

Time Series Analysis and the Forecasting of Menhaden Catch and CPUE

A. L. JENSEN

School of Natural Resources
University of Michigan
Ann Arbor, Michigan 48109

ABSTRACT

Catch and catch per unit of effort (CPUE) data for Atlantic menhaden (*Brevoortia tyrannus*) and the gulf menhaden (*Brevoortia patronus*) of the Gulf of Mexico were analyzed with autocorrelation analysis to test for time lags and to develop forecasting equations. Analysis of data for two closely related species and for CPUE, as well as catch, provides an opportunity to evaluate time-series methods. The results for the two species and for catch and CPUE data were considerably different. A 6-year time lag was found for both the Atlantic menhaden catch and the gulf catch but the correlation was positive in the Atlantic and negative in the Gulf. A 3-year time lag, corresponding to the age at maturity, was found for the Atlantic menhaden CPUE. No significant time lags were found for CPUE in the Gulf of Mexico. Environmental factors such as changes in sea temperature could have caused the observed delayed density dependence but the time lags also could have arisen from a reproductive delay. The autocorrelation structure of the catch and CPUE data were applied to develop autoregressive models for year-ahead forecasts and confidence intervals. The coefficients of determination for the catch series were high for both the gulf menhaden and the Atlantic menhaden but the confidence intervals for year-ahead predictions were wide.

Reliable forecasts of catch are essential for fisheries management, and time-series analysis is an economical method for forecasting catches that could be widely applied (as one of several methods) in fishery management. Time-series analysis provides a measure of the precision of forecasts and describes the time structure of the catch data. This analysis has been applied successfully to several fisheries (Jensen 1976; Van Winkle et al. 1979; Saila et al. 1980; Mendelssohn 1981; and Stocker and Hilborn 1981). In this study, time-series analysis was applied to analyze the time structure, evaluate the predictability, and develop equations for year-ahead forecasting of the Atlantic menhaden (*Brevoortia tyrannus*) and gulf menhaden (*Brevoortia patronus*) fisheries.

The catch histories of the Atlantic Ocean and Gulf of Mexico menhaden fisheries are considerably different. The fishery developed earlier in the Atlantic than in the Gulf; the catch of Atlantic menhaden rose to a peak from 1955 to 1958 and then gradually declined to a level similar to that which existed prior to 1952 (Fig. 1A). In the Gulf of Mexico, there has been a gradual increase in yield (Fig. 2A). A comparison of results for these two closely related fisheries and for catch and CPUE provides an opportunity to evaluate time-series methods.

Time series such as catch records can be analyzed using several different methods of analysis (Box and Jenkins 1960); two widely applied methods are correlation analysis developed by Yule (1927) and spectral analysis (Box and Jenkins 1976). Spectral analysis assumes that the underlying process can be described in terms of sine and cosine functions, an assumption that is not reasonable for populations (Moran 1953). Autocorrelation analysis makes few assumptions about the underlying process and seems more appropriate for ecological analyses (Moran 1953). In autocorrelation analysis of a catch series, it is assumed that the sequence of catches is a realization of jointly distributed random variables with some joint probability function. The correlogram of autocorrelation analysis and the spectrum of spectral analysis are identical mathematically (Box and Jenkins 1976).

In the approach to time-series analysis proposed by Box and Jenkins (1976), which is now popular, both moving average terms and autoregressive terms are tested in a standardized procedure that overfits the data to seek the best-fitting model. The resulting model usually has no physical interpretation and provides little or no understanding; it is a tool for forecasting. Autoregressive terms are easily interpreted in terms

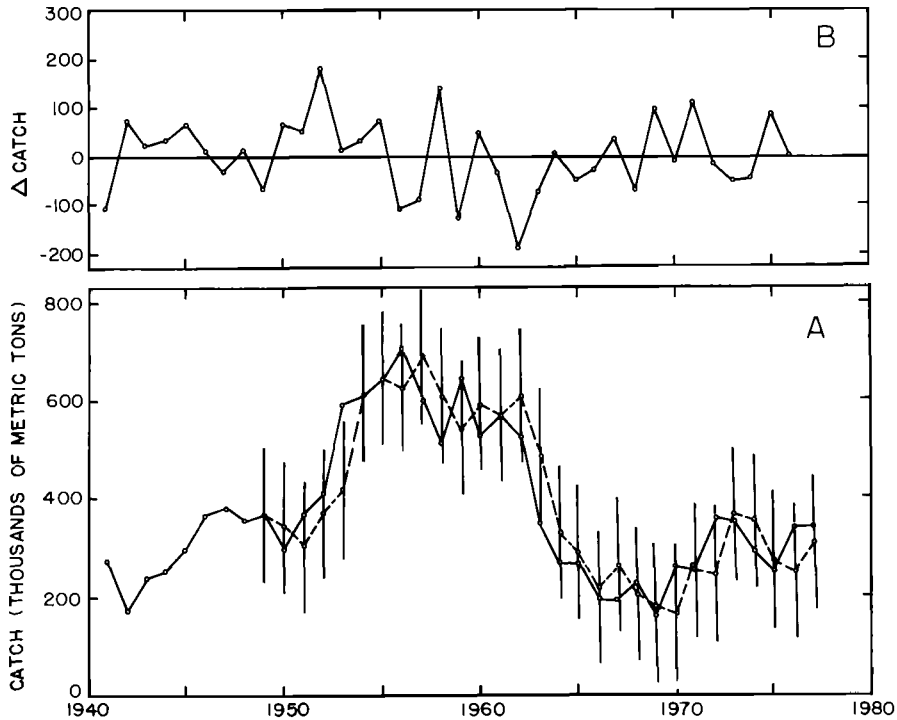


Figure 1. (A) Observed catch of Atlantic menhaden (solid line) and year-ahead forecasts of catch (dashed line). The vertical lines are 95% confidence limits for year-ahead predictions. (B) First difference of catch of Atlantic menhaden.

of such biological phenomenon as reproductive time lags, whereas moving average terms that are based on differences between predicted and observed values of a series are not easily interpreted. Moving average terms always can be expressed as a series of autoregressive terms but are used because they may result in simpler models (Box and Jenkins 1976). In this study, the objective is to identify biologically meaningful time lags and only autoregressive terms are used. This approach was used by Moran (1953). The small increase in the coefficient of determination with increase in the order of the model indicates that moving average terms would not increase the precision of forecasts. The data for menhaden were provided by P. J. Doody, Zapata Haynie Corporation, Houston, Texas.

METHODS

Methods of autocorrelation analysis are applicable only to stationary time series. In a stationary series, the data fluctuate about some mean

level and the mean, variance, and autocovariances are not dependent on time. Two tools of autocorrelation analysis are the autocorrelation function and the partial autocorrelation function. The correlation between C_t and C_{t-p} over all values of t is termed an autocorrelation with time lag p . For a stationary time series, the autocorrelation depends only on the lag. The set of autocorrelations associated with a time series is called the autocorrelation function. The partial autocorrelations, analogous to the partial correlations of multiple linear regressions, are the correlations between the time-series values at time t and a period k time units earlier, $t - k$, taking into consideration the correlation between C_t and C_{t-p} as p goes from 1 to $k - 1$. A plot of the partial autocorrelations against the lag is called the partial autocorrelation function.

Under the hypothesis that a time-series process is autoregressive of order p , the estimated partial autocorrelations of order $p + 1$ and higher are approximately independently distributed with

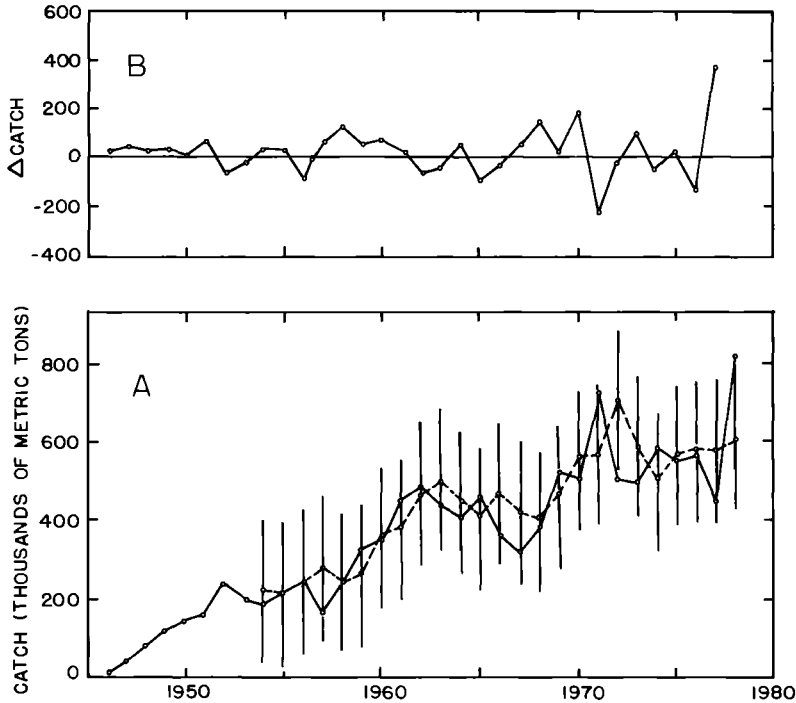


Figure 2. (A) Observed catch of gulf menhaden (solid line) and year-ahead forecasts of catch (dashed line). The vertical lines are 95% confidence limits for year-ahead predictions. (B) First difference of catch of gulf menhaden.

a variance of $1/T$ where T is the length of the series (Box and Jenkins 1976). Partial autocorrelations more than 1.65 standard deviations away from zero are significant at the 0.10 level (Box and Jenkins 1976). An alpha of 0.10 is used to indicate significance because the series are of relatively short length.

The raw series for menhaden catch (C_t) were nonstationary but the first differences of the series, $Z_t = C_t - C_{t-1}$ for $t = 2, 3, \dots, T$, were stationary. A p^{th} order autoregressive model for a differenced series Z_t is

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + e_t \quad (1)$$

where:

- C_t = catch for year t ,
- $Z_t = C_t - C_{t-1}$, the first difference of the catch series,
- ϕ_i = parameters,
- e_t = independently and identically distributed random variables with mean zero.

For forecasting catch, Equation (1) can be written as (Box and Jenkins 1976):

$$C_t = (1 - \phi_1)C_{t-1} + (\phi_2 - \phi_1)C_{t-2} + \dots + (\phi_p - \phi_{p-1})C_{t-p} + \phi_p C_{t-p-1} + e_t \quad (2)$$

The predictability of a time series of order j can be evaluated with the coefficient of determination R_j^2 which is calculated iteratively for autoregressive models as (Nelson 1976):

$$\begin{aligned} R_1^2 &= \phi_1^2 \\ R_j^2 &= \phi_j^2 (1 - R_{j-1}^2) + R_{j-1}^2, \\ j &= 1, 2, 3, \dots \end{aligned} \quad (3)$$

The coefficient of determination R_j^2 gives the proportion of the total variation predicted by an autoregressive model of order j . A 95% confidence interval for a 1-year-ahead forecast C_t made with a p^{th} order autoregressive model is calculated as $C_t \pm 1.96S$, where

$$\begin{aligned} S^2 &= S_0^2 - \sum_{i=1}^p \phi_i r_i S_0^2 \\ S_0^2 &= \frac{1}{T} \sum_{t=1}^T (C_t - \bar{C})^2 \end{aligned} \quad (4)$$

The r_i are autocorrelations and the ϕ_i are the coefficients of the autoregressive model of order

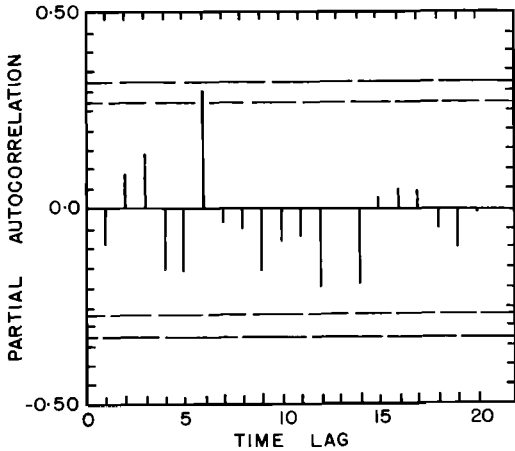


Figure 3. Partial autocorrelations for first differences of Atlantic menhaden catch. The dashed lines indicate statistically significant lags ($\alpha = 0.05$ and $\alpha = 0.10$).

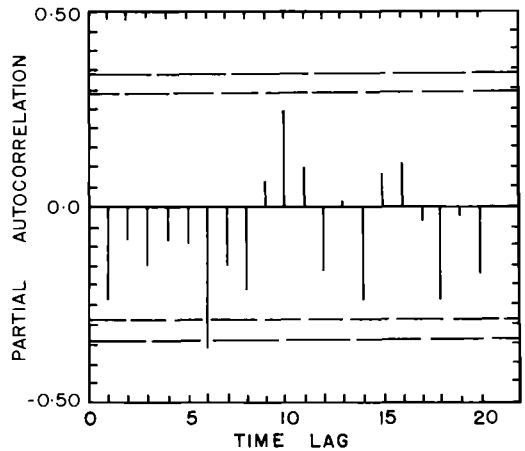


Figure 4. Partial autocorrelations for first differences of gulf menhaden catch. The dashed lines indicate statistically significant lags ($\alpha = 0.05$ and $\alpha = 0.10$).

p. Model parameters were estimated with the University of Michigan Interactive Data Analysis System which uses the Yule-Walker equations (Box and Jenkins 1976).

RESULTS AND DISCUSSION

Neither the gulf menhaden nor the Atlantic menhaden catch series is stationary but the first difference of both series appears to be stationary and the changes in catch fluctuate about zero, with the magnitude of fluctuation somewhat higher at higher yields (Figs. 1B and 2B).

The partial autocorrelations for the differenced catch series are complex and none of the autocorrelations are large. The largest and only statistically significant partial autocorrelations occur at a lag of 6 years in both the Atlantic and gulf menhaden fisheries (Figs. 3 and 4). The positive partial autocorrelation at a lag of 6 years in the Atlantic menhaden fishery indicates a tendency to cycle over a 6-year period. The correlation is negative for the Gulf. In an earlier study of the Atlantic menhaden, Jensen (1976) analyzed the raw series and failed to detect time lags.

Many analyses have been based on catch data but variation in catch may not be closely related to variation in abundance. Abundance is often assumed to be proportional to catch per unit of effort although recent work indicates the relation between abundance and CPUE may be more complex (Bannerot and Austin 1983). The At-

lantic menhaden CPUE data are considerably different from the gulf menhaden CPUE data (Figs. 5A and 6A) and the CPUE data for both fisheries are considerably different from the catch data. In the Atlantic, CPUE appears to increase slowly. The first difference of CPUE in the Atlantic menhaden fishery appears to be stationary (Fig. 5B). In the Gulf, the CPUE decreased consistently as catches increased. The first difference of the CPUE appears stationary for the gulf menhaden fishery (Fig. 6B).

The partial autocorrelations for Atlantic menhaden CPUE indicate a high negative autocorrelation at 3 years (Fig. 7), suggesting a delayed density dependence of 3 years which corresponds with the 3-year age at maturity of menhaden. Abundance of menhaden is negatively correlated with abundance one reproductive cycle earlier; small year classes tend to produce larger year classes. No correlations are significant for the differenced gulf menhaden CPUE but the highest correlations occur at a lag of 9–12 years, similar to the pattern observed for the differenced catch data (Fig. 8).

Time lags observed in fish catch have been associated with environmental variables. Regular fluctuations have been reported in the catch of several species of fish in the Gulf of St. Lawrence and along the eastern coast of North America. Koo (1970) reported a pattern in the catch of striped bass (*Morone saxatilis*), with alter-

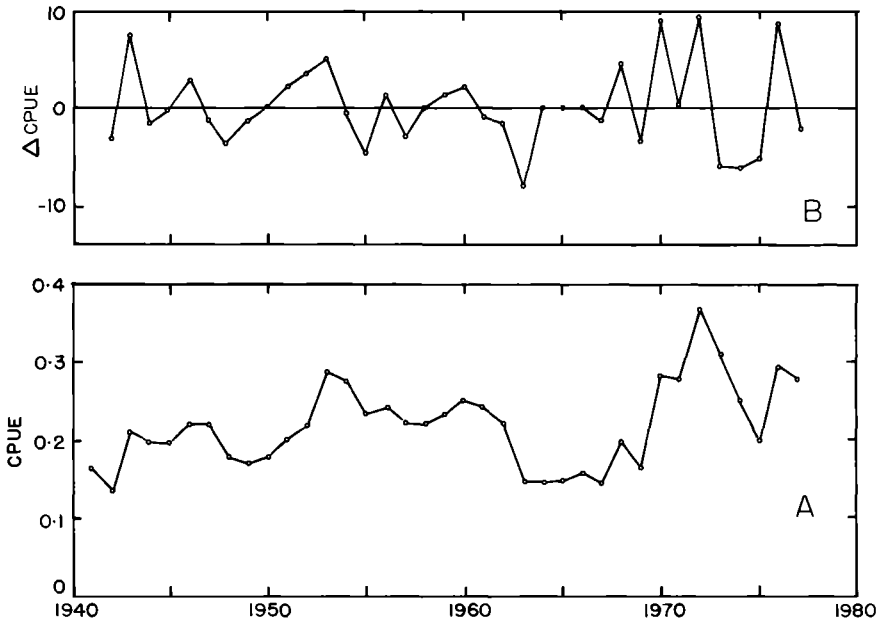


Figure 5. (A) Observed CPUE of Atlantic menhaden. (B) First difference of CPUE of Atlantic menhaden.

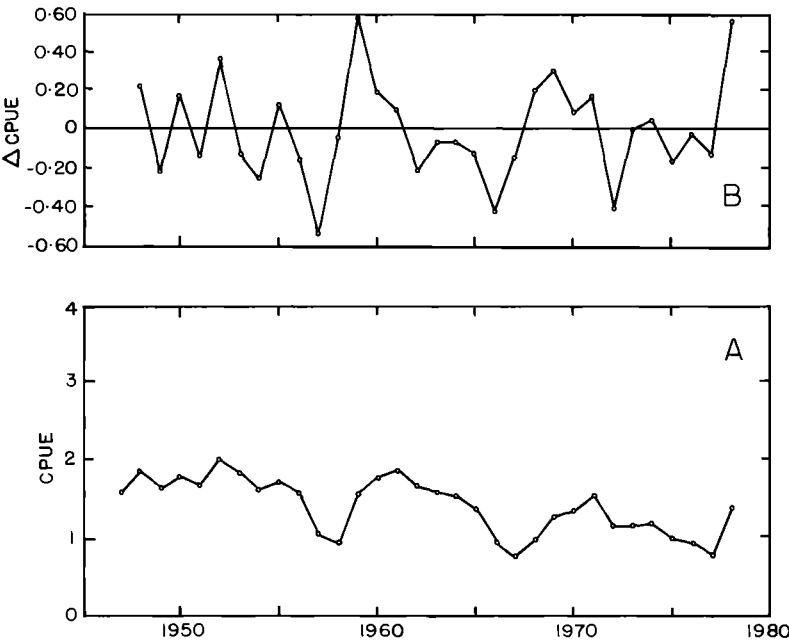


Figure 6. (A) Observed CPUE of gulf menhaden. (B) First difference of CPUE of gulf menhaden.

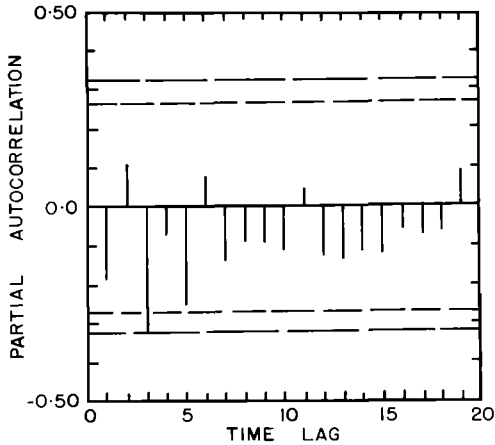


Figure 7. Partial autocorrelations for first differences of Atlantic menhaden CPUE. The dashed lines indicate statistically significant lags ($\alpha = 0.05$ and $\alpha = 0.10$).

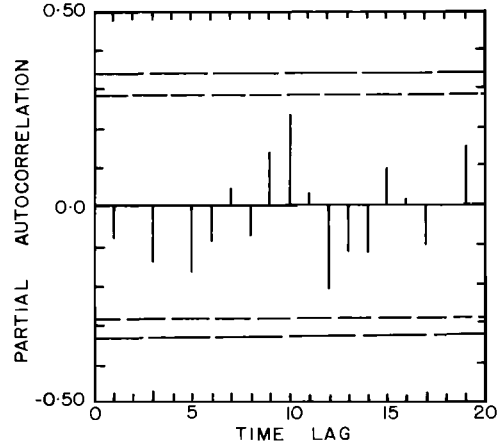


Figure 8. Partial autocorrelations for first differences of gulf menhaden CPUE. The dashed lines indicate statistically significant lags ($\alpha = 0.05$ and $\alpha = 0.10$).

nately high and low landings at fairly regular intervals of 6 years. He could find no intrinsic biological factor that operated over a 6-year time period and concluded the physical factors that enhanced the survival of young were likely to be operative. Van Winkle et al. (1979) applied time-series analysis to striped bass catch data and detected weak 20-year and 6–8-year periodicities. They concluded that the likely cause was environmental factors that enhanced survival of young rather than intrinsic factors of the life history. Yields of the Maine lobster (*Homarus americanus*) fishery are related to water temperature, and the inshore catch of Maine lobster can be predicted with precision ($R^2 = 0.76$) from the mean annual temperature 6 years in the past (Dow 1969; Flowers and Salla 1972). A lag of 9 years was found in the Atlantic lobster catches and a lag of 10 years was found in the Atlantic halibut (*Hippoglossus hippoglossus*) catch (Sutcliffe 1973). Atlantic menhaden landings in New England are correlated with sea temperature at St Andrews (Canada) 3 years in the past (Sutcliffe et al. 1977). Lett et al. (1975) found that both stock biomass and temperature were associated with recruitment of Atlantic cod (*Gadus morhua*) in the Gulf of St. Lawrence. Skud (1982) proposed that, although catch is correlated with environmental factors, the correlation between catch of one species and a physical factor such as temperature

may depend on the abundance of other species in the area.

Although the observed time lags might be associated with environmental factors, delayed density dependence could arise entirely from the delay between conception and maturity or through the combined effects of environmental factors and a reproductive time lag. Hutchinson (1948) developed a model in which cycles in abundance resulted from delayed density dependence, Ricker (1954) developed a model in which a reproductive time lag caused fluctuations in population size, and Nicholson (1957) demonstrated experimentally that delayed density dependence resulting from a reproductive time lag produced cycles in abundance of blowfly (*Lucilia curprina*) populations.

The partial autocorrelations of the catch data were applied to develop autoregressive models to evaluate the predictability of the series and to make year-ahead predictions of catch. Applying the partial autocorrelations listed in Table 1 gives the forecasting equation for Atlantic menhaden catch as

$$\begin{aligned}\hat{C}_t = & 1.06C_{t-1} + 0.24C_{t-2} + 0.08C_{t-3} \\ & - 0.31C_{t-4} + 0.07C_{t-5} \\ & + 0.43C_{t-6} + 0.29C_{t-7}\end{aligned}\quad (7)$$

where \hat{C}_t is the year-ahead catch forecast. The

forecasting equation for Atlantic menhaden CPUE is

$$\hat{U}_t = 1.13U_{t-1} + 0.20U_{t-2} - 0.31U_{t-3} \quad (8)$$

where \hat{U}_t is the year-ahead CPUE forecast. The forecasting equation for gulf menhaden catch is

$$\begin{aligned} \hat{C}_t = & 1.32C_{t-1} + 0.11C_{t-2} - 0.05C_{t-3} \\ & + 0.09C_{t-4} - 0.01C_{t-5} \\ & - 0.17C_{t-6} - 0.36C_{t-7} \end{aligned} \quad (9)$$

The forecasting equation for the CPUE of gulf menhaden was not determined because the correlations were not significant.

The above models are not parsimonious representations of the series. Application of the Box and Jenkins (1976) methods for analysis of "seasonal" time series would give simpler models; however, the series are of short length and higher order differencing would shorten the series available for estimation. The above models are adequate for evaluation of the predictability of the series and for forecasting.

The 1-year-ahead forecasts of catch reflect the observed values well for both the Atlantic menhaden and gulf menhaden fisheries (Figs. 1A and 2A). The differenced gulf menhaden catch series is somewhat more predictable than the differenced Atlantic menhaden series (Table 1), but only about 17% of the total variation of the differenced series for the Atlantic menhaden and 22% of the differenced series for the gulf menhaden are predicted by the autoregressive models—the proportion of the variation in abundance described by delayed density dependence.

The correlations between the observed catches and the catches predicted using the integrated models are high, with $R^2 = 0.79$ for the Atlantic menhaden and $R^2 = 0.67$ for the gulf menhaden. The difference in predictability between the differenced series and the undifferenced data results from trends in the catch data. Both trends and a small scatter about the trends increase R^2 . In the differenced series, the trends have been removed. In the short run (e.g., for 1-year-ahead forecasting), the higher R^2 values are realistic. Both the source of the trends and when the trends will change direction are unknown so that the lower R^2 values are more realistic for predictions several years ahead.

CPUE and differenced CPUE are somewhat less predictable than catch. Only about 14% of the variation of the differenced CPUE series is predicted (Table 1). For the Atlantic menhaden

Table 1. Coefficients of determination (R^2) for autoregressive models applied to menhaden catch and CPUE.

| Lag | Coefficient of determination | | |
|-----|------------------------------|----------------|---------------|
| | Gulf catch | Atlantic catch | Atlantic CPUE |
| 1 | 0.05 | 0.01 | 0.04 |
| 2 | 0.06 | 0.02 | 0.05 |
| 3 | 0.08 | 0.04 | 0.14 |
| 4 | 0.09 | 0.06 | |
| 5 | 0.10 | 0.09 | |
| 6 | 0.22 | 0.17 | |
| 7 | 0.24 | 0.17 | |

CPUE, $R^2 = 0.49$; $R^2 = 0.58$ for the gulf menhaden. The trends in CPUE are less marked than the trends in catch and the trend in the gulf menhaden CPUE is more noticeable than the trend in the Atlantic menhaden CPUE. These differences in trends account for much of the difference in the coefficients of determination.

Confidence intervals were calculated for year-ahead predictions of the catch (Figs. 1A and 2A). The 95% confidence interval for a 6th order autoregressive process for the differenced series for the Atlantic menhaden gives a bound on the year-ahead predictions of catch of $\hat{C}_t \pm 144$; for the gulf menhaden, the bound is $\hat{C}_t \pm 179$. The widths of the confidence intervals indicate there is considerable uncertainty in predicting catch 1 year ahead (Figs. 1A and 2A). Although the confidence intervals are wide, they give a useful measure of the likely range of catch and would be useful for planning.

REFERENCES

- BANNEROT, S. P., AND C. B. AUSTIN. 1983. Using frequency distributions of catch per unit effort to measure fish-stock abundance. *Transactions of the American Fisheries Society* 112:608–617.
- BOX, G. E. P., AND G. M. JENKINS. 1976. *Time series analysis forecasting and control*. Holden-Day, San Francisco, California, USA.
- DOW, R. L. 1969. Cyclic and geographic trends in sea water temperature and abundance of American lobster. *Science* 164:1060–1063.
- FLOWERS, J. M., AND S. B. SAILA. 1972. An analysis of temperature effects on the inshore lobster fishery. *Journal of the Fisheries Research Board of Canada* 29:1221–1225.
- GULLAND, J. A. 1969. *Manual of methods of fish stock assessment. Part I. Fish population analysis*. FAO Manuals in Fisheries Science, Number 4. United Nations Food and Agricultural Organization, Rome, Italy.

- HUTCHINSON, G. E. 1948. Circular causal systems in ecology. *Annals of the New York Academy of Science* 50:221-246.
- JENSEN, A. L. 1976. Time series analysis and forecasting of Atlantic menhaden catch. *Chesapeake Science* 17:305-307.
- KOO, T. S. Y. 1970. The striped bass fishery in the Atlantic states. *Chesapeake Science* 11:73-93.
- LETT, P. F., A. C. KOHLER, AND D. N. FITZGERALD. 1975. Role of stock biomass and temperature in recruitment of southern Gulf of St. Lawrence Atlantic cod *Gadus morhua*. *Journal of the Fisheries Research Board of Canada* 32:1613-1627.
- MENDELSSOHN, R. 1981. Using Box-Jenkins models to forecast fishery dynamics: identification, estimation, and checking. *Fishery Bulletin* 78:887-896.
- MORAN, P. A. P. 1953. The statistical analysis of the Canadian lynx cycles. I. structure and prediction. *Australian Journal of Zoology* 1:163-173.
- NELSON, C. R. 1976. The interpretation of R^2 in autoregressive moving average time series models. *The American Statistician* 30:175-180.
- NICHOLSON, J. A. 1957. The self adjustment of populations to change. *Cold Spring Harbor Symposia on Quantitative Biology* 22:153-173.
- RICKER, W. E. 1954. Stock and recruitment. *Journal of the Fisheries Research Board of Canada* 11:559-623.
- SAILA, S. B., M. WIGBOUT, AND R. J. LERMIT. 1980. Comparison of some time series models for the analysis of fisheries data. *Journal du Conseil Permanent International pour l'Exploration de la Mer* 39:44-52.
- SKUD, B. E. 1982. Dominance in fishes: the relation between environment and abundance. *Science* 216:144-149.
- STOCKER, M., AND R. HILBORN. 1981. Short-term forecasting in marine fish stocks. *Canadian Journal of Fisheries and Aquatic Science* 38:1247-1254.
- SUTCLIFFE, W. H., JR. 1973. Correlations between seasonal river discharge and local landings of American lobster (*Homarus americanus*) and Atlantic halibut (*Hippoglossus hippoglossus*) in the Gulf of St. Lawrence. *Journal of the Fisheries Research Board of Canada* 30:856-859.
- SUTCLIFFE, W. H., JR., K. DRINKWATER, AND B. S. MUIR. 1977. Correlations of fish catch and environmental factors in the Gulf of Maine. *Journal of the Fisheries Research Board of Canada* 34:19-30.
- VAN WINKLE, W., B. L. KIRK, AND B. W. RUST. 1979. Periodicities in Atlantic coast striped bass (*Morone saxatilis*) commercial fisheries data. *Journal of the Fisheries Research Board of Canada* 36:54-62.
- YULE, G. U. 1927. On a method of investigating periodicities in disturbed series. *Philosophical Transactions* 226:267-298.