IDENTIFYING CLIMATIC FACTORS INFLUENCING COMMERCIAL FISH AND SHELLFISH LANDINGS IN MARYLAND[1]

ROBERT E. ULANOWICZ,[2] MOHAMMED LIAQUAT ALI,[3] ALICE VIVIAN,[4] DONALD R. HEINLE,[5] WILLIAM A. RICHKUS,[6] AND J. KEVIN SUMMERS[6]

ABSTRACT

In five of the seven most important commercial fisheries of Maryland an appreciable portion of the annual variations in catch can be linked with past fluctuations in the physical environment. The harvest figures were compared with appropriate annual characterizations of 40 years of daily environmental records using a variation of the stepwise multiple linear regression technique. The criterion for entry of a term into the regression was how well the given variable improved the prediction of a randomly chosen independent subset of catch figures. The identification of spurious predictor variables becomes less probable under this criterion. The results should help in the organization of further research and management concerning these species and may afford estimates of catches 1 or more years into the future.

Annual population levels of commercially harvested fish and shellfish usually fluctuate widely over the years. Such variation is often attributed to the influence of important environmental variables, such as water temperature, upon spawning success (Sissenwine 1978). Environmental variables may directly affect the mortality rates of prerecruits or indirectly exert influence by altering the abundance of forage or predators. Many other aspects of the ecosystem may also alter population levels (Cushing 1975); however, exact causative mechanisms in most fisheries are seldom known.

Year-to-year fluctuations in the abundance of exploited species will determine in part the magnitude of annual harvest of those species. But the relationship will not be completely deterministic, since landings will also be influenced by socioeconomic factors (e.g., prices and costs as they affect effort) as well as biological factors unrelated to exploitation (Ricker 1978). Despite these many complicating factors, significant correlations between various environmental variables and commercial landings of various species have been found in a number of fisheries. Dow (1977), for example, showed that temperature correlates well with the landings of 24 species of finfish, crustacea, and mollusks off the coast of Maine. Sutcliffe (1972) found freshwater input to St. Margaret's Bay to be a good indicator of fisheries production, possibly because of the stimulation of production caused by the nutrients in the runoff water. However, in neither case were the observed relationships demonstrated to help in predicting harvests, nor were the specific mechanisms responsible for the observed relationships rigorously delineated. In contrast, a regression model of brown shrimp landings off North Carolina, using temperature and salinity as independent variables, was found to be a reasonably accurate predictor of future landings (Hunt et al.[7]). Hunt's model has proven to be a useful management tool for this fishery, helping fishermen to decide how to gear up for the coming season.[8] Multiple linear regression has likewise been employed to explain variations in catch (e.g., Flowers and Saila 1972; Driver 1976). Only

[1]Contribution 1235, Center for Environmental and Estuarine Studies of the University of Maryland.
[2]University of Maryland, Chesapeake Biological Laboratory, Solomons, MD 20688.
[3]University of Maryland, Chesapeake Biological Laboratory, Solomons, Md.; present address: Senior Scientific Officer, Fisheries Campus Chandpur, P.O. Baburhat, Dist Comilla, Bangladesh.
[4]University of Maryland, Chesapeake Biological Laboratory, Solomons, Md.; present address: 517 Flag Harbor Drive, St. Leonard, MD 20685.
[5]University of Maryland, Chesapeake Biological Laboratory, Solomons, Md.; present address: CH2M Hill, 1500 114th Ave. SE., Bellevue, WA 98004.
[6]Martin Marietta Corporation, Environmental Center, Baltimore, MD 21227.

[7]Hunt, J. H., R. J. Carroll, V. Chinchilli, and D. Frankenburg. 1979. Relationship between environmental factors and brown shrimp production in Pamlico Sound, North Carolina. Completion Report for Project 2-315-R, North Carolina Department of Natural Resources, Morehead City, N.C., 37 p.
[8]M. W. Street, Chief, Fisheries Management Section, Division of Marine Fisheries, North Carolina Department of Natural Resources and Community Development, P.O. Box 769, Morehead City, NC 28557, pers. commun. April 1981.

Manuscript accepted January 1982.

in the former instance, however, was there an attempt to validate the regression using independent data.

Thus, the value of correlative or regression models of fisheries is twofold: First, significant correlations can serve to guide research into identifying the causes of annual variation in catch; secondly, if validated, such models may forecast harvest in cases where more detailed deterministic models cannot be developed for lack of information.

The issue of model validation is especially important. Correlational models which best regress to the available data are often not the best predictors (Saila et al. 1980). In an effort to overcome this difficulty we have employed an amended version of stepwise regression analysis which does not rely solely on overall goodness-of-fit, but rather identifies those variables most likely to provide good predictions of data points not used in the actual regression.

To our knowledge multiple correlational models of the relationships between environmental variables and commercial landings in the Maryland portion of Chesapeake Bay have not been attempted in an algorithmic fashion. Because most of the dominant species reproduce in Maryland waters, the influence of environmental variation on harvest of these species may be particularly strong. Thus, we have developed multivariate regressions of the landings of major commercial species, using as predictors those environmental variables considered to have biological significance for the species being examined. Although measures of catch per unit effort (CPUE) would have been preferable as indicators of stock size, adequate effort data were not available. The results provide valuable insight into factors which may contribute to determining the population dynamics of these species and may also prove to be of value in establishing management practices.

SPECIES ADDRESSED AND DATA AVAILABLE

Seven dominant species in Maryland landings were selected for analysis. American oyster, *Crassostrea virginica*; blue crab, *Callinectes sapidus*; soft shell clam, *Mya arenaria*; and striped bass, *Morone saxatilis*, were chosen because they are the four species which yield the greatest dollar value to the Maryland economy. Menhaden, *Brevoortia tyrannus*, and alewife, *Alosa pseudoharengus* and *A. aestivalis* arbitrarily combined, were selected because they have been dominant in number of pounds harvested. The bluefish, *Pomatomus saltatrix*, was included because its harvest has increased dramatically in recent years, and there was interest in determining if this increase might be related to environmental variation.

A 33-yr record of annual catch data for 24 commercial species was available from records maintained by the Chesapeake Biological Laboratory and the NOAA Current Fisheries Statistics series. These records report the total Maryland landings (Chesapeake Bay and Atlantic Ocean) for each year. The Chesapeake Bay portion of the harvest heavily dominates the catches of the chosen species (85% or more of total). Because of the difficulty in obtaining sufficient information to separate Bay catch from the State total, the total was assumed to be representative of Chesapeake Bay.

ANNUAL CHARACTERIZATIONS OF ENVIRONMENTAL DATA

The environmental variables for which long-term records exist are water temperature, air temperature, salinity, and precipitation. All four have potential relevance to the levels of commercial harvest. Cross correlative relationships among these variables would be accounted for in the stepwise multivariate regression procedure employed in the analysis (discussed below). Daily recordings of these variables exist for a period exceeding 40 yr as taken from the Chesapeake Biological Laboratory pier at the mouth of the Patuxent estuary in Solomons, Md. Because this location is central to the Maryland portion of Chesapeake Bay, these data were assumed to be characteristics of conditions in the bay as a whole. Gaps in air temperature and precipitation from 1960 onwards were filled by data taken at the nearby Patuxent River Naval Air Test Center in Lexington Park, Md.

While catch figures represented the total landings for a season, environmental data existed with much finer temporal resolution. Our goal was to pair each annual catch figure with a value of an environmental property which might be representative of the effect that variable had on the stock during the year the daily readings were accumulated. One straightforward way of characterizing a year is to calculate the annual average of the variable in question. The stock, however, may be more sensitive to shorter term deviations from this average. In an effort to quantify these deviations we devised four different ways of treating each of the original four time series to yield 26 annual series of environmental data.

The first of these methods, calculating the annual average, has already been mentioned. But the annual mean conveys little information on the cumulative amount of stress or benefit experienced by the populations because of the extreme high or low values of environmental variables. To portray the cumulative effects of these deviations, we defined variables analogous to the degree-days of agricultural science. Here the effect of a variable is assumed to be manifested only when its value goes beyond a certain "bias-level." If, for example, the organism is assumed to be cold stressed when the water temperature falls below 4°C, then 3 successive days of 1°C water temperature will contribute 9 degree-days towards the index of cold stress.

For each of the four variables recorded, a high and a low bias level were chosen so that when conditions exceeded these bounds at Solomons, we estimated that there were significant regions throughout the Maryland section of the Chesapeake Bay where fish and shellfish were probably stressed (or benefited) by the large excursions from the norm. These bias levels are shown in Table 1.

Of course, the fishery might be responding to individual episodes of stress, rather than the yearly cumulative value. We, therefore, elected to measure the lengths of the longest episodes during a year that a variable was beyond the bias values. These episodes were intermediate time-scale phenomena (on the order of 1 to several weeks), and we wished to avoid contamination from high frequency events. For example, salinity may have remained above 16.2 ppt for all of a 28 d period, save on the 15th day when it dropped to 16.1 ppt. To characterize the episode as 14 d in duration would clearly be erroneous. To avoid such contamination we chose a "gap-interval" for each variable ranging from 3 to 5 d. If the variable went beyond the bias level for a duration not exceeding the gap interval, the episode was not terminated, although the days on which the lapse occurred were not tallied in the episode length. Thus, the episode of high salinities mentioned above would be counted as 27 d.

TABLE 1.—Parameters used in calculating cumulative variables and episodes.

Variable	High bias	Low bias
Salinity	16.2 ‰	10.5 ‰
Water temperature	26.5°C	4°C
Air temperature	30°C	0°C
Precipitation	3 cm/d[1]	0 cm/d

[1]This value becomes 0.01 cm/d in calculating rain episodes, i.e., any day it rains is counted.

Finally, the possibility remains that the stocks might be acutely affected by short-term, intense stresses. We felt this eventuality would be reflected in the annual extremes of each variable.

These four operations, when applied to the four daily time series, yielded 26 annual time series of interest. (Cumulative and extreme low precipitations are uniformly zero by definition, and provide no information.) These series constituted the possible "predictor vectors" from which those yielding the best multiple regressions would be chosen. The values for the 26 variables calculated for the years 1938-76 are listed in Ulanowicz, Caplins, and Dunnington (1980).

REGRESSION METHODOLOGY

In most fish stocks, year class size is considered to be established by the juvenile stage (Cushing 1975). For example, oyster spat set (analogous to juvenile stages of finfish) is a good indicator of spawning success (Galtsoff 1964). Thus, recruitment (and subsequently harvest) is often correlated to those conditions in the past which helped determine the level of juvenile success. In populations where all individuals in a year class are recruited into the fishery at the same age, and annual landings consist primarily of a single year class, a significant correlation might be obtained when the environmental variable in question was lagged against landings by the number of years equivalent to the age at recruitment.

For most species harvested in Maryland, recruitment is not simultaneous for all members of a year class; landings in 1 yr may consist of members of several or many year classes. Thus, environmental characteristics important to establishing year class strength may be partially correlated with landings recorded over several years, and vice versa. In order to account for such extended partial recruitment, stepwise regressions were employed, allowing the contribution of a given environmental variable to be assessed by successively lagging that independent variable by annual increments so as to encompass the lifespan of most of the stock being fished, i.e., harvests are regressed against conditions during the same year, 1 yr ago, 2 yr ago, etc.

In the case of species which do not spawn in Maryland and where environmental conditions in the Chesapeake Bay would not influence year class size (i.e., menhaden and bluefish), any significant correlations arising would either be the result of how the Chesapeake Bay environmental conditions influence the availability of the species to Maryland fishermen, or how its conditions might be correlated with critical conditions at the remote spawning site. Oysters and striped bass, being the longer lived of the species of interest, were regressed against conditions as long ago as 9 yr in the past. Conditions affecting the remaining species were investigated over the past 5 yr.

As mentioned in the introduction, we wished to limit our attention to those variables which are most likely to be good predictors of future harvests. In conventional stepwise regression, that variable which increases the goodness-of-fit by the greatest amount (usually measured by R^2, the percentage of the variance explained by the model) is included as the next variable in the regression equation. We chose instead to enter that variable which improved the model prediction of independent data points by the greatest amount.

To implement this alternate criterion we randomly chose 25% of our data to be reserved for testing. At each step in the regression all of the remaining variables were entered in turn into a least squares multiple regression using the remaining 75% of the data (employing subroutine GLH from the Univac[9] STAT-PAK library). The coefficients derived for each entry were then used to see how well they would predict the test values of the dependent variable. That variable whose inclusion generated the greatest improvement in fitting the test data (as measured by the sum of the squares of the deviations) was entered into the prediction equation.

Ivakhnenko et al. (1979) suggested that one should continue to include terms until the prediction can no longer be improved. It became apparent during the first few runs, however, that with six or eight degrees of freedom in the test data, statistically insignificant improvements in predicting the independent data were occurring. Accordingly, no variable was added to the prediction equation when its F-to-enter statistic (calculated on the fit to the test data) dropped below 3.5. This somewhat low level of confidence (a little below 90%) was chosen so as not to exclude potential predictors early in the screening process. It should also be pointed out that because of the small number of points in the test data, a relatively large percentage of the error in the test data must be explained to meet this F-to-enter criterion (40-60% in our trials).

By separating test and regression data in a random manner, it was always possible that by chance the set of test data chosen for any single run was unduly influenced by high or low production years. Such bias in the test data could result in a predictor accurate only under particular circumstances. Hence, it was necessary to run several (and in the later stages of the screening process, many) trials with different randomly chosen sets of test data. Presumably, the predominance of any single sequence of predictor vectors among the various trials would be an indication that the associated model might be a robust tool for forecasting. Once the functional form of the best predictor has been chosen, the parameters of this equation are redetermined using the full data set.

The sequence of searches outlined above should provide a necessary (although not sufficient) test for prediction formulae.

RESULTS AND DISCUSSION

To facilitate easy recognition of the environmental variables in the regression equations that are to follow, we adopt a two-letter, one-digit code to designate each of the 260 possible predictor vectors. The first letter will be either A, C, E, or X according to whether the processed variable represented an annual average, cumulative deviation, episode, or extremum, respectively. The second letter will designate air temperature, water temperature, daily precipitation, or salinity by T, W, P, or S, respectively. When it is necessary to distinguish between high or low deviations of these variables, the low values will be designated by writing the second letter in the lower case. Finally, the digit will designate the number of years lag behind the harvest figures. As examples, Cs3 would indicate cumulative low salinity 3 yr in the past, whereas EW2 would denote the longest episode of high water temperatures 2 yr ago.

After the field of predictor variables for each fishery had been narrowed to five or fewer, 1,000

[9]Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

Monte Carlo trials were run for each species using different random combinations of test data. In five of the seven species considered, pairs of two variables were identified frequently enough to warrant their being cited as potential predictor formulae in Table 2. In one case (blue crab) no variables appeared often enough in the trials to warrant reporting a predictor equation. By contrast several sequences of oyster predictors appeared often, but no sequence predominated in the trials. About five separate sequences appeared with almost equal frequency. Hence, no formula for oysters is cited.

In the clam regression, Cw1 appeared as the primary predictor in over 50% of the trials. In roughly 20% of these instances CS2 was included as secondary predictor. No predictor was chosen 12% of the time. When the two selected variables were finally regressed against the entire data set, 60% of the total variance was explained, 49% by Cw1 alone. Environmental data were available to assess the predictive value of this equation for the year 1977. As can be seen in Figure 1, this projected value has a large deviation from the recorded measure, but this deviation falls within the range of errors in the hindcast.

Interpretation of this equation in terms of causality is complicated by the absence of effort data. For example, the effects of the rise in number of licensed clammers (from 3 in 1952 to 100 in 1957 to 200 in 1979 [Richkus et al.[10]]) on catch cannot be accounted for, and they may have been substantial. Still, the strong correlation between cumulative low water temperature lagged 1 yr and catch suggests a causal relationship.

In Maryland, soft shell clams spawn in spring and fall (Pfitzenmeyer 1962). However, the spring set each year is almost totally eradicated because of predation by benthic feeding fish and crabs which migrate onto Maryland clam grounds each spring and leave each fall (Holland et al.[11]). Factors influencing the strength of the fall set (which occurs from October through December) and the ensuing survival of juveniles have not been identified. It appears that these factors are the ones most likely to have the greatest effect on the magnitude of commercial clam landings. Since Maryland is near the southern

TABLE 2.—Potential predictors of landings (in metric tons) of designated species; see text for code to predictor variables.

Species	Regression	Multiple R^2	F	df
Soft clam, *Mya arenaria*	$H_s = 372.4 + 13.56Cw1 + 3.765CS2$	0.60	11.1	22
Menhaden, *Brevoortia tyrannus*	$H_m = 888.4 + 45.99Ep5 - 9.126CT4$	0.53	11.1	30
Bluefish, *Pomatomus saltatrix*	$H_b = -48.20 + 1.948Ep5 + 0.1693Cs2$	0.60	14.8	29
Alewives, *Alosa aestivalis* and *A. pseudoharengus*	$H_a = 3{,}344 - 24.19Ep3 + 93.80Xt2$	0.51	10.5	30
Striped bass, *Morone saxatilis*	$H_r = 7{,}414 - 446.5AT3 + 2.435Ct1$	0.45	8.3	30

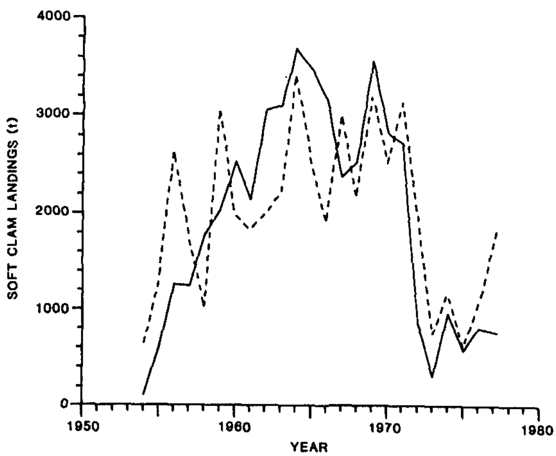

FIGURE 1.—Maryland soft clam landings in metric tons from 1952 to 1977 (solid line) and landings predicted using the regression model (dotted line) (Table 2). (Landings for 1977 did not enter into the derivation of the model.) Environmental factors were a cumulative low deviation in water temperature (with a 1-yr lag behind the harvest figure), and a cumulative high deviation in salinity (with a 2-yr time lag).

[10]Richkus, W. A., J. K. Summers, T. T. Polgar, and A. F. Holland. 1980. A review and evaluation of fisheries stock management models. Martin Marietta Laboratories, Baltimore, Md., 177 p.

[11]Holland, A. F., N. K. Mountford, M. Hiegel, D. Cargo, and J. A. Mihursky. 1979. Results of benthic studies at Calvert Cliffs. Final Report to Maryland Power Plant Siting Program. Ref. No. PPSP-MP-28, 229 p. Martin Marietta Laboratories, Baltimore, Md.

boundary of the geographical range of soft-shell clams (Manning[12]), cold water temperatures may, in some unexplained way, enhance the survival of a previous year's set.

Manning and Dunnington (1956) showed that Maryland clams grow at a rapid rate, achieving legal size (2 in, 5.1 cm) at an age of 16 to 22 mo. Hence clams spawned in the fall of one year would enter the commercial fishery during the spring 2 yr later. Extreme low water temperatures generally occur in January or February of each year and during some years coincide with periods of high salinities. Thus, both variables in the regression model could be exerting an effect on juvenile clams from set to the age of 6-7 mo, when they are approximately 0.5 cm (0.2 in) in size. Low water temperature may delay movement of predators into Maryland waters, permitting juvenile clams to grow to a less vulnerable size. High salinities during the juvenile life stages could also favor growth and rapid maturation of clams.

The remaining four regressions are composed of variables which are less readily explained. The two terms entering the menhaden regression have 4 and 5 yr lags (Table 2). But menhaden which make up Maryland landings are of ages 2 and 3, with almost no 4-yr-old fish taken (Merriner[13]). Thus, any causal mechanism suggested by the regression would have to be a second generation response, remembering that menhaden do not become sexually mature until ages 3 or 4. Ep5 appeared as the most important predictor in 37% of the trials with CT4 following as a secondary predictor in 23% of those cases. Menhaden catch is depicted in Figure 2.

Juvenile bluefish use Chesapeake Bay as a nursery area and there is the possibility that a distinct Chesapeake Bay stock of bluefish exists (Kendall and Walford 1979). Bluefish, being a marine species, are generally not found in low salinity waters, and their distributions can be well defined by salinity patterns (Lippson et al. 1980). Thus, the precipitation variable entering the regression (Table 2) may reflect diminished nursery habitat caused by high precipitation, resulting in a decline of harvestable fish in future years. However, age composition of Maryland bluefish catch is unknown, and the particular lags in the regression are not readily explained. Bluefish harvests are illustrated in Figure 3. Ep5 appears as the primary predictor in 56% of the trials run with Cs2 following in 23% of those instances. It is noteworthy that the same variable (Ep5) appears as the most useful predictor of both menhaden and bluefish. Both species are coastal spawners, and it is entirely possible that

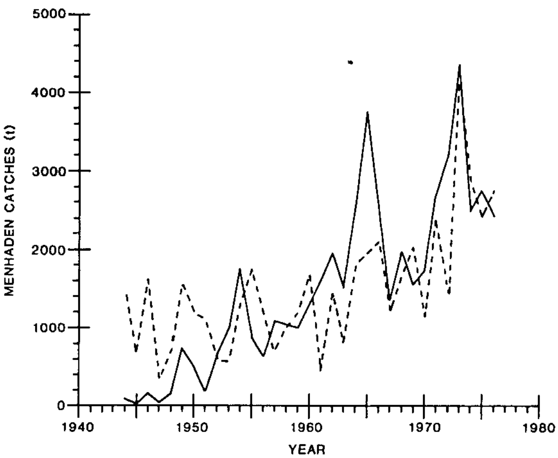

FIGURE 2.—Actual (solid line) and predicted (dotted line) catches in metric tons of menhaden, 1946-76, based on the regression model in Table 2. Environmental factors were an episode of low daily precipitation (with a 5-yr time lag behind the harvest figure), and cumulative high deviation in air temperature (4-yr time lag) which had a negative effect.

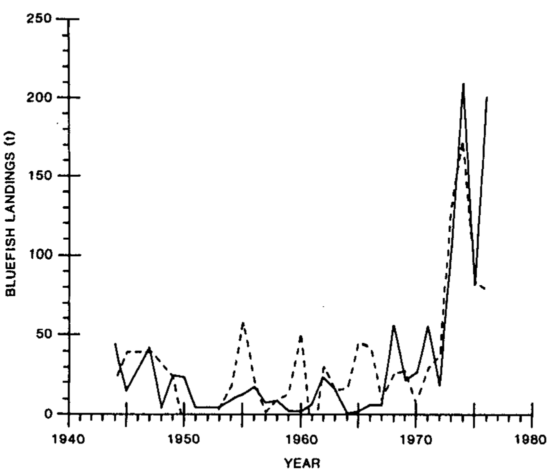

FIGURE 3.—Predicted (dotted line) and recorded (solid line) weights of bluefish landings in metric tons, 1947-76, based on the regression model in Table 2. An episode of low daily precipitation (5-yr time lag), and cumulative low deviation in salinity (2-yr time lag) were the environmental factors.

[12]Manning, J. H. 1957. The Maryland soft-shell clam industry. Study Report 2, 25 p. Maryland Department of Research and Education, Solomons, Md.

[13]J. V. Merriner, Chief, Division of Fisheries, Southeast Fisheries Center Beaufort Laboratory, National Marine Fisheries Service, NOAA, Beaufort, NC 28516, pers. commun. September 1980.

the same causal mechanism is affecting the catches of both fishes.

Both species included in landings as alewives, *A. pseudoharengus* and *A. aestivalis*, are anadromous, tributary spawners in Maryland (Hildebrand and Schroeder 1927). Thus, poor spawning success could readily be related to low freshwater runoff caused by low precipitation. However, the age of first spawning of these species is from 3 to 5 yr, with the majority spawning at 4 or 5 (Davis et al.[14]). The regression (graphed in Fig. 4) suggests the possibility that recruitment in Maryland occurs at a younger age, but data on the age distribution of the catch are unavailable to confirm or refute this suggestion. The other variable entering the alewife regression (Table 2) is lagged by 2 yr. Since fish taken in a given year would have been present in the Atlantic Ocean 2 yr before being harvested, this correlation is difficult to explain in a causal manner. Ep3 is the major predictor in 50% of the trials and is followed by Xt2 in 20% of those cases.

The least significant of all the predictors cited is the one for striped bass (see Fig. 5). Both terms show a favorable correlation with cold air temperatures over a season. Cold seasons are conducive to greater amounts of ice formation along river edges. The scouring from ice floes contributes high quality detritus to the riverine system to supplement the food source for zooplankton, in turn providing the larvae with abundant food (Heinle et al. 1976). Boynton et al.[15] have previously remarked that year class success correlates jointly with cold winters and high runoff. The chosen variables did not dominate the trials heavily. AT3 was the major predictor in only 33% of the trials, one-third of which were accompanied by the variable Ct1. Xs4 appeared as the major predictor almost as often (25% of the time there was no major predictor), but did not result in a significant correlation with the full data.

An examination of the power spectra of the errors of the five models (using subroutine POWDEN in Univac STAT-PAK) revealed no appreciable differences from the spectral pattern of random noise.

[14]Davis, J., J. V. Merriner, W. J. Hoagman, R. H. St. Pierre, and W. L. Wilson. 1971. Annual Progress Report, Anadromous Fish Project. Proj. No. Va. AFC7-1, 106 p. Virginia Institute of Marine Science, Gloucester Point, Va.

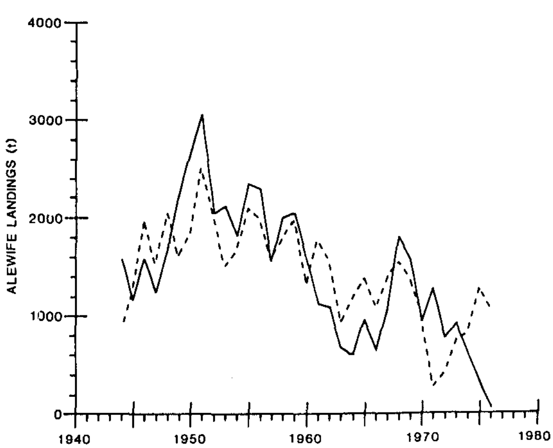

FIGURE 4.—Annual landings (solid line) in metric tons of alewife, 1944-76, as compared with predicted values (dotted line), based on the regression model in Table 2. Environmental factors were an episode of low daily precipitation (3-yr time lag) (negative effect) and an extremum of low air temperature (2-yr time lag).

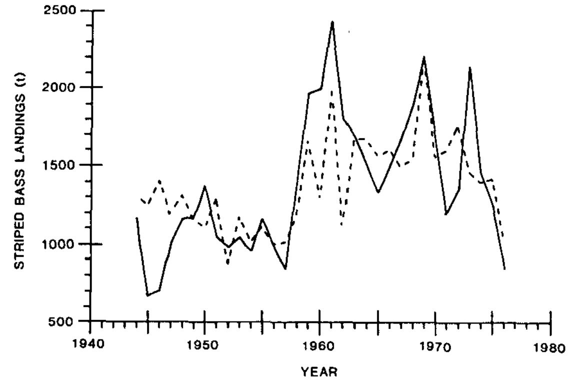

FIGURE 5.—Predicted (dotted line) and tabulated (solid line) landings in metric tons of striped bass from 1944 to 1976, based on the regression model in Table 2. Environmental factors were high annual average air temperature (3-yr time lag) (negative effect) and cumulative low air temperature (1-yr time lag).

CONCLUDING REMARKS

Perhaps the most significant observations to be made from this exercise involve the comparison of the results reported herein with those reported previously from a conventional search for

[15]Boynton, W. R., E. M. Setzler, K. V. Wood, H. H. Zion, M. Homer, and J. A. Mihursky. 1976. Potomac River fisheries program ichthyoplankton and juvenile investigations. Ref. No. 77-169CBL, 328 p. Center for Environmental and Estuarine Studies, Solomons, Md.

predictors using the same data (Ulanowicz, Ali, and Richkus 1980). The major predictors cited in the previous work were either identical to, or qualitatively similar to, the initial variables selected by the more usual analysis. In that earlier work up to seven terms appeared in one regression equation (F-to-enter criterion of 4.0), and R^2 values ranged as high as 0.86 with four variables. Despite having dropped the F-to-enter criterion below the 90% confidence level, the joint criterion that the variables chosen also be reasonably good "internal predictors" appears to have resulted in a more stringent combined test for selecting variables. Fewer spurious predictors are likely to appear using the new criteria. Although the regression with the full set of data will not be as tight as might otherwise be possible, there is less likelihood that predictions on independent data will be wildly in error. In the words of Ivakhnenko et al. (1979), the "fan of predictions" has been narrowed.

When the number of possible predictor vectors is large compared with the number of observations (as it is in this study), there is concern that multiple regression R^2 values can be inflated (Rencher and Pun 1980). Fortunately, the method described herein does not rely on R^2 values alone. Before a variable is chosen for further consideration, it must explain a significant fraction of the variance in several randomly assembled groups of test data. To see how well this might screen against including spurious variables, the search procedure for the menhaden predictor was rerun with the yearly observations randomly scrambled. Out of 28 possible trials with the original data, at least one variable was added in exactly half the trials (with an average F-to-enter of 9). In all but 2 of those 14 successes the first variable entered was identical (Ep). By contrast, only 5 successful trials were recorded with the scrambled data (average F-to-enter was 5), although one variable did appear in 3 of those successful trials. Nonetheless, there is an evident decrease in the frequency and number of variables with successful F-to-enter ratios in the trials with scrambled data. The only species studied giving results nearly as poor as the scrambled data was blue crab, and those findings were disregarded.

Unfortunately, the results of the present analysis must still be viewed with caution. Although the possibility of identifying a spurious correlation as a predictor has been decreased, it cannot be totally eliminated. The fact that substantial portions of the variability in landings of all the species considered can be explained by a few environmental variables suggests the important role which environmental conditions play in determining stock size. However, our inability to interpret many of these relatinships in a causal manner reflects both a lack of knowledge of mechanisms influencing fish population dynamics as well as an unfamiliarity with the auto and cross correlative relationships between the variables introduced into the regression process. (Because the procedure employed was stepwise, true causal variables may have been displaced in the regressions by spurious variables which by chance were closely correlated. No detailed analysis of the independent variable data sets was performed to address this issue. More analyses would be required to fully account for this possibility.)

Despite these limitations, the analyses appear to have been fruitful, particularly in the case of the soft clam. As for the other species, the value of the models will be determined when sufficient data are available to assess their predictive value for future landings. Only then can it be ascertained whether any chosen predictor was spurious or reflected some unknown causal relationship between environmental variation and stock dynamics. Meanwhile, the terms appearing in the regressions may engender new research projects into the mechanisms determining the sizes of these important fisheries stocks.

ACKNOWLEDGMENTS

This work is a result of research sponsored in part by the NOAA Office of Sea Grant, Department of Commerce, under grant no. 04-7-158-44016. Mohammed Liaquat Ali's internship at the Chesapeake Biological Laboratory was jointly sponsored by the Directorate of Fisheries, Government of Bangladesh, and the World Bank. Alice Vivian was supported by the Environmental Protection Agency's Chesapeake Bay Program, Eutrophication Project, grant no. Sub R806189010. William A. Richkus and J. Kevin Summers were supported under a contract to the Martin Marietta Environmental Center from the Coastal Resources Division, Maryland Tidewater Administration. Edward Houde and Martin Wiley provided comments helpful in revising the manuscript. The Computer Science Center of the University of Maryland donated some of the computer time used in this project. The authors would like to thank the journal referees for suggesting the separation of data into regression and verification subsets and the scrambling of the criterion vector to check for spurious predictors.

LITERATURE CITED

CUSHING, D. H.
1975. Marine ecology and fisheries. Camb. Univ. Press, Camb., 278 p.

DOW, R. L.
1977. Effects of climatic cycles on the relative abundance and availability of commercial marine and estuarine species. J. Cons. Cons. Int. Explor. Mer 37:274-280.

DRIVER, P. A.
1976. Prediction of fluctuations in the landings of brown shrimp (*Crangon crangon*) in the Lancashire and Western Sea Fisheries District. Estuarine Coastal Mar. Sci. 4:567-574.

FLOWERS, J. M., AND S. B. SAILA.
1972. An analysis of temperature effects on the inshore lobster fishery. J. Fish. Res. Board Can. 29:1221-1225.

GALTSOFF, P. S.
1964. The American oyster, *Crassostrea virginica* Gmelin. U.S. Fish Wildl. Serv., Fish. Bull. 64:1-480.

HEINLE, D. R., D. A. FLEMER, AND J. F. USTACH.
1976. Contributions of tidal marshlands to mid-Atlantic estuarine food chains. *In* M. L. Wiley (editor), Estuarine processes. Vol. II. Circulation, sediments, and transfer of material in the estuary, p. 309-320. Acad. Press, N.Y.

HILDEBRAND, S. F., AND W. C. SCHROEDER.
1927. Fishes of Chesapeake Bay. Bull. U.S. Bur. Fish. 43(1), 366 p.

IVAKHNENKO, A. G., G. I. KROTOV, AND V. N. VISOTSKY.
1979. Identification of the mathematical model of a complex system by the self-organization method. *In* E. A. Halfon (editor), Theoretical systems ecology, p. 325-352. Acad. Press, N.Y.

KENDALL, A. W., AND L. A. WALFORD.
1979. Sources and distributions of bluefish, *Pomatomus saltatrix*, larvae and juveniles off the east coast of the United States. Fish. Bull., U.S. 77:213-227.

LIPPSON, A. J., M. S. HAIRE, A. F. HOLLAND, F. JACOBS, J. JENSEN, R. L. MORAN-JOHNSON, T. T. POLGAR, AND W. A. RICHKUS.
1980. Environmental atlas of the Potomac estuary. Johns Hopkins Univ. Press, Balt., Md., 279 p.

MANNING, J. H., AND E. A. DUNNINGTON.
1956. The Maryland soft shell clam fishery: A preliminary investigation report. Proc. Natl. Shellfish. Assoc. 46:100-110.

PFITZENMEYER, H. T.
1962. Periods of spawning and setting of the soft-shell clam, *Mya arenaria*, at Solomons, Maryland. Chesapeake Sci. 3:114-120.

RENCHER, A. C., AND F. C. PUN.
1980. Inflation of R^2 in best subset regression. Technometrics 22:49-53.

RICKER, W. E.
1978. Computation and interpretation of biological statistics of fish populations. Dep. Environ. Fish. Mar. Serv., Ottawa, Bull. 191.

SAILA, S. B., M. WIGBOUT, AND R. J. LERMIT.
1980. Comparison of some time series models for the analysis of fisheries data. J. Cons. Cons. Int. Explor. Mer 39:44-52.

SISSENWINE, M. P.
1978. Is MSY an adequate foundation for optimum yield? Fisheries (Bethesda) 3(6):22-42.

SUTCLIFFE, W. H., JR.
1972. Some relations of land drainage, nutrients, particulate material, and fish catch in two eastern Canadian bays. J. Fish. Res. Board Can. 29:357-362.

ULANOWICZ, R. E., M. L. ALI, AND W. A. RICHKUS.
1980. Assessing harvests of pelagic and invertebrate fisheries of northern Chesapeake Bay in terms of environmental variations. Int. Counc. Explor. Sea Code C. M. 1980/H:43, 8 p.

ULANOWICZ, R. E., W. C. CAPLINS, AND E. A. DUNNINGTON.
1980. The forecasting of oyster harvest in central Chesapeake Bay. Estuarine Coastal Mar. Sci. 11:101-106.