR: 104 p.

Turchin, P., and A. D. Taylor. 1992. Complex dynamics in ecological time series. Ecology. 73(1): 289-305.

Velicer, W. F. and S. M. Colby. 2005. A comparison of missing-data procedures for ARIMA time-series analysis. Educational and Psychological Measurement. 65(4): 596-615.

Warren, J. L. 2000. Maine Sardine Industry History 1875 – 2000. Brewer, Maine: Self-published. 160 p.

Walters, C. and J. Maguire. 1996. Lessons for stock assessment from the Northern Cod Collapse. Rev. Fish Biol. Fisheries. 6:125–138.

Webber, F. C. 1921. The Maine Sardine Industry. In: Maine Sardine Council eds. 1986. Maine Sardine Industry History: An Anthology. Maine Sardine Council.

Worm, B., E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, E. Sala, K. A. Selkoe, J. J. Stachowicz, and R.Watson. 2006. Impact of biodiversity loss on ocean ecosystem services. Science. 314(5800):787-790.

Yoo, S., and C. I. Zhang. 1993. Forecasting of hairtail (*Trichiurus leplurus*) landings in Korean waters by time series analysis. Bull. Korean Fish. Soc. 26(4): 363-368.

Zivot, E. 2006. Modeling financial time series with S-Plus. Springer, New York, NY: 998p

APPENDICES

APPENDIX A:

LIST OF SOURCES BY LOCATION AND YEARS COVERED, INCLUDING A SUMMARY OF CONTENT

Source	Location	Years	Contents
Annual Reports of Fisheries - Canada	Canada	1871-75	**1871-75:** Qualitative report of fisheries inspectors on all fisheries by county Herring caught by county in barrels, in boxes (smoked) starts in 1872 **1872-75:** Total men, weirs, nets, boats, "fishing material" by county Recapitulation of yield and value for entire province by species Yield by district in each province begins in 1873 Tables at end of inspector's reports begin in 1873 (instead of at end of report)
		1877 – 1916 (some years missing)	Detailed reports by counties (at fine scale) begin in 1877 Total recapitulation/comparative statement of total fish & value by province and fishery General qualitative report by inspectors by county for each province Total fish (by smoked in boxes or in barrels) and gear by fishery for each province - by county, several locations within each Gear = total boats, vessels, fishing materials by area (not spec by fishery) Products, value; Prices, totals **1875 onward:** Imports/exports, goods by fishery, province
		1916 - 1920 1922 - 1930	**1917-19:** Qualitative report for all of Canada, cannery inspection, discussion of herring, landings by province, quantity, and value Number of fishermen, boats, gear for all of Canada **1920:** Qualitative report by county and province for all fisheries Summaries: totals (#, value) of each fishery by province, for entire country Qualitative discussion of fish canning and curing industries Totals for gear, equipment by province and by county Yield and value of fish by species, county (some detail) **1922-30**: Qualitative report only - no tables. Landings by county in text. Licenses by fishery and province **By 1940**: No longer reported by province - for entire fishery Much less info by 1940s, and even less by 1950s.
Special Appended Reports	Canada	1899	Qualitative discussion of pollution and it's effects on fish

Meteorological Reports	Canada	1871-73, 1875	General discussion of meteorological stations & methods Mean temps by month, season, & station (many per county) - for 1871-72 High/low temps for the year and station -1871-72 Mean daily temp by station (corrected for diurnal variation) Rainfall & snowfall by month/station, # days of rain, ave amount of rain
Canadian Fisheries Statistics	Canada	1917-1962	Qualitative overall discussion of fisheries and fishing season Discussion of canning and curing industry Fish caught & marketed: summaries by province, also by district (summaries only by province for 1947-1950 - back to districts in 1951) Gear, vessels, boats, employees by fishing district (cross fisheries) Employees in canning industry by month and county (for all fish) Summary stats of fish processing and canning **1920s:** Percentage of fish taken offshore **1920s:** Characterization of vessels by fishery and district **1940:** No qualitative reports **1950s:** Much less industry information, still quantity and value
Monthly Review of Canadian Fisheries Stats	Canada	1947 - 1960	Monthly catch stats by species and province Qualitative overall text for all fisheries by month
Dept of Marine & Fisheries - Monthly Statistic Returns	Canada	Late 1920s-50s?	Landings of all fish by county (with more detail: several locations per county) and month Includes info on loss of property/life, weather, general fishery comments
Report of the Atl. Biol. Station, Biological Board of Canada	SABS	1931-45	Qualitative report on work done at SABS - oceanographic, fishery-related, etc.
Bulletin of the United States Fish Commission	US	1882, 1884, 1885	Qualitative articles on the fisheries of the US
US Fishery Statistics	US	1931 - 1947	

Statistics of the Fisheries of the New England States	New England	1905	Catch by state and species in pounds and value Qualitative text on total fisheries, specific for canning and smoking by state Breakdown of catch by product, species, and county for each state
Sea & Shore Fisheries	Maine	1895 - 1919 (some years missing)	Qualitative information on herring industry for the state **1897-1902, 1905-06:** Herring fishery stats by county, product, and value; also packing stats by county
Maine Landings	Maine	1939-1960	Summary total by gear and county at end of year (starting in 1942) Landings by month, species, value, pounds
Report of the Commissioners on Fisheries & Game	Mass	1878 – 1932 (some years missing)	**1878-1900** : Catch by species, gear, town/place and proprietor **1905:** Qualitative text, catch in barrels/lbs by gear for Gloucester landings **1909, 1919:** Catch by town **1920:** Text and catch total for state

APPENDIX B:

LIST OF INDIVIDUALS INTERVIEWED

Contact	Location	Additional Information
David Libby – *Maine DMR*	Boothbay, ME	Will plan trip to visit ME DMR, invitation to give presentation in the future. D. Libby has sent on some very useful information.
John Annala – *GMRI*	Portland, ME	Gulf of Maine Research Institute (GMRI)
Jeff Kaelin – *Research Associates*	Portland, ME	Final director of Marine Sardine Council, archived Sardine Council documents, has been very helpful in attempting to relocated information
Bill Overholtz – *NMFS*	Woods Hole, MA	Was NMFS assessment biologist – involved in herring assessments and NMFS acoustic surveys.
Matthew Cieri – *ME DMR*	Boothbay, ME	M. Cieri is a ME state herring biologist.
Kohl Kanwit – *ME DMR*	Boothbay, ME	K. Kanwit now works on groundfish but has previously done extensive work on herring.
Mike Fogarty – *NMFS*	Woods Hole, MA	
Robert Stephenson – *DFO*	St. Andrew's Bay, NB (Director)	Dr. Stephenson is very interested in current work.
V. Anthony	Boothbay, ME	Met at Fishermen's Forum in March, visited in August
A. West –*industry*	Prospect Harbor, ME	Met at Cannery in Prospect Harbor – gave additional fisheries contacts
Tony Hooper – *industry*	NB Canada	Sent on add'l current cannery info
Mike Power – *DFO*	St. Andrew's, NB	
Peter Baker – *CHOIR*		Met at GMRI herring meeting in May
Ted Ames – *lobster*	Mt. Desert, ME	Met at GMRI herring meeting in May, T. Ames has done a lot of historical cod research.
James Warren – *Author*	Brewer, ME	Still need to track down contact information – would like to discuss where he found his information.
John Gilman - *"Canned"*	Deer Island, NB	

APPENDIX C

SUMMARY OF CONVERSION FACTORS

Unit	Pound	Canned Fish	Fresh Fish	Salted/ Pickled Fish	Smoked fish	Alternative Unit	Notes	Year	Source
Barrel	200lbs herring/barrel		300 fresh fish	330 round herring	5 boxes of 100 bloaters each		In general, 5 pecks salt to cure a barrel of herring Law: 200 lbs fish/ barrel Little wt lost in pickling	1898	Bulletin of the US Fish Commission 1898 p438, 486
Barrel						40gal barrel = 3,000-8,000 pressed sardines		1898	Bulletin of the US Fish Commission 1898 p465
Barrel - Fresh herring			one barrel			yeilds ~5 boxes 100 bloaters		1898	Bulletin of the US Fish Commission 1898
Barrel - herring oil			5hh in spring, 2-2.5hh in fall (fish are fatter)				$9/barrel but $3 per hh of fish due to cost of labor in pressing	1880	The Sardine Industry - 1880 (Earll).
Barrel - Pickled herring	200lbs herring/barrel		250lbs fresh - for pickled herring	200lbs			200lbs for both pickled and Russian sardines	1880	The Fisheries & Fishery Industries of the U.S. SXn II (Goode) p13-4.
Barrel - Russian sardines	200lbs herring/barrel		~323 lbs fresh	200lbs			200lbs for both pickled and Russian sardines	1880	The Fisheries & Fishery Industries of the U.S. SXn II (Goode) p13-4.

Basket			1.5 bushels in Eastport, further W 1.75 bushels					1886	The Sardine Industry – 1880 (Earll)/ The American Sardine Industry (Earll & Smith)
Box							Box: 15.5" x 11.5" x 7.5"	1898	Bulletin of the US Fish Commission 1898 pg527
Box - "Length-wise"	~6				15-20 fish		12"x6.5"x 2.75"	1898	Bulletin of the US Fish Commission 1898 pg482
Box - "Med-scaled"					30-40 lg fish, 40-50 sm		12"x6.5"x 2.75"	1898	Bulletin of the US Fish Commission 1898 pg482
Box - "No. 1"					55-75 fish		12"x6.5"x 2.75"	1898	Bulletin of the US Fish Commission 1898 pg482
Box - pickled herring				100 fish				1886	The Sardine Industry – 1880 (Earll)/ The American Sardine Industry (Earll & Smith)
Box - Smoked herring			~11.57 fresh		8.5lbs prepared			1880	The Fisheries & Fishery Industries of the U.S. SXn II (Goode) p13-4.
Box - Bloaters			~44.5 lbs fresh		33 1/3 prepared			1880	The Fisheries & Fishery Industries of the U.S. SXn II (Goode) p13-4.

Bushel (US standard)	70							1957	Maine Landings - Dept of Sea & Shore Fisheries
Can			~12 fish, never less than 7-8					1898	Bulletin of the US Fish Commission 1898
Can - Quarter			9-12 fish dressed (3.5 - 4" cut from 6"long)				packed in oil 4.5" x 3" x 1"	1898	Bulletin of the US Fish Commission 1898 pg527
Can - Half			10-16 fish dressed (4-4.5" cut from 8" long)				4.5"x3.5"x2"	1898	Bulletin of the US Fish Commission 1898 pg527
Can - Sardines			~0.866lbs fresh/can					1880	The Fisheries & Fishery Industries of the U.S. SXn II (Goode) p13-4.
Case						100 qrtr & half cans, 50 3/4 cans		1895	Sardine Business in Maine (Sardine Council Anthology
CWT	100							1910s	Canadian Annual Reports
Herring - Bloated (100)	25-35lbs				100 bloaters		Wt depends on size and extent of smoking	1898	Bulletin of the US Fish Commission 1898 pg486
Herring - Round	200		211lbs fresh fish make 200 lbs round herring					1898	Bulletin of the US Fish Commission 1898 pg439

Herring - Gibbed	200		228 lbs round herring make 200 lbs gibbed					1898	Bulletin of the US Fish Commission 1898 pg439
Herring - Split	144		200 lbs fresh herring make 144 split herring					1898	Bulletin of the US Fish Commission 1898 pg440
Hogshead (hh)	~1225		~51,000 4" fish, ~33,500 5" fish, 22,300 6" fish, 9,900 7" fish, 7,200 8" fish			171/2 bushel		1921/ 1998	Fishermen's Voice online
Hogshead (hh)						52.5 gal/~238.5 liters	British system - but same as US	2000	Dictionary of Units of Measure: http://www.unc.edu/ ~rowlett/units/ dictH.html
Hogshead (hh)			10 baskets (Eastport and other areas)			5 flour barrels = 15 bushels in Eastport	Some places increased to 6 barrels or 15 baskets	1880/ 1886	The Sardine Industry - 1880 Earll/The American Sardine Industry (Earll & Smith)
Hogshead (hh)		~ 16 cases of qrtr cans, 1-3 Cases of 1/2 cans (extremes: 12-22 cases dep. on cond of fish)						1886	The Sardine Industry - 1880 Earll/The American Sardine Industry (Earll & Smith)

Hogshead (hh)	1,000	20 cases of canned fish						1921	The Maine Sardine Industry (Sardine Council Anthology)
Keg - Russian Sardines	4, 7, 11 lbs each							1895	Sardine Business in Maine (Sardine Council Anthology
Meal			15 hh fish, 1.5 tons waste					1921	The Maine Sardine Industry (Sardine Council Anthology)
Pail - Russian sardines						4 quarts	Fish packed in spices and vinegar - barrel "worthless" in NYC where marketed	1886	The Sardine Industry - 1880 Earll/The American Sardine Industry (Earll & Smith)
Sardine - Russian			7-12" alive, 5-9" dressed (similar in size to those in oil					1898	Bulletin of the US Fish Commission 1898 pg438, 486
Ton - metric	2204.62 262								

APPENDIX D:

QUALITATIVE TIMELINE

YEAR	EVENT
1797	Smoked herring industry begins in Lubec
1854-55	Frozen herring trade from Newfoundland begins as bait for cod
1861-65	Civil War: destroyed herring trade with the south (smoked herring) – reduced demand
1864	Herring scrap (byprod of oil) for poultry, cattle, sheep, hogs feed was an est practice – short period of time BEFORE sardine industry when market favored scrap over smoked herring or oil
1865	Beginning of building deep-water weirs (imp for Cobscook Bay) Smoked herring industry in Lubec begins to fall off
1866-67	Frozen herring industry begins in Eastport - majority of fish caught in NB, some in US, Passamaquoddy only area in US where herring extensively frozen
1873	Washington treaty (July 1) disastrous for many herring industries in US = no duty on Canadian prods (bloaters, smoked)
1870-71	Franco-German war provides opportunity for US sardine business Beginning of winter herring fishery in Eastport-Lubec area (lasts only a few years)
1871	A peak of the smoked herring industry (or in 1870 - 18) – declined from here to 1890 Smoked fish no longer required inspection, did need to be labeled by processor
1874	First attempt at Russian sardines
1875	Sardine industry begins in Eastport (Eagle Preserved Fish - J Wolff) and confined there until 1880 Duty of $4 on imported sardines
1880	Sardine industry begins in Lubec Seining beginning to be used to limited success Practice of using cottonseed oil instead of olive oil generally in use Oven for drying sardines introduced by Henry Sellman – much less time than open-air Eastport controls frozen herring trade

Smoked herring trade fallen off greatly – due to destruction of trade with South (Civil War) followed by years of overproduction to drive down the price
Weir fishery in state of neglect prior to increase in canneries being built around this time
Drag-seining/purse seining introduced to limited success

1883 Process of bloater herrings begins in Gloucester with fish from Newfoundland

1884 Seining (haul and purse) not used very much until around 1884
Frying method using oil heated with coils of pipe introduced – advantageous over furnace heat
Concave top/bottom can makes venting unnecessary in qrtr size cans
Torching was popular method up until about this time – used for winter herring industry
Haul and purse seining being used more extensively due to increased demand

1885 Canneries organized to regulate price of fish @ weirs and of product (ave $5/hh) – worked until Nov when contract was broken and price of fish increased rapidly = competition
Pack of oil sardines low due to scarcity of sm fish for qrtr cans
Large quantity of cans dumped into market at $4.50/case – bad but "panic averted"
End of Washington treaty (June 30) and duty-free Canadian imports (duty of $2.50 enforced -3)
Smoked herring industry importance growing – regained importance

1886 Canadian weir fishermen have law successfully passed against seining
American law re:menhaden effectively shut out seining in US as well
Canneries organize to regulate price of fish @ weirs ($5/hh) and of products
Sardine industry overcrowded
Bad fire in Eastport destroys several factories – rebuilt but no increase until 1892
Weirs in early March – but fish too big for canneries, some smoked or used as bait
~Mid-Apr lg fish gone, sm herring abundant for sardines – sardine fishery at height by 1May
Law prohibiting canning after Dec. 15 in effect by this time
Birch-bark for torching decreased locally – have to go further to get it. Fishermen using cotton and kerosene – weir fishermen complaining about oil in water

Changes since 1880: Tails of sardines no longer removed
Stove rooms more complete, less drying in sun
Cooking on ovens done more often by steam than by furnace
Most cooking done in ovens vs. frying in oil
Concave cover instead of flat on quarter cans – no venting needed after bathing
American labels being used more often vs. "foreign" ones
Frozen herring/winter herring trade: height in Jan dep on weather, very abundant fish

1887 Fishing season for weirs: 1 April – Jan for canneries (prior only ~1 June – beg of Sept)
Increased weir construction in hopes of more constant supply to canneries – inconsistent supply due to inefficient fishing, not decrease or low abundance

1889 Apparatus introduced that converted solder bars to wire = decreased costs/waste
Also introduced conveyor belt and hot air chambers for cooking = less labor (use is "clumsy")
Introduced this yr? Machine to decorate plate in house – no need to ship, decrease cost
Smoked herring industry as extensive as ever by this year

1890 Apx this time practice of frying fish in oil replaced by cooking with steam – few still fry
Solder principle expense for canning industry
Smoked herring at low point – begins increasing again

1892 Some increase in the industry for first time since early 1880s.
Foreign companies sue over use of French label – suit lost but companies begin fine print

After 1892 – Field overcrowded again, prices dropped ($4/case in 1892 - $2/cases in 1896)

1893 Electricity introduced in factories

1895 Partial strike over an attempt to cut wages (summer) in Eastport & Lubec – lessened output
Capacity of factories exceeds demand, prices are very low
Few canneries in Canada (therefore no comp) due to tariff – only one near St. Andrews – Canadian market too small to support and industry

1896 Reported as very bad season – no reason given

1897	Canning season shortened by 40 days Other regulationsQuantity of oil, decoration of tin, # fish/can, manner of cooking, etc. Also reported as a bad year, some factories closed early
1898	Canning industry improved somewhat over 1897
1899	Apx this year canning plant for machine-made cans built in N Lubec by American Can Co. – for three-piece cans, previously all hand-made i.e. decrease employees/labor First syndicate attempted this year to regulate industry – two companies controlled most Smoked industry at peak again – new demand out west with defeat of Spanish First drawn cans introduced
1900	Sea Cost Packing gains control of AM Can CO. plant Russian sardine industry at height – declines from here to practically gone by 1913 Height of Eastport population – declined from here, some recovery w/1930s tidal project
1901	New fish driers introduced by Sea Coast Packing replacing reel ovens
1902	Sea Coast Packing introduces sealing machine to its factories this yr and next
1903	Machine-made and sealed cans generally replaced handmade soldered cans – increased prod (Cans made by independ company, decreased labor and cut costs – less men esp) Peak of smoked herring industry (apx) declined after this year Sea Coast Canning Co buys N. Lubec canning plant from Sea Coast Packing Approx failure of the syndicates Drawn cans in general use by now
1903-1914	Sardine industry reasonably prosperous – machines decreased employees but almost doubled output – new factories in operation along ME coast, inc comp + greater prod red prices some
1904	First drawn can made in factory – can plant in Eastport New sealing machines introduced

1905 First can plant in Lubec
Better sealing machines for independent factories

1906 American Can Co regains control of Sea Coast Canning Plant & produces all cans for SeaCoast

~1905-1910 Boats equipped with water-tight tanks for transportation of herring to canneries in which salt or pickle may be added for preservation – increased distance fish can be transported w/o spoiling

1907 French labels no longer used – Pure Food & Drug Act goes into effect
Child labor laws enacted but not enforced to great degree
ME statute forbids children under 14 in factories but permits in packing perishable goods as long as there is supervision. Law not enforced effectively for some yrs
N. Lubec American Can Co plant burns – rebuilt in Lubec village, 2-piece drawn cans
Sea Coast Canning builds new canning plant 1907-08 in Eastport – later sold to American Can
All cans bought are 2-piece drawn cans from American Can Co factories

1908 Flaking machines experimented with in some factories in Eastport – in use by now
Railroad comes to Eastport – competes with Boston steamship line

1913 Practically no Russian sardines produced by this time

1913-16 Some efforts to improve quality – study of packing process
Majority of canneries formed assoc to better industry and much was done to improve sanitation – ME sanitary legislation expanded and inspectors more numerous and vigilant

1914 3 Sardine manufacturers taken to court & fined over child labor law – stricter enforcement, removal of children from factories

1914-18 World War I – boom in sardine industry – cheap food, prices high

1915 Great decrease in French sardines
End of year embargo on Norwegian sardines – increase demand/price on US
Maine enacted laws restricting seining – need details

1917 Punch type machine discarded for 2 spindle Max Ams machines

1921	Dept of Agriculture publishes recommendations for canning industry – carried out 1913, 14, 16: faster improved carriers, refrigerated holds, improved handling and holding methods introduced All carriers converted to gas by this time Conditions "ripe" for increase of Russian sardine industry – foreign supply cut off Fish meal used globally for stock protein but not in lg extent in US – on the rise/encouraged
1920s	Slump in sardine industry with end of war – no recovery to previous levels during this time) Sharp depression in industry at begin of 1920s
1929-33	Economic collapse in US – Great Depression (continues to 1941): important impact – many plants closed, prod fell well below 1900 levels – Eastport/neighboring towns bankrupt → power project
~1930	Packing season Apr – Dec
1931	Last steamer leaves route
Mid-1930s	Some increase in downeast population (Eastport) due to tidal project
1936	Power project dropped by government.
1939-45	World War II: Boom in sardine industry again – govt purchased 80% of pack, Eastport and industry revived and maintained up to 1948 season.
1941	Official end of Great Depression, US enters WWII
1948	Poor run of herring, decline in business (esp with end of war) Herring used in pet food begins – encouraged by ME Devo Commission/Dept of Sea & Shore Fisheries – another market for catch and increased ME income
1950s	Catches decline
1961	Georges Bank fishery begins, intense pressure from USSR (gill nets)
1963-5	USSR diverts attention from herring to other species, herring still abundant
1964-65	Nova Scotia adult purse seine fishery beings Great increase in catch of adult herring off NS

1964-8	USSR primarily fishes with otter trawls
1966	Poland begins fishing for herring on Georges Bank
1967	Adult herring western GOM/Jeffreys Ledge fishery begins Germany begins fishing for herring, followed by Romania, Iceland, Japan, Norway, Bulgaria, and Cuba
1968	Peak of Georges Bank fishery, catch declines thereafter USSR introduces purse seines into the fishery
1971	German Democratic Republic introduces midwater trawling
1970s	Refocus on inshore waters, majority of catch taken nearshore and fixed gear dominant
1972	Internat'l Commission for NW Atlantic Fisheries (ICNAF) attempts to begin management of the 3 adult fisheries: Georges Bank, NS, western GOM Change to midwater gear – possibly from quotas enforced by ICNAF First national catch quotas made and management "begins in earnest"
1970-2	Adult herring Jeffreys Ledge fishery peaks, declines and collapses thereafter
1972-76	Herring is managed by the International Commission for the Northwest Atlantic Fisheries (ICNAF)
1976	U.S. declares 200 mile limit in federal fishery management plan (FMP) to help rebuild western GOM stocks
1976-78	NMFS regulates foreign fishing through a preliminary FMP
1977	Offshore Georges Bank herring fishery collapses
1978	U.S. adopts its own FMP for Atlantic herring to manage herring stocks on Georges Bank and in the Gulf of Maine to achieve higher levels of spawning biomass and stable recruitment, and to rebuild the juvenile herring resource and sardine fishery in the GOM.
1982	NMFS rescinds the 1978 FMP because of conflicts between state and federal regulations

Herring is placed on prohibited species list, eliminating directed fisheries for herring by foreign fleets within the U.S. EEZ and requiring any herring bycatch be discarded

1983 ASMFC adopts Interstate FMP for Atlantic Herring

Mid-1980s Georges Bank herring population begins to rebuild

Mid 1980s – 1990s Collapse of U.S. nearshore fixed gear fishery
Mobile gear gains in importance (purse seines, midwater trawlers)

1991 Improvements in assessment procedures, stocks combined into a single stock complex for management

1994 ASMFC adopts new FMP to address the growth of the herring resource and nternal Water Processing (IWP) operations
Mid-water trawling by U.S. and Canada begins

1996 Magnusen-Stevens Fishery and Conservation Act

1999 ASMFC adopts Amendment 1 to the Herring FMP to complement federal FMP.

2003 ASMFC and NEFMC develop new amendments to address to limited entry and other issues

2006 Limited entry for herring vessels implemented through Amendment 1 to the Herring FMP

APPENDIX E:

SCRIPT FOR CHAPTER II

```
maine37=scan("C:/Documaine37nts and Settings/Emily Klein/Desktop/Math &
Stats/Final Analysis/Data/Maine Data/maine37-06.txt")

maine37.ts=ts(me37, start=1937)

par(mfrow=c(1,1))
plot(maine37.ts)

par(mfrow=c(2,1))
acf(maine37.ts, 100)
pacf(maine37.ts, 100)
```

DIFFERENCING

```
maine37.diff = diff(maine37.ts)

par(mfrow=c(2,1))
acf(maine37.diff, 100)
pacf(maine37.diff, 100)

maine37.ln  = log(maine.37.ts)
maine37.lndf = diff(maine37.ln)

plot(maine37.lndf, main="Diff & Ln-Transformaine37d")
abline(c(0,0), col="red")
```

ACF/PACF of LN/DIFFERENCED DATA

```
acf(maine37.lndf, 100)
pacf(maine37.lndf, 100)
```

ARIMA FITS

```
maine37.ar=arima(maine37.lndf, order=c(1,0,0))
maine37.ma=arima(maine37.lndf, order=c(0,0,1))
maine37.arma=arima(maine37.lndf, order=c(1,0,1))
```

maine37.ma

```
Call:
arima(x = maine37.lndf, order = c(0, 0, 1))
Coefficients:
          ma1  intercept
      -0.4960     0.0150
s.e.   0.0902     0.0242
sigma^2 estimated as 0.1543:  log likelihood = -33.58,  aic = 73.16
```

maine37.ar

```
Call:
arima(x = maine37.lndf, order = c(1, 0, 0))
Coefficients:
          ar1  intercept
      -0.5836     0.0167
s.e.   0.1029     0.0289
sigma^2 estimated as 0.1428:  log likelihood = -30.96,  aic = 67.92
```

maine37.arma

```
Call:
arima(x = maine37.lndf, order = c(1, 0, 1))
Coefficients:
          ar1      ma1  intercept
      -0.5554  -0.0397     0.0166
```

```
s.e.    0.2009   0.2364      0.0283
sigma^2 estimated as 0.1427:  log likelihood = -30.95,  aic = 69.89
```

##AIC/AICc/SIC - LN/DIFF, ARIMA(0,1,1)

```
length(maine37.lndf)
[1] 69
```

ARIMA(0,1,1)

```
> #AICc
> log(maine37.ma$sigma2)+((69+1)/(102-1-2))
[1] -1.161626
> #SIC
> log(maine37.ma$sigma2)+((1*log(69))/69)
[1] -1.807333
```

ARIMA(1,1,0)

```
> #AICc
> log(maine37.ar$sigma2)+((69+1)/(102-1-2))
[1] -1.239409
> #SIC
> log(maine37.ar$sigma2)+((1*log(69))/69)
[1] -1.885116
```

ARIMA(1,1,1)

```
> #AICc
> log(maine37.arma$sigma2)+((69+2)/(102-2-2))
[1] -1.222361
> #SIC
> log(maine37.arma$sigma2)+((2*log(69))/69)
[1] -1.824123
```

DIAGNOSTICS

```
par(mfrow=c(2,1))
hist(maine37.ar$resid, br=20)
qqnorm(maine37.ar$resid)
qqline(maine37.ar$resid)

tsdiag(maine37.ar, gof.lag=20)
```

BACKCASTING FOR MAINE TIME SERIES

```
sillyrev = rev(maine37.ts)
revlndf = ts(diff(log(sillyrev)), start = 1937)
maine.rv.ar = arima(revlndf, order=c(1,0,0))

maine.rv.ar
Call:
arima(x = revlndf, order = c(1, 0, 0))
Coefficients:
          ar1  intercept
      -0.5836    -0.0167
s.e.   0.1029     0.0289
sigma^2 estimated as 0.1428:  log likelihood = -30.96,  aic = 67.92


pred1=predict(maine.rv.ar, n.ahead=1
pred2=pred1$pred

sillynew=ts(rep(0,71), end=2006)
```

```
log = log(sillyrev)
sillynew[1]= log[1]

for (i in 2:70){
sillynew[i] = sillynew[i-1] + revlndf[i-1]}

sillynew[71]=sillynew[70] + pred2[1]


## CHECKING

par(mfrow=c(1,1))
rev = ts(rev(maine37.ts), start=1937)

plot(sillynew)
lines(log(rev), col="blue")

maine.new = ts(rev(sillynew), end = 2006)
maine.new2=exp(maine.new)
```

1928:

```
maine.new2[1:3]
[1] 55368450 70683349 48187271
```

1924:

```
[1] 44780504 38988767 49773051 33932001
```

1919:

```
[1] 60214323 52426436 66927575 45626831
```

1900:

```
[1] 50987871 53038521 50580108 55619193 48425622 61820137 42144915
```

1890:

```
[1] 31405075 21409920
```

1881:

```
[1] 22161720 21134492 23240033 20234257 25831047 17609914
```

FINAL MODEL FOR COMPLETED MAINE TIME SERIES

```
maine.1880=scan("C:/Documents and Settings/Emily Klein/Desktop/Math &
Stats/Final Analysis/Data/Maine Data/ME1880-06.txt")
maine1880.ts=ts(maine.1880, start=1880)

plot(maine1880.ts)

par(mfrow=c(2,1))
acf(maine1880.ts, 100)
pacf(maine1880.ts, 100)

plot(maine1880.df)
abline(mean(maine1880.df),0, col="red")

maine1880.ln = log(maine1880.ts)
maine1880.lndf = diff(maine1880.ln)
par(mfrow=c(1,1))

plot(maine1880.lndf, main="Diff & Ln-Transformed")
abline(c(0,0), col="red")
```

```
acf(maine1880.lndf, 100)
pacf(maine1880.lndf, 100)

maine1880.ar=arima(maine1880.lndf, order=c(1,0,0))
maine1880.ma=arima(maine1880.lndf, order=c(0,0,1))
maine1880.arma=arima(maine1880.lndf, order=c(1,0,1))
```

maine1880.ma

```
Call:
arima(x = maine1880.lndf, order = c(0, 0, 1))
Coefficients:
          ma1  intercept
      -0.5435     0.0106
s.e.   0.0670     0.0160
sigma^2 estimated as 0.1523:  log likelihood = -60.41,  aic = 126.83
```

maine1880.ar

```
Call:
arima(x = maine1880.lndf, order = c(1, 0, 0))
Coefficients:
          ar1  intercept
      -0.5592      0.010
s.e.   0.0736      0.022
sigma^2 estimated as 0.1471:  log likelihood = -58.25,  aic = 122.49
```

maine1880.arma

```
Call:
arima(x = maine1880.lndf, order = c(1, 0, 1))
Coefficients:
          ar1      ma1  intercept
      -0.4106  -0.2183     0.0102
s.e.   0.1514   0.1684     0.0189
sigma^2 estimated as 0.1453:  log likelihood = -57.45,  aic = 122.89
```

##AIC/AICc/SIC - LN/DIFF, ARIMA(0,1,1)

```
length(maine1880.lndf)
[1] 126
```

ARIMA(0,0,1)

```
#AICc
> log(maine1880.ma$sigma2)+((126+1)/(126-1-2))
[1] -0.8492037
#SIC
> log(maine1880.ma$sigma2)+((1*log(126))/126)
[1] -1.843341
```

ARIMA(1,0,0)

```
#AICc
> log(maine1880.ar$sigma2)+((126+1)/(126-1-2))
[1] -0.8837981
#SIC
> log(maine1880.ar$sigma2)+((1*log(126))/126)
[1] -1.877935
```

ARIMA(1,0,1)

```
#AICc
> log(maine1880.arma$sigma2)+((126+2)/(126-2-2))
[1] -0.8800766
#SIC
> log(maine1880.arma$sigma2)+((2*log(126))/126)
[1] -1.852491
```

APPENDIX F:

BACKTRANSFORMATION OF MODELS

MAINE MODEL

y_i *the natural log-transformed and difference Maine time series*
x_t *the Maine pounds time series*

$y_i = 0.01 - 0.5592\ y_{i-1} + \hat{w}_i$ *the model for the transformed Maine series*
$y_i = 0.01 - 0.5592\ y_{i-1}$ $\hat{w}_i$ *can be ignored*
$y_i = \ln(x_t) - \ln(x_{t-1})$ y_i *is the transformed Maine series*

Therefore - $y_i = \ln(\frac{x_{t-1}}{x_t})$ *solving for* y_i

Solving for x_t -

$$\ln(\frac{x_{t-1}}{x_t}) = 0.010 - 0.5592[\ \ln(\frac{x_{t-1}}{x_t})\ ]$$

$$\frac{x_{t-1}}{x_t} = e\hat{}[\ 0.01 - 0.5592[\ \ln(\frac{x_{t-1}}{x_t}))\ ]$$

$x_t = x_{t-1} * [e\hat{}(0.01 - 0.5592(\ln(\frac{x_{t-1}}{x_t}))\ ]$ *final Maine herring pounds model*

CANADIAN MODEL

y_i *the natural log-transformed and difference Canadian time series*
x_t *the Canadian pounds time series*

$y_i = 0.0148 - 0.5480\ y_{i-1}$ *the model for the transformed Canadian series*
$y_i = \ln(x_t) - \ln(x_{t-1})$ y_i *is the transformed Canadian series*

Therefore - $y_i = \ln(\frac{x_{t-1}}{x_t})$ *solving for* y_i

Solving for x_t - *same as above*

$x_t = x_{t-1} * [e\hat{}\ (0.0148 - 0.5480 * w_{t-1})\ ]$ *final Maine herring pounds model*

where $w_{t-1} = y_{i-1(obs)} - y_{i-1(pred)}$

and $y_{i-1(obs)} = \ln(\frac{x_{t-1}}{x_t})$

APPENDIX G:

SCRIPT FOR CHAPTER III

Maine S+ and Intervention

Initial data exploration

```
metot.ts

par(mfrow=c(2,1))
acf(metot.ts, 100)
acf(metot.ts, 100, type="partial")

metot.lndf=diff(log(metot.ts))
plot(metot.lndf)

par(mfrow=c(2,1))
acf(metot.lndf, 100)
acf(metot.lndf, 100, type = "partial")
```

ARIMA MODEL - confirming model choice with R from Chap 2

```
metot.ar=arima.mle(x=metot.lndf, model=list(order=c(1,0,0)))
metot.ma=arima.mle(x=metot.lndf, model=list(order=c(0,0,1)))
metot.arma=arima.mle(x=metot.lndf, model=list(order=c(1,0,1)))
```

metot.ar

```
Call: arima.mle(x = metot.lndf, model = list(order = c(1, 0, 0)))
Method:  Maximum Likelihood
Model :  1 0 0
Coefficients:
      AR : -0.55863
Variance-Covariance Matrix:
            ar(1)
ar(1) 0.005503488
Optimizer has  converged
Convergence Type: relative function convergence
```

AIC: 117.4614

metot.ma

```
Call: arima.mle(x = metot.lndf, model = list(order = c(0, 0, 1)))
Method:  Maximum Likelihood
Model :  0 0 1
Coefficients:
      MA : 0.53884
Variance-Covariance Matrix:
            ma(1)
ma(1) 0.005632125
Optimizer has  converged
Convergence Type: relative function convergence
```

AIC: 123.25375

metot.arma

```
Call: arima.mle(x = metot.lndf, model = list(order = c(1, 0, 1)))
Method:  Maximum Likelihood
Model :  1 0 1
Coefficients:
      AR : -0.40792

      MA : 0.22084
Variance-Covariance Matrix:
           ar(1)      ma(1)
ar(1) 0.02004488 0.01749155
ma(1) 0.01749155 0.02287330
Optimizer has  converged
Convergence Type: relative function convergence
```

AIC: 117.73119

 # INTERVENTIONS

```
metot.rob = arima.rob((metot.lndf)~1, p=1, innov.outlier = T)
```

metot.rob

```
Call:
arima.rob(formula = (metot.lndf) ~ 1, p = 1, innov.outlier = T)

Regression Coefficients:
 (Intercept)
 0.0394

AR Coefficients:
   AR(1)
 -0.9751
Degrees of freedom: 126 total; 124 residual
Innovations standard deviation: 0.2709

 Number of outliers detected:  3

Outlier index
[1]  71  84 102

Outlier type
[1] "IO" "AO" "AO"

Outlier impact
[1] -1.0030 -0.9661 -0.9772

Outlier t-statistics
[1] 3.1840 4.1972 4.1390
> summary(metot.rob)

Call:
arima.rob(formula = (metot.lndf) ~ 1, p = 1, innov.outlier = T)

Regression model:
 (metot.lndf) ~ 1

ARIMA model:
Ordinary differences: 0 ; AR order: 1 ; MA order: 0

Regression Coefficients:
             Value Std. Error t value Pr(>|t|)
(Intercept) 0.0394 0.0160     2.4594  0.0153

AR Coefficients:
         Value Std. Error  t value Pr(>|t|)
AR(1)  -0.9751    0.0211   -46.1262   0.0000

Degrees of freedom: 126 total; 124 residual

Innovations standard deviation: 0.2709

 Number of outliers detected:  3

Outliers detected:

       |Time    |Type    |Impact |t-value|
-------+--------+--------+-------+-------+
```

```
1      |1951   |IO     |-1.003 |3.184  |
-------+-------+-------+-------+-------+
2      |1964   |AO     |-0.9661|4.197  |
-------+-------+-------+-------+-------+
3      |1982   |AO     |-0.9772|4.139  |
-------+-------+-------+-------+-------+

Innovation scale estimate before correcting outliers:
  0.3244

Innovation scale estimate after correcting outliers:
  0.3525
```

Canada S+ and Intervention

Initial data exploration

```
can1871.ts = timeSeries(Can1871.07, pos=timeCalendar(y=1871:2007,
   format="%Y"))
plot.timeSeries(can1871.ts)

par(mfrow=c(2,1))
acf(can1871.ts, 100)
acf(can1871.ts, 100, type="partial")

can1871.lndf=diff(log(can1871.ts))
plot(can1871.lndf)

par(mfrow=c(2,1))
acf(can1871.lndf, 100)
acf(can1871.lndf, 100, type = "partial")
```

ARIMA MODEL - confirming model choice with R from Chap 2

```
can1871.ar=arima.mle(x=diff(log(can1871.ts)), model=list(order=c(1,0,0)))
can1871.ma=arima.mle(x=diff(log(can1871.ts)), model=list(order=c(0,0,1)))
can1871.arma=arima.mle(x=diff(log(can1871.ts)), model=list(order=c(1,0,1))) >

summary.arima(can1871.ar)
Call:
arima.mle(x = can1871.lndf, model = list(order = c(1, 0, 0)))
Method:  Maximum Likelihood  with likelihood conditional on  1 observations
ARIMA order:  1 0 0
       Value Std. Error t-value
ar(1) -0.383     0.0795  -4.817

Variance-Covariance Matrix:
            ar(1)
ar(1) 0.006320968

Estimated innovations variance:  0.1372

Optimizer has  converged
Convergence Type: relative function convergence
AIC: 116.9455

Time period: from 1872 to 2007

summary.arima(can1871.ma)
```

```
Call: arima.mle(x = can1871.lndf, model = list(order = c(0, 0, 1)))
Method:  Maximum Likelihood  with likelihood conditional on  0 observations

ARIMA order:  0 0 1
       Value Std. Error t-value
ma(1) 0.5323    0.07259   7.333
Variance-Covariance Matrix:
            ma(1)
ma(1) 0.005269364

Estimated innovations variance:  0.1282

Optimizer has  converged
Convergence Type: relative function convergence
```

AIC: 108.903

```
Time period: from 1872 to 2007

summary.arima(can1871.arma)
Call: arima.mle(x = can1871.lndf, model = list(order = c(1, 0, 1)))
Method:  Maximum Likelihood  with likelihood conditional on  1 observations
ARIMA order:  1 0 1

        Value Std. Error t-value
ar(1) 0.06962      0.1615   0.431
ma(1) 0.57970      0.1319   4.394

Variance-Covariance Matrix:
           ar(1)      ma(1)
ar(1) 0.02609102 0.01805150
ma(1) 0.01805150 0.01740731

Estimated innovations variance:  0.1283

Optimizer has  converged
Convergence Type: relative function convergence
```

AIC: 110.3403

 # INTERVENTIONS

```
can1871.rob = arima.rob((can1871.lndf)~1, q=1, innov.outlier=T)
can1871.rob

> summary(can1871.rob)

Call:
arima.rob(formula = (can1871.lndf) ~ 1, q = 1, innov.outlier = T)

Regression model:
 (can1871.lndf) ~ 1

ARIMA model:
Ordinary differences: 0 ; AR order: 0 ; MA order: 1

Regression Coefficients:
             Value Std. Error t value Pr(>|t|)
(Intercept) 0.0294 0.0198     1.4853  0.1398

MA Coefficients:
       Value Std. Error t value Pr(>|t|)
MA(1) 0.0748 0.0976     0.7659  0.4451
```

Degrees of freedom: 136 total; 134 residual

Innovations standard deviation: 0.217

Number of outliers detected: 67

Outliers detected:

	Time	Type	Impact	t-value
1	1874	AO	0.2852	4.945
2	1875	IO	0.1483	9.025
3	1881	IO	-0.7588	4.227
4	1882	AO	1.386	6.715
5	1886	IO	-0.1484	8.643
6	1890	AO	-0.5776	3.906
7	1893	IO	-0.3655	3.882
8	1894	IO	0.4433	3.584
9	1896	IO	-0.2369	5.29
10	1897	IO	0.4834	3.788
11	1900	IO	-0.8382	4.548
12	1901	IO	0.3948	3.836
13	1906	IO	-0.1541	7.877
14	1907	IO	-0.1966	7.033
15	1908	IO	-0.2321	5.474
16	1909	IO	0.1541	7.543
17	1910	IO	-0.2784	5.296
18	1911	IO	0.6637	4.01
19	1913	AO	-0.2982	4.399
20	1915	AO	0.2234	5.864
21	1917	IO	0.1486	8.281
22	1919	IO	-0.2302	5.723
23	1921	IO	-0.4906	3.732
24	1922	IO	0.4214	3.771
25	1923	IO	-0.6646	3.908
26	1924	IO	0.6635	4.12
27	1925	IO	-0.2517	5.039

28	1926	IO	-1.755	8.088
29	1927	IO	1.575	7.43
30	1936	IO	0.1412	9.442
31	1937	AO	-0.4259	3.683
32	1939	IO	0.4919	3.636
33	1940	IO	-0.1975	6.762
34	1941	IO	-1.136	5.85
35	1942	AO	1.289	6.41
36	1945	IO	-0.1698	7.605
37	1946	IO	0.3	4.205
38	1947	IO	-0.9144	4.832
39	1953	IO	-0.202	6.335
40	1954	IO	0.2839	5.141
41	1955	IO	-0.6164	4.044
42	1956	AO	0.2938	4.624
43	1957	IO	0.2053	6.159
44	1958	AO	0.3318	4.049
45	1959	IO	0.1982	6.495
46	1961	IO	-0.3192	4.066
47	1962	AO	0.1441	9.215
48	1964	AO	0.3721	3.793
49	1965	IO	0.2904	4.776
50	1966	AO	0.4338	3.632
51	1968	IO	0.1671	7.827
52	1969	AO	-0.5228	3.646
53	1970	IO	-0.154	8.217
54	1971	IO	-0.6704	3.836
55	1972	IO	0.6586	4.204
56	1976	IO	-0.3125	4.176
57	1979	IO	-0.4174	3.886
58	1985	IO	0.1923	7.513

```
59     |1988   |IO     | 0.1827|7.469  |
-------+-------+-------+-------+-------+
60     |1989   |IO     |-0.1819|7.788  |
-------+-------+-------+-------+-------+
61     |1991   |AO     |-0.3345|3.908  |
-------+-------+-------+-------+-------+
62     |1993   |IO     |-0.1941|7.254  |
-------+-------+-------+-------+-------+
63     |1994   |AO     |-0.3594|4.011  |
-------+-------+-------+-------+-------+
64     |1995   |AO     |-0.2132|6.144  |
-------+-------+-------+-------+-------+
65     |1998   |IO     | 0.219 |6.005  |
-------+-------+-------+-------+-------+
66     |2005   |AO     |-0.4976|3.569  |
-------+-------+-------+-------+-------+
67     |2007   |AO     | 0.2391|5.051  |
-------+-------+-------+-------+-------+

Innovation scale estimate before correcting outliers:
  0.217

Innovation scale estimate after correcting outliers:
  0.01875
```

DECREASED SENSITIVITY

```
can1871.rob2 = arima.rob((can1871.lndf)~1, q=1, innov.outlier=T, critv=4)

can1871.rob2
Call:
arima.rob(formula = (can1871.lndf) ~ 1, q = 1, innov.outlier = T, critv = 4)

Regression Coefficients:
 (Intercept)
 0.0294

MA Coefficients:
  MA(1)
 0.0748

Degrees of freedom: 136 total; 134 residual
Innovations standard deviation: 0.217

 Number of outliers detected:  8

Outlier index
[1] 10 11 29 55 56 70 71 76

Outlier type
[1] "IO" "AO" "IO" "IO" "IO" "IO" "AO" "IO"

Outlier impact
[1] -0.7588  1.3856 -0.8382 -1.7555  1.5747 -1.1362  1.2888 -0.9144

Outlier t-statistics
[1] 4.2273 6.7153 4.5479 8.0884 7.4300 5.8502 6.4100 4.8324


> summary(can1871.rob2)

Call:
```

```
arima.rob(formula = (can1871.lndf) ~ 1, q = 1, innov.outlier = T, critv = 4)

Regression model:
 (can1871.lndf) ~ 1

ARIMA model:
Ordinary differences: 0 ; AR order: 0 ; MA order: 1

Regression Coefficients:
             Value  Std. Error t value Pr(>|t|)
(Intercept) 0.0294 0.0198     1.4853  0.1398

MA Coefficients:
       Value  Std. Error t value Pr(>|t|)
MA(1) 0.0748 0.0976     0.7659  0.4451

Degrees of freedom: 136 total; 134 residual

Innovations standard deviation: 0.217

 Number of outliers detected:  8

Outliers detected:

       |Time   |Type   |Impact |t-value|
-------+-------+-------+-------+-------+
1      |1881   |IO     |-0.7588|4.227  |
-------+-------+-------+-------+-------+
2      |1882   |AO     | 1.386 |6.715  |
-------+-------+-------+-------+-------+
3      |1900   |IO     |-0.8382|4.548  |
-------+-------+-------+-------+-------+
4      |1926   |IO     |-1.755 |8.088  |
-------+-------+-------+-------+-------+
5      |1927   |IO     | 1.575 |7.43   |
-------+-------+-------+-------+-------+
6      |1941   |IO     |-1.136 |5.85   |
-------+-------+-------+-------+-------+
7      |1942   |AO     | 1.289 |6.41   |
-------+-------+-------+-------+-------+
8      |1947   |IO     |-0.9144|4.832  |
-------+-------+-------+-------+-------+

Innovation scale estimate before correcting outliers:
  0.217

Innovation scale estimate after correcting outliers:
  0.1802
```

APPENDIX H:

SCRIPT FOR CHAPTER IV

```
## CROSS CORRELATION ANALYSIS

## Combing data, creating time series

sabs24.06.ts=timeSeries(sabs24.06, pos=timeCalendar(y=1924:2006, format="%Y"))
sabs24.06.ts

can24.06.ts=timeSeries(Can24.06, pos=timeCalendar(y=1924:2006, format="%Y"))
can24.06.ts

## Transforming for stationarity

sabs24.06sst.diff=diff(sabs24.06.ts)
can24.06.diff=diff(can24.06.ts)
maine24.06.diff=diff(maine24.06.ts)
maine24.06.lndf=diff(log(maine24.06.ts))

ME05.06.diff=diff(ME05.06)
ME05.06.lndf=diff(log(ME05.06))
length(ME05.06.diff)

## SERIES MERGE

can24.06.lndf=diff(log(can24.06.ts))

sabs.sal.tot=timeSeries(sabs.sal24.06.tot, pos=timeCalendar(y=1924:2006,
   format="%Y"))
x=seriesMerge(sabs24.06sst.diff, can24.06.diff, maine24.06.diff,
   maine24.06.lndf)
x=seriesMerge(sabs24.06sst.diff, maine24.06.lndf, sabs24.06sst.diff)
x=seriesMerge(sabs.sal.tot, sabs24.06sst.diff, maine24.06.lndf,can24.06.lndf)

## Pounds.3 = Maine, Pounds.4 = Canada


## ORDINARY LEAST SQUARES

t=1924:2006

## With Model parameters: Salinity = none, SST = AR(1), ME = AR(1), Can = none

## MAINE

# Maine x Salinity

Sal.ME.fit = OLS(Pounds.3~Salinity + ar(1), data=x, correction="nw")
summary(Sal.ME.fit)

> summary(Sal.ME.fit)
Call:
OLS(formula = Pounds.3 ~ Salinity + ar(1), data = x, correction = "nw")

Residuals:
     Min      1Q  Median      3Q     Max
 -1.0147 -0.1674  0.0675  0.2318  0.6577

Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept) -7.0063  6.1220    -1.1444  0.2559
   Salinity  0.2207  0.1925     1.1463  0.2552
```

```
      lag1 -0.6169  0.0885    -6.9711  0.0000

Regression Diagnostics:

         R-Squared 0.3904
Adjusted R-Squared 0.3748
Durbin-Watson Stat 2.0393

Residual Diagnostics:
              Stat P-Value
Jarque-Bera  9.2609  0.0098
  Ljung-Box 15.4690  0.6924

Residual standard error: 0.3785 on 78 degrees of freedom
Time period: from 1926 to 2006
F-statistic: 24.98 on 2 and 78 degrees of freedom, the p-value is 4.139e-009
```

Maine x SST

```
SST.ME.fit = OLS(Pounds.3 ~ SST + ar(1) + tslag(SST), data = x, correction =
   "nw")
summary(SST.ME.fit)
Call:
OLS(formula = Pounds.3 ~ SST + ar(1) + tslag(SST), data = x, correction = "nw")

Residuals:
     Min      1Q  Median      3Q     Max
 -1.0452 -0.1735  0.0784  0.2247  0.6994

Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept)  0.0099  0.0425     0.2327  0.8166
        SST  0.0524  0.0707     0.7411  0.4609
 tslag(SST)  0.0402  0.0731     0.5498  0.5841
       lag1 -0.6183  0.0896    -6.9027  0.0000

Regression Diagnostics:

         R-Squared 0.3852
Adjusted R-Squared 0.3613
Durbin-Watson Stat 2.0197

Residual Diagnostics:
              Stat P-Value
Jarque-Bera 10.2788  0.0059
  Ljung-Box 13.5630  0.8085

Residual standard error: 0.3825 on 77 degrees of freedom
Time period: from 1926 to 2006
F-statistic: 16.08 on 3 and 77 degrees of freedom, the p-value is 3.282e-008
```

CANADA

Canada x Salinity

```
sal.can.fit = OLS(Pounds.4~Salinity, data=x, correction="nw")
summary(sal.can.fit)
Call:
OLS(formula = Pounds.4 ~ Salinity, data = x, correction = "nw")

Residuals:
```

```
     Min      1Q  Median      3Q     Max
 -1.6714 -0.1649  0.0134  0.1509  1.7018

Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept)  4.3835  6.8121     0.6435  0.5217
   Salinity -0.1376  0.2142    -0.6424  0.5224

Regression Diagnostics:

         R-Squared  0.0051
Adjusted R-Squared -0.0073
Durbin-Watson Stat  2.6707

Residual Diagnostics:
                Stat  P-Value
Jarque-Bera 134.1155   0.0000
  Ljung-Box  32.4103   0.0281

Residual standard error: 0.4218 on 80 degrees of freedom
Time period: from 1925 to 2006
F-statistic: 0.4127 on 1 and 80 degrees of freedom, the p-value is 0.5224
```

Canada x SST

```
SST.can.fit = OLS(Pounds.4 ~ SST + tslag(SST), data = x, correction = "nw")
summary(SST.can.fit)

Call:
OLS(formula = Pounds.4 ~ SST + tslag(SST), data = x, correction = "nw")

Residuals:
     Min      1Q  Median      3Q     Max
 -1.7395 -0.1251  0.0139  0.1679  1.7169

Coefficients:
              Value Std. Error t value Pr(>|t|)
(Intercept)  0.0117  0.0474     0.2468  0.8057
        SST -0.0314  0.0788    -0.3980  0.6917
 tslag(SST) -0.0403  0.0813    -0.4961  0.6212

Regression Diagnostics:

         R-Squared  0.0038
Adjusted R-Squared -0.0217
Durbin-Watson Stat  2.5660

Residual Diagnostics:
                Stat  P-Value
Jarque-Bera 164.9544   0.0000
  Ljung-Box  36.5829   0.0089

Residual standard error: 0.4263 on 78 degrees of freedom
Time period: from 1926 to 2006
F-statistic: 0.1502 on 2 and 78 degrees of freedom, the p-value is 0.8608
```