PAPER

International Council for the Exploration of the Sea

ICES C.M. 1991/D:26
Statistics Committee
Ref.No. [UMCEES]CBL 91-150

FISHMAP - AN EXPERT DECISION SUPPORT SYSTEM FOR EFFICIENT FISH SAMPLING AND STOCK ASSESSMENT

by

Brian J. Rothschild, Jerald S. Ault and Gerard T. DiNardo

University of Maryland
Center for Environmental and Estuarine Studies
Chesapeake Biological Laboratory
Solomons, MD 20688 USA

Abstract

We describe development of an expert decision-support sampling system which provides fishery managers a tool for identification of efficient and cost-effective sampling designs. The system, FISHMAP, links techniques in sampling, estimation and optimization through use of expert systems technology. FISHMAP's modular design provides: (1) strategy to determine the number and real-time dimensions of strata, (2) definition of spatial sampling characteristics (i.e where to sample) for individual stock segments (i.e. size classes), (3) methods to determine necessary sample sizes within strata, and (4) methods to assess data collected in real-time, and if necessary suggest design modifications. When applied to stocks of white perch in northern Chesapeake Bay we observed significant gains in precision due to real-time stratification.

Introduction

Effective fishery resource management strategy relies in part on precise measures of population dynamics and abundance indices, and functional relationship(s) which exist between these indices and fishing effort. Fishery-independent trawl surveys were recognized for their value to fisheries management over 30 years ago (e.g. Grosslein, 1969). The number of fish caught per tow reflects an index of species' abundance (Clark, 1981; Halliday and

Koeller, 1981; Pitt et al., 1981), and attribute measurements of the catch are used to generate population dynamics parameters. The impact of decision-making advice depends on the quality of parameter estimates.

Because fish are usually distributed heterogeneously, stratified random sampling (STRS) is often used as the basis for survey design (Doubleday and Rivard, 1981). STRS exploits the heterogeneity by lumping items of similar variance, usually resulting in more precise (lower variance) estimates than simple random sampling (SRS). Further, STRS provides the synoptic spatial population coverage which is not ensured using SRS. However, despite use of stratification, time-series of abundance for many species contain high variance making it difficult to detect statistically real differences between consecutive years (Survey Working Group, Northeast Fisheries Center, 1988). These problems are thought to result from 1) a combination of incorrectly specified strata and 2) inappropriate allocation of sampling effort to each stratum (Gavaris and Smith, 1987). Including prior information on catch rates and sources of variability to set stratum boundaries and effort allocations would result in more efficient estimates (Francis, 1984). Physical variables related to local variability of fish abundance can be incorporated as covariates into statistical models of catch-per-tow and these relationships used to formulate a more efficient stratification scheme. The precision of estimates maximized when factors affecting availability are observed in real-time. However, the identification of factors affecting variability, and their incorporation into a real-time sampling scheme, is complex requiring a systems approach.

In this paper we describe development of an expert fishery-independent decision support sampling system, FISHMAP (Fishery Independent Sampling and Habitat MAPping), which provides fishery managers with a computer-based tool for identifying efficient and cost effective sampling designs (e.g. number of strata), as well as algorithms to facilitate real-time cost-effective data collection. Using expert systems technology, our system links techniques in sampling, estimation, and optimization. FISHMAP supports a generalized framework which can be applied to any species and sampling gear configuration. We configured the system for trawling and applied it to Chesapeake Bay white perch (<u>Morone americana</u>) stocks.

Methods

The CONCEPTUAL MODEL

Many applications of artificial intelligence and expert systems exist in the natural resource area. Recently, there have been applications in crop management (Lemmon, 1986), water resources management (Reboh et al., 1982; Palmer and Holmes, 1988), fishery resource management (Ryan and Smith, 1985; Stagg, 1990), and sampling in contaminated sites (Van der Gaast and de Jonge, 1988). An expert system is an intelligent computer program employing knowledge and inference procedures to solve specific problems and assist decision-making (Turban, 1988). Expert systems are designed to imitate the reasoning process of experts.

We developed an expert decision support system to facilitate and improve estimates derived from fishery-independent finfish resource surveys. FISHMAP is composed of three modules 1) database management, 2) sampling design, and 3) field sampling. Each module is hierarchically structured (Figure 1). The database management module contains historical trawling data and reside in three integrated databases (i) trawl characteristics, (ii) catch, and (iii) population dynamics. Information derived from these databases are used as inputs into the sampling design and field sampling modules.

The sampling design module is intended for use in the office prior to field sampling. It has two components; strategic information and tactical models. The strategic information identifies where to sample. The tactical models determine stratification variable, number of strata, stratum boundaries and stratum sample sizes.

The field sampling module is intended for field use. It is comprised of two components, real-time sampling and population estimation. The real-time sampling models update design characteristics identified during consultation with the sample design module. Population estimation models 1) randomly chooses stratum sampling sites, 2) allows for direct input of catch data, and 3) estimates mean catch-per-unit effort (CPUE) and associated variance for each species.

PROTOTYPE COMPUTER MODEL

We have developed an operational prototype microcomputer-based FISHMAP system. Module integration is facilitated using the rule-based expert system shell LEVEL5 (Information Builders, Inc., 1988). Statistical operations within modules and general system linkage is supported through use of generalized external programs (GEP) written in the "C" programming language. FISHMAP supports all components of the sampling design and field sampling modules. At present, it does not support the database management module. Data required by the sample design and field sampling modules reside as data files. Operations within the sampling design and field sampling modules are shown in Figure 2, and function as follows.

Sample Design Module

The sample design module is initialized through a series of queries and responses relative to target species, locations (general area to sample), size-class and month of sampling. Initialization results in access of strategic information and a system recommendation of sampling area. Sampling area recommendations are determined by system analysis of the spatial distribution of a species. This allows FISHMAP to modify spatial boundaries of sampling, including only those areas occupied by the species. Data to arrive at sample area recommendations is available for viewing in the form of density maps.

Upon accessing the tactical models component a stratification submodel which will be used to determine sample design characteristics is developed. Submodel development uses stepwise regression procedures (Neter et al., 1985). The dependent variable is catch per 0.5 nm trawl tow (CPUE). Independent variables include four physical variates 1) bottom salinity, 2) bottom water temperature, 3) bottom dissolved oxygen, and 4) trawling depth. Plots of CPUE aganist the physical variates were usually parabolic. Therefore, in the modeling process squares of the variables are also included as independent variables. Statistically significant independent variable(s) identify stratification variable(s). Criterion for variate addition to the model is a statistically significant reduction in MSE. Because trawl catch data is highly skewed catch data is log-transformed (ln(CPUE + 1)) and modeled as the polynomial

$$C_j = \beta_o + \beta_n X_n^i + \xi \qquad \forall\, k, \tag{1}$$

where C_j is estimated ln(CPUE + 1) in the *j-th* time period (e.g. month); β_o is a constant; β_n are regression coefficients; X_n are the independent variables; k is a location index; n is a independent variable index; i=1,2 (Neter et al., 1985). To facilitate ease in model interpretation no interactions are considered.

To determine number of strata (L) we used a 10% relative gain stopping rule is used. First data from the stratification submodel is used to determine the reduction in variance due to stratification by 1 to 20 strata (Cochran, 1977). Next, relative gain in precision due to the addition of another stratum is determined (Figure 3). The minimum strata number resulting in at least a 10% relative gain in variance reduction equals L. To determine stratum boundaries we used the cum $\sqrt{f(y)}$ rule of Dalenius and Hodges (1959). This method computes cumulative square-root values of percent total catch frequency for each stratification variable interval. If the cumulated total over the interval H, the approximate stratum boundaries are

$$y_1 = H/L, \quad y_2 = 2H/L, \ldots, y_h = hH/L, \qquad (2)$$

where L is number of strata and h a stratum boundary index (h=1,...,L-1). In other words, the cumulative value of $\sqrt{f(y)}$ is made constant within all strata.

To determine total sample size the following equation from Cochran (1977) is used

$$n = (\sum_i W_i S_i)^2 \ / \ (V_o + N^{-1}(\sum_i W_i S_i^2)), \qquad (3)$$

where i is stratum number (1,2,...,L); N is the number of sampling units in the population, $N = N_1 + N_2 + \ldots + N_L$; N_i is the number of sampling units in stratum i; S_i^2 is sample variance in stratum i; S_i is sample standard deviation in stratum i; V_o is the desired sampling variance and defined as $V_o = (B/t_{\alpha,0.95}) = B^2/4$; ; B is the amount of acceptable error for our estimate; W_i are stratum weights and of the form N_i/N. FISHMAP calculates total sample size for error tolerances of 0.5, 0.4, 0.3, 0.2, 0.1 and 0.05. Stratum weights (W_i) are determined by first developing a regression equation of latitude verses stratification variable metrics from historical data, solving the equation for each stratum boundary value identified by the cum $\sqrt{f(y)}$ procedure, and than enumerating the number of possible sampling sites (N_i) within each stratum. CPUE data is assigned to appropriate strata and CPUE variance metrics calculated. Possible sampling sites within target locations reside as data files within FISHMAP and were identified by placing a 0.5nm by 0.5nm grid over a nautical chart of each sampling location and recording the location coordinates of each grid node in waters greater than 10 feet.

Stratum sample size is determined for each error tolerance level using Neyman allocation (Neyman, 1934) and expressed as

$$n_i = n\{(N_i S_i) \ / \ (\sum_i N_i S_i)\}. \qquad (4)$$

To determine sample size, strata must be capable of (i) supporting sampling and (ii) providing prior metrics of stratum mean CPUE and variance. Two fail-safe procedures reside in the sample size and allocation of sampling effort submodels component which check the adequacy of defined strata to support these requirements. The fail-safe procedures reside as an integrated set of rules and support a conservative approach to sampling when historical trawl data is scant. The first fail-safe procedure checks the adequacy of defined stratum to support sampling. Within FISHMAP minimum stratum sample size is set at three. Inadequacies in stratum size (stratum unable to support minimum sample size) result in recalculation of strata number and characteristics (e.g.

boundary measures). The second fail-safe procedure ensures non-zero estimable stratum metrics of mean CPUE and variance. Rule-based approximations are used to estimate mean CPUE and variance metrics when stratum sample size equals 1. For zero variance strata the following heuristic is applied--stratum sample size is equal to 20% of the total possible sample size (N_i) or 3 which ever is larger. As length of the trawl time series increases these rules can be relaxed.

Stratum sample sizes, for each error tolerance level, are displayed in tabular form. This table provides the user with some reference of tradeoffs in precision associated with sample sizes, and the user must specify a desired acceptable error tolerance level. The user is asked if the estimated total sample size for the specified error level is acceptable. If so a summary table of sample design characteristics is displayed. If the determined total sample size is infeasible the user is queried for total allowable sample size. Stratum sample sizes for the desired acceptable error tolerance level, based on total allowable sample size, is determined and a new summary table displayed. It should be noted that the displayed latitudinal boundaries reflect historical locations of stratification variable bounds and are <u>not</u> real-time locations.

This concludes consultation with the sample design module. Consultation output includes where to sample, stratification variable, number of strata, stratification variable boundary values, total sample size for a specified allowable error tolerance level and stratum sample size. All data generated during consultation is saved in ASCII formatted output files.

Field Sampling Module

Prior to trawling, stratification variable measures and associated location data (longitude and latitude where measurement is taken) are collected in the sampling area and used to formulate the linear regression model

$$Latitude = \beta_o + \beta_1 (Stratification\ Variable). \qquad \textbf{(5)}$$

Real-time location of stratum boundaries (latitude) and stratum area metrics are determined by solving the equation for each stratum boundary measurement identified during consultation with the sample design module, and summing total number of possible sampling sites. Real-time area calculation is pivotal since a change in stratum size affects sampling strategies; in particular, stratum sample size. For example, prior to the identification of real-time stratum boundaries suppose a sample size of 35 was required in stratum 1 (n_1=35) and that this accounted for 70% of all possible sampling sites (N_1 = 50) in the stratum. Because N_i identifies all possible sampling sites in stratum i it also serves as a metric of stratum area. Suppose that after collecting stratification variable data and identifying real-time stratum boundaries, stratum area decreased such that $N_1^* = 30$ (where the superscript * indicates real-time stratum metrics). Because $N_1 > N_1^*$ a decrease in stratum sample size must occur (in this example the original sample size is now larger then new total possible sampling sites). Adjusted stratum sample size is calculated as

$$n_i^* = N_i^* [n_i / N_i], \tag{6}$$

where n_i^* is the adjusted sample size in stratum i, N_i^* is the real-time total number of possible sample sites in stratum i, n_i is the original sample size in stratum i, and N_i is total number of sample sites in stratum i prior to the identification of real-time stratum boundary locations. Because n_i is proportional to variance in stratum i, as well as its size, the ratio (n_i / N_i) in the above equation serves as a variance and areal weighting factor. This procedure assumes that the error structure within a stratum remains the same relative to the other stratum, regardless of changes in stratum size. Adjusted total sample size, n^*, is the sum of all adjusted stratum sample sizes, $\sum_i n_i^*$.

Because of the dynamic nature of stratum boundary locations, adjusted total sample sizes may be greater than original sample size. However, constraints on resources (e.g. money) may deny collection of adjusted total sample size. In many cases, the original sample size is all that can be collected. Procedures in the real-time sampling models component allow for the allocation of the original sample size among the recalculated strata by the equation

$$n_i' = N_i^* (n_i / N_i) (n / n^*), \tag{7}$$

where n_i' is the allocated sample size in stratum i when the total sample size for reallocation equals n. The real-time sampling table is then displayed listing real-time locations of stratum boundaries and stratum sample sizes. Stratum sampling sites are randomly chosen and their location coordinates displayed.

We assume trawling will be conducted at each sampling site. The population estimation models component allows for direct entry of trawl data through a series of queries. At the conclusion of sampling, species-specific tables of mean CPUE and variance estimates are displayed. Stratum means and variances are calculated as

$$\overline{x}_{ij} = \sum_{k} X_{ijk} / n_i, \tag{8}$$

and

$$VAR(\overline{x}_{ij}) = \sum_{k} (x_{ijk} - \overline{x}_{ij})^2 / (n_i - 1), \tag{9}$$

where $\overline{x}_{ij}$ is the mean catch per tow in the i-th stratum for species j, x_{ijk} is the catch in stratum i of species j in the k-th sample, and n_i is the sample size in stratum i. Stratified mean and variance estimates are calculated as

$$\overline{x}_{STRS,j} = W_1\overline{x}_{1j} + W_2\overline{x}_{2j} + . . . + W_i\overline{x}_{ij}, \tag{10}$$

and

$$VAR(\overline{x}_{STRS,j}) = W_1^2 (VAR(\overline{x}_{1j}) / n_1) + W_2^2 (VAR(\overline{x}_{2j}) / n_2) + ... + W_i^2 (VAR(\overline{x}_{ij}) / n_i), \tag{11}$$

where the subscript "STRS" indicates stratified random sampling, $\overline{x}_{st,j}$ is the stratified mean catch per tow for species *j*, $\overline{x}_{ij}$ is the mean catch per tow in stratum i for species j, W_i are weights of the form N_i/N; where N_i is the total number of sampling units in stratum i and $N = N_1 + N_2 + . . . + N_i$.

This concludes consultation with the field sampling module. Consultation output includes real-time locations of stratum boundaries, modified stratum sample sizes, locations of sampling sites, and species specific estimates of mean CPUE and variance. All data generated during consultation is saved in ASCII formatted output files.

FISHMAP APPLICATION SYSTEM

FISHMAP has been applied to stocks of white perch (Morone americana) in both the mainstem of the Chesapeake Bay and Patuxent River. This species was chosen because of its importance to Maryland's commercial and recreational fisheries (DiNardo et al., 1991). White perch are year round residents of Chesapeake Bay waters, exhibiting fairly restricted movement patterns. In many cases, distribution is limited to natal tributaries (Mansueti, 1961; Mulligan, 1987).

Because white perch distribution varies with life stage (Mansueti, 1961; Setzter-Hamilton, 1990), FISHMAP identifies size-class specific sample design characteristics. Four size classes of white perch are identified, small (0-100mm TL), medium (101-180mm TL), large (> 180mm TL), and all sizes combined coinciding with ages 0-1+, 2-5+, >6, and all ages, respectively. Ages were assigned to the three size classes (small, medium and large) based on age-length analyses of Mansueti (1961).

Historical and recent trawl data sets from the Chesapeake Bay, presently housed at and collected by the Chesapeake Biological Laboratory were used to support system development. Data set characteristics are outlined in Table 1. The historical trawl data set comprises six surveys (A-F), spanning the years 1961 to 1975. Approximately 5,000 one-half nautical mile trawls were made during the 15 years. Trawl types varied over the 15 years as did number and location of sample sites; the location of sites depending on survey objective. Sampling frequency was monthly, one sample collected at each site during the early 1960's, increasing to two samples collected at each site (one sample was collected in "shallow" waters and one in "deep" waters) during the mid to late 1960's and 1970's. Throughout the surveys, primarily three bottom trawls were used; a cotton otter trawl with 30 foot sweep, a nylon otter trawl with 40-foot sweep, and a marlon otter trawl with 40-foot sweep.

The recent trawl data set spans the years 1988 and 1989. During 1988, sampling was bimonthly and restricted to the Patuxent River and adjacent Chesapeake Bay transects in the vicinity of Solomons, Maryland. During January and February 1989, sampling was restricted to the same area sampled in 1988. In March 1989 the breadth of sampling was increased and a Maryland baywide survey initiated. Sampling was conducted in the mainstem of the Chesapeake Bay from the Virginia/Maryland state line to the Chesapeake and Delaware Canal in water deeper than 10 feet, as well as in the Patuxent and Choptank Rivers. Trawling was standardized to a distance of 0.5nm using a four seam high-rise otter trawl with 30 foot sweep. For a description of the 1988 and 1989 sampling protocol the reader is referred to Rothschild et al. (1989) and Rothschild (1990).

Data collected during each haul included species-specific information such as total number caught, total biomass, total length, sex, and scales for age analysis. In addition, measurements of the physicochemical parameters water depth, tide, salinity, water temperature, and dissolved oxygen were also recorded.

Results

APPLICATION SYSTEM

CPUE has been standardized to catch per 0.5 nm trawl. Because trawling gear varied historically, and there were no trawl calibration studies, the survey design characteristics stratification variable, number of strata, and sample size are determined using only 1989 data. Stratum boundary determination uses all data. Insofar as stratum boundary determination relies on relationships of catch (numbers) and stratification variable (e.g. salinity) expressed in histogram form, and not on CPUE modeling, all historical trawl data could be pooled and histograms developed. Data manipulation and analysis for data file development was facilitated using SAS-PC (SAS institute, Inc.,1990).

Monthly size-specific CPUE contour maps and X-Y plots for white perch collected in both the mainstem Chesapeake Bay and Patuxent River from 1961-1975 and 1989 are included in the system and available for viewing. Contour maps are used to delineate density and distribution in the mainstem Chesapeake Bay. Delineation in the Patuxent River is facilitated using X-Y CPUE by river mile plots. A total of 94 mainstem CPUE contour maps and 131 Patuxent River X-Y plots reside in FISHMAP.

Monthly size-class specific stratification submodels developed for white perch in the mainstem Chesapeake Bay and Patuxent River identify salinity, temperature and their squares as factors affecting distribution. The adequacy of stratification submodels to identify factors affecting white perch distribution is substantiated by significant correlations between CPUE and environmental parameters ($R^2 > 0.4$, $p < 0.05$).

Presently FISHMAP limits stratification to a single variable. However, in many submodels temperature and salinity significantly affect white perch distribution. For these months, salinity alone is identified as stratification variable. Justification for FISHMAP choosing salinity was its greater importance in determining species distribution. In all months, except July and August, high correlations between salinity and temperature ($R^2 \geq 0.90$) were consistently observed . During the summer months (July and August) water temperature was relatively uniform while salinity continued to exhibit a north-south gradient. White perch distribution in the summer months was similar to that observed in other months, limited to the upper Chesapeake Bay, suggesting the greater importance of salinity in determining distribution.

FIELD TESTING AND EVALUATION

To determine the effectiveness of FISHMAP, in particular, the field sampling module, field testing was conducted during July 25 and 26, 1990. Testing was restricted to the mainstem Chesapeake Bay, north of latitude 39°00.00'. During testing, total sample size was constrained to 14. Sampling was conducted using a 20 foot high-rise bottom trawl. Trawling distance was standardized at 0.5 nm at a speed of approximately 3 knots. Data collected during each haul was the same as those in 1988 and 1989.

Prior to going in the field, sample design characteristics were identified through consultation with FISHMAP's sample design module. Salinity was identified as stratification variable and 4 strata required with the following boundary values:

Stratum Number	Lower Strat. Variable	Upper Strat. Variable	Lower Latitude Boundary	Upper Latitude Boundary
1	0.0	1.5	39°16.10'	39°27.00'
2	1.5	4.5	39°08.98'	39°16.10'
3	4.5	7.5	39°03.43'	39°08.98'
4	7.5	16.5	39°00.00'	39°03.43'

It should be noted that the above latitudinal boundaries are based on historical trawl data and are not real-time boundary locations. In the following discussion on gains in precision, this stratification scheme will be referred to as the historical stratification scheme.

Once in the field real-time measurements of the stratification variable salinity were collected. The information (salinity and location coordinates at each stratification variable sample site) was directly input into FISHMAP and the real-time location of stratum boundaries identified as:

Stratum Number	Lower Strat. Variable	Upper Strat. Variable	Lower Latitude Boundary	Upper Latitude Boundary
1	0.0	1.5	39°22.41'	39°30.51'
2	1.5	4.5	39°17.76'	39°22.41'
3	4.5	7.5	39°13.12'	39°17.76'
4	7.5	16.5	39°00.00'	39°13.12'

Sample size was allocated among strata, sample sites randomly chosen, and the trawl fished.

To determine the efficacy of real-time stratification with FISHMAP, the method of Sukhatme and Sukhatme (1970) was used to estimate gains in precision of relative abundance estimates (CPUE) due to stratification (both historical and real-time schemes) relative to simple random sampling. To estimate gains in precision, a "sufficiently" large sample size was required, one larger than that collected during FISHMAP testing. Because field testing occurred during the last two days of a regularly scheduled

Maryland baywide trawling cruise, which sampled in the same vicinity of testing (upper Chesapeake Bay), these data (n=15) were pooled and gains in precision evaluated. Pooling provided a larger sample size (n=29), while also ensuring complete spatial coverage of the sampling area (upper Chesapeake Bay), which was not accomplished using test data alone. Because trawling operations (gear, vessel, crew) between the two cruises were identical and occurred consecutively, it is unlikely that pooling introduced serious bias.

Comparisons between historical and real-time stratum boundary locations indicate significant changes in stratum boundary locations (Figure 4). Stratum 1 decreased by 43%, while stratum 2 moved further up the Bay with a concurrent decrease in area of 78%. Likewise, stratum 3 moved further up the Bay with a concurrent decrease in area of 41%, while stratum 4 showed a five fold increase in stratum area. Sampling site locations relative to stratum boundary locations of each stratification scheme are shown in Figure 5. Stratified mean and variance estimates, as well as percent gains in precision (relative efficiency) of mean CPUE estimates owing to each stratification scheme were estimated as:

Stratification Scheme	Stratified Mean ($\overline{X}_{st}$)	Stratified Variance ($VAR(\overline{X}_{st})$)	Relative Efficiency (RE)
Historical	45.8	305.8	21.0
Real-time	32.5	142.2	9.8
Simple Random Sampling	61.7	1455.0	100.0

The relative efficiency estimates indicate substantial gains in precision owing to both stratification schemes. However, gain in precision owing to the real-time stratification scheme of FISHMAP (90.23%) is substantially larger than that observed using the historical stratification scheme (78.71%). It should be noted however, that the results presented are from a single sampling period. The choice of different sampling period would result in different gains in efficiency, where the amount of the gains would be different but the relative position of each would probably remain the same.

Discussion

The development and application of an expert decision support system designed to aid in sampling has been defined and described. FISHMAP's modular design is intended for use in monitoring, interpreting, and diagnosing incoming streams of physical data to optimize real-time collection of population data. For demonstration purposes the system was configured for trawling.

Specific operational objectives of FISHMAP were identified at the outset of model development, and the model constructed to meet the goals. FISHMAP's modular design provides:

1) strategy to determine number and real-time dimensions of strata

(2) definition of the spatial sampling characteristics (where to sample)

(3) identification of stratification variable(s)

(4) strategy for efficient allocation of sampling effort among strata, and

(5) methodologies to assess data recovered in real-time and if necessary suggest efficient design modifications.

These operations represent two phases of pre-survey design activity: (1) survey design development based on historical data (model operations 1 through 4) and (2) survey design modification using data collected in real-time (model operation 5). While contemporary survey design operations employ pre-survey activities, they are limited to activities concerning survey development using historical data (model operations 1 through 4). Factors affecting distribution are considered invariant.

These differences in survey design operations represent fundamental differences in sampling philosophies. Contemporary design methodologies assume stability of survey design characteristics. In contrast, FISHMAP explicitly recognizes the dynamics of design characteristics, quantifies changes, and provides updated survey design metrics. This approach is well suited given the variability of the medium sampled---aquatic environments. The dynamics of stratum boundary locations was illustrated during field testing, demonstrating the utility of FISHMAP to represent activities necessary for survey design development. Failure to determine proper design characteristics often results in surveys which are of little direct use to the resource scientist because of highly variable estimates. In most cases, parameter estimates are less precise than those of simple random sampling (Dalenius, 1950; Sukhatme and Sukhatme, 1970).

Computational procedures embedded in FISHMAP allow for the determination of sample design characteristics using regression analysis and standard sampling formulae. The canonical intent of these procedures was maintained through development of generalized external programs. The modular design of FISHMAP facilitated development of a menu driven system promoting user friendliness.

For FISHMAP to be effective, the embedded computational rocedures must be efficient. It is obvious that the best characteristic for stratification is the value of the variable under study, CPUE. In practice, however, we cannot stratify by values of CPUE. Large gains in precision can still be realized by satisfying the following three conditions: (i) the principle variate being measured, CPUE, must be closely related to another variate (e.g. salinity), (ii) the related variate must vary widely in measure (producing varying measures of CPUE), and (iii) accurate measures of the related variate must be available for setting up strata (Cochran, 1977). FISHMAP explicitly allows for the identification of variables related to CPUE through regression analysis and development of stratification submodels. To determine number of strata, two questions relevant to the decision are: (i) at what rate does the variance of the stratified mean CPUE decrease as strata number is increased? and (ii) how is cost of the survey affected by increases in strata number? The importance of knowing this information when determining number of strata was demonstrated by Dalenius (1957). As number of strata increases gains in precision decrease with a concomitant increase in survey cost. The 10% relative gain stopping rule identifies strata number for a specified gain in precision. Extant FISHMAP procedures assume that costs have no effect, a rather simplistic assumption. To provide optimal estimates of strata number, a cost function is required.

Although a number of approximation methods to determine stratum boundaries are available the cum$\sqrt{f}$rule of Dalenius and Hodges (1957) is used by FISHMAP. Cochran (1961) compared the various approximations to populations having skewed distributions and found that the cum$\sqrt{f}$ rule performed well in all cases.

For purposes of illustrating the utility of FISHMAP, the system was applied to stocks of white perch in Chesapeake Bay. The feasibility of FISHMAP to guide researchers through the development of sampling designs has been demonstrated. Field testing of the system punctuated the importance of assimilating real-time data into stratified sampling design procedures. This was demonstrated to be of particular importance when sample design characteristics are linked to environmental conditions.

The white perch stratification submodels punctuate the importance of analyses to determine stratification variable. Model results suggested that salinity and temperature are factors influencing white perch distribution. This contrasts with most existing stratified trawling surveys whose purpose is to estimate abundance of white perch; strata are developed according to depth. For many of these surveys, justification for choosing depth is out of convenience or an inferred but unanalyzed relationship between depth and catch.

A users guide for FISHMAP is available (DiNardo and Li, 1991). Those interested in obtaining a copy should contact the senior author.

REFERENCES

Clark, S. 1981. Use of trawl survey data in assessments. In Bottom trawl surveys, W.G. Doubleday and D. Rivard (eds.) Can. Spec. Publ. Fish. Sci., 58:82-92.

Cochran, W. G. 1961. Comparison of methods for determining strata boundaries. Bull. Inter. Stat. Inst., 38(Part II):345-358.

Cochran, W. G. 1977. Sampling Techniques. John Wiley and Sons, New York, 428 p.

Dalenius, T. 1950. The problem of optimum stratification. Skandinavisk Aktuarietidskr, 3-4:203-213.

Dalenius, T. 1957. Sampling in Sweden: Contributions to the methods and theories of sample survey practice. Almqvist Och Wiksell, Stockholm.

Dalenius, T., and J.L. Hodges, Jr. 1957. The choice of stratification points. Skandinavisk Aktuarietidskr, 3-4:198-203.

Dalenius, T., and J.L. Hodges, Jr. 1959. Minimum variance stratification. J. Amer. Stat. Assoc., 54:88-101.

DiNardo, G., and Huaixiang Li. 1991. FISHMAP: Users manual and reference guide. Manuscript.

DiNardo, G.T., J.S. Ault, B.J. Rothschild and M. Holloway. 1991. Population dynamics and stock assessment of the white perch (Morone americana) in the northern Chesapeake Bay. Final report to MD Dept. of Natural Resources, Project No.CB90-001-003, University of Maryland, CEES, Chesapeake Biological Laboratory, Solomons, MD.

Doubleday, W.G., and D. Rivard. 1981. Bottom trawl surveys. Canadian Special Publication of Fisheries and Aquatic Sciences, No. 58, Department of Fisheries and Oceans, Ottawa, 273 p.

Francis, R. I. C. C. 1984. An adaptive strategy for stratified random trawl surveys. New Zealand Journal of Marine and Freshwater Research, 18:59-71.

Gavaris, S., and S. J. Smith. 1987. Effect of allocation and stratification strategies on the precision of survey abundance estimates for Atlantic cod (Gadus morhua) on the eastern Scotian shelf. Journal of Northwest Atlantic Fisheries Science, 7:137-144.

Grosslein, M.D. 1969. Groudfish survey program of BCF Woods Hole. Commer. Fish. Review 31:22-35.

Halliday, R. G., and P. A. Koeller. 1981. A history of Canadian groudfish surveys and data usage in ICNAF Divisions 4TVWX. In: Bottom trawl surveys, W. G. Doubleday and D. Rivard (eds.) Can. Spec. Publ. Fish. Sci., 58:27-41.

Information Builders, Inc. 1988. Level 5, Expert Systems Software. Version 1.2. Information Builders, Inc., New York.

Lemmon, H. 1986. COMAX : An expert system for cotton crop management. Science, 233:29-33.

Mansueti, R. 1961. Movements, reproduction, and mortality of the white perch, Roccus americanus, in the Patuxent River Estuary, Maryland. Chesapeake Science, 2:142-205.

Mulligan, T. J. 1987. Identification of white perch (Morone americana) stocks in Chesapeake Bay based on otolith composition and mitochondrial DNA analysis. Ph.D Thesis, University of Maryland, 108 p.

Neter, J., W. Wasserman, and M.H. Kutner. 1985. Applied linear statistical models, Second Edition. Richard D. Irwin, Inc., Homewood, Illinois, 1127 p.

Neyman, J. 1934. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. J. Royal Stat. Soc., 97:558-606.

Palmer, R.N., and K.J. Holmes. 1988. Operational guidance during droughts: Expert system approach. J. Water Res. Plan. and Mgmt., 114:647-666.

Pitt, T. K., R. Wells, and W. D. McKone. 1981. A critique of research vessel otter trawl surveys by the St. John's research and resource services. In: Bottom trawl surveys, W. G. Doubleday and D. Rivard (eds.) Can. Spec. Publ. Fish. Sci., 58: 42-61.

Reboh, R., J. Reiter, and J. Gaschnig. 1982. Development of a knowledge-based interface to a hydrological simulation program. SRI Int., Menlo Park, California.

Rothschild, B.J. 1990. Final Report. Development of a sampling expert system: "FISHMAP." Submitted to MD Dept. Natural Resources and U.S. Dept. of Int., Fish and Wildlife Ser., Project No. F171-89-008. Univ. of Maryland CEES Ref.No. [UMCEES] CBL90-090; Chesapeake Biological Laboratory; Solomons. 609p.

Rothschild, B.J., G. DiNardo, M. Holloway, M. Bhandary, D. Levy, D. Turlington, C-F Tsai, L. Baylis, and M. Wiley. 1989. CHESFISH, A cooperative study of the fishery resources of Chesapeake Bay. Final Report. Submitted to Md Dept. Natural Resources and U.S. Dept Int., Fish and Wildlife Ser., Project No. F130-88-008. Chesapeake Biological Laboratory, Solomons.

Ryan, J. D. and P. E. Smith. 1985. An "Expert System" for fisheries management. Southwest Fisheries Center, National Marine Fisheries Service, La Jolla.

SAS Institute, Inc. 1990. SAS for the Personal Computer, Version 6. Cary, NC.

Setzter-Hamilton, E.M. 1990. Habitat requirements for white perch, Morone americana. University of Maryland CEES Ref. No. UMCEES[CBL]90-129; Chesapeake Biological Laboratory. 73p.

Stagg, C. 1990. The expert support system as a tool in fishery stock management. pp 299-314, In A.G. Rodrigues (Ed.). Operations research and management in fishing. Kluwer Academic Publishers. 340p.

Sukhatme, P. V., and B. V. Sukhatme. 1970. Sampling Theory of Surveys With Applications. Iowa State University Press, Ames, Iowa, 452 p.

Survey Working Group, Northeast Fisheries Center. 1988. An Evaluation of the Bottom Trawl Survey Program of the Northeast Fisheries Center. NOAA Technical Memorandum NMFS-F/NEC-52, 83 pp.

Turban, E. 1988. Decision Support and Expert Systems. Macmillan Pub. Co., New York, 697 p.

Van der Gaast, N. G. and L. H. de Jonge. 1988. The use of expert systems for determining sampling strategies in contaminated site investigations. Trends in Analytical Chemistry, 7:283-285.

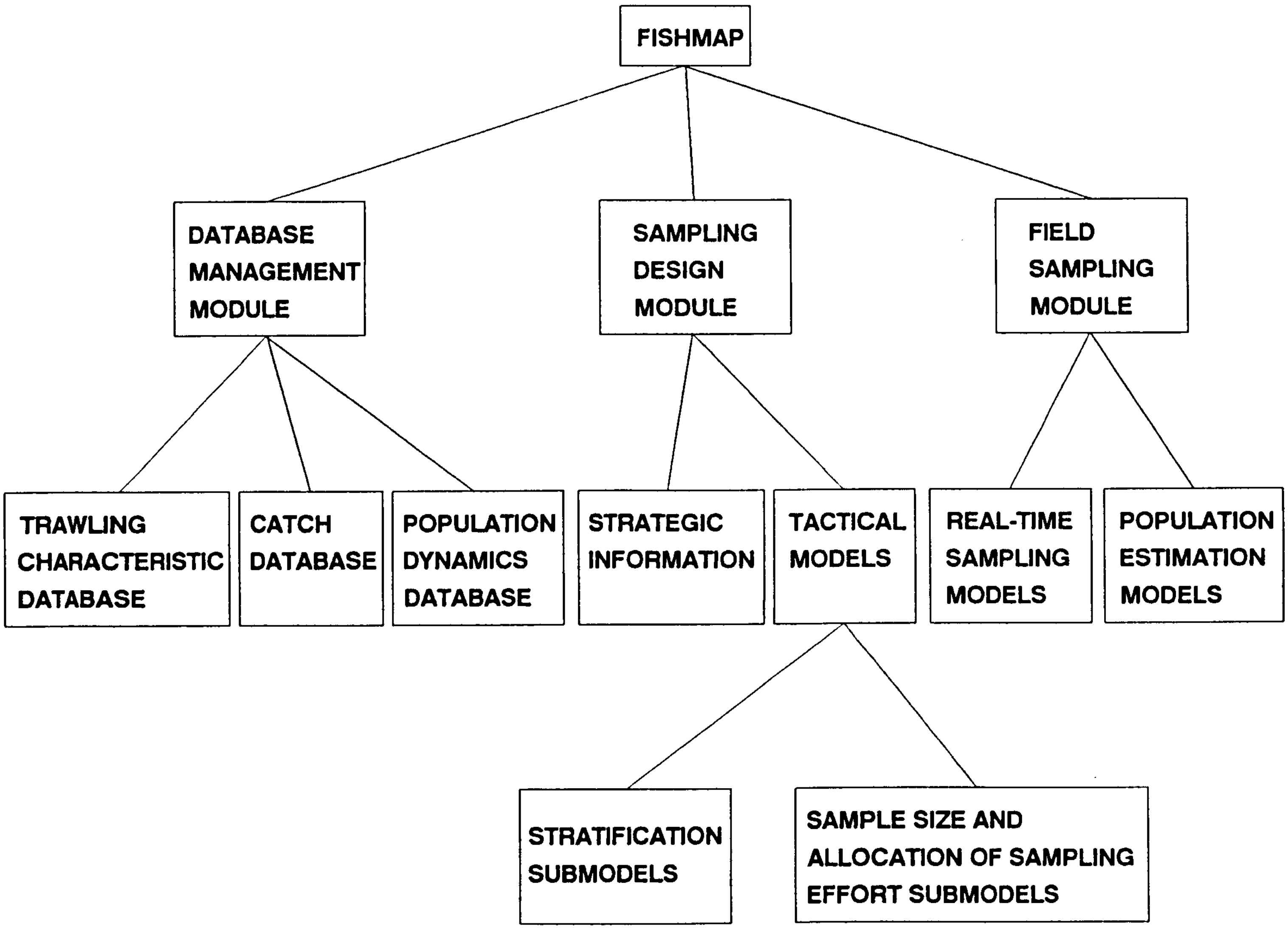
Figure 1. Architecture of the expert decision support system FISHMAP.

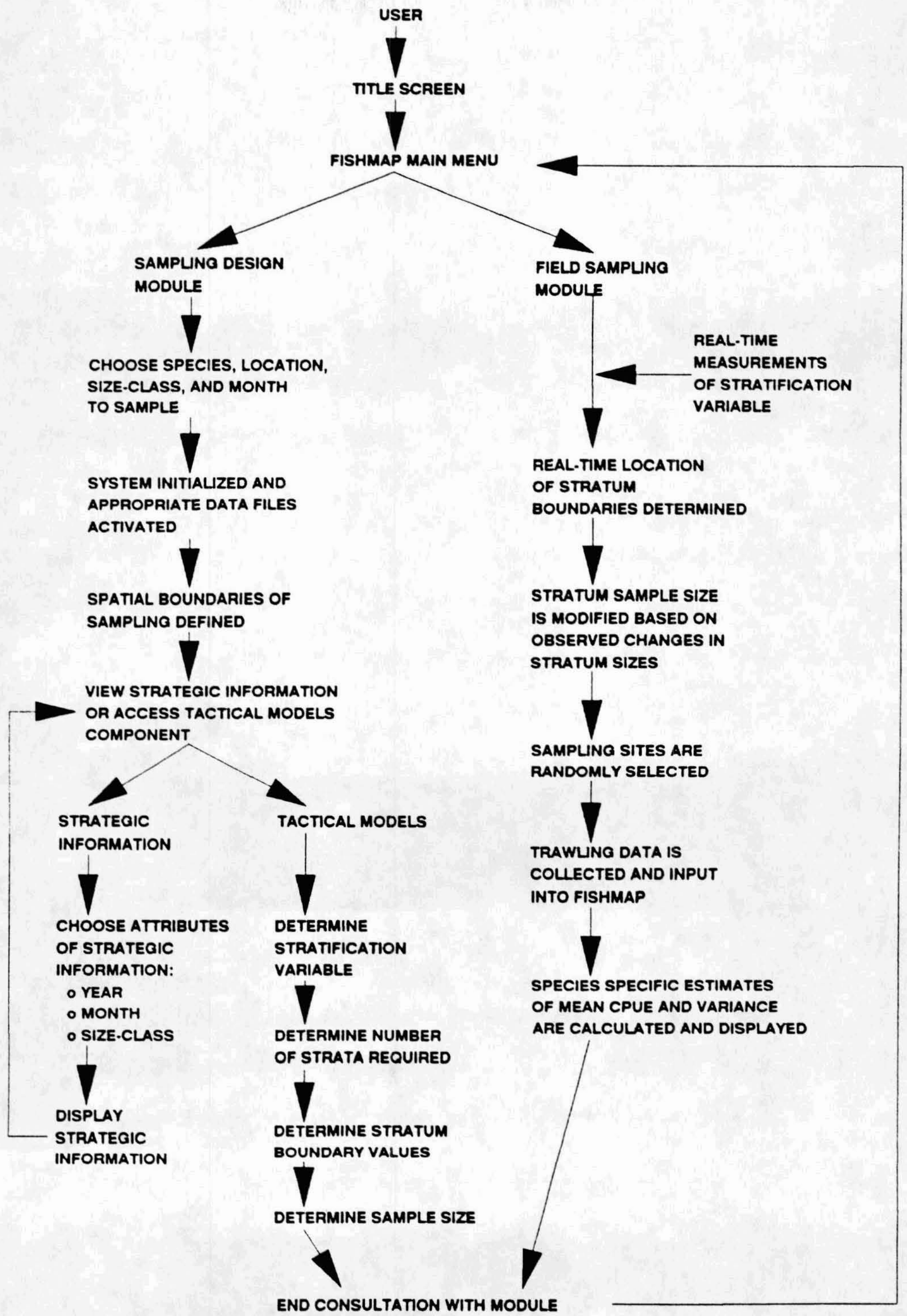
Figure 2. Operation of FISHMAP's sampling design and field sampling modules.

Figure 3. Relative gain in variance reduction plot.

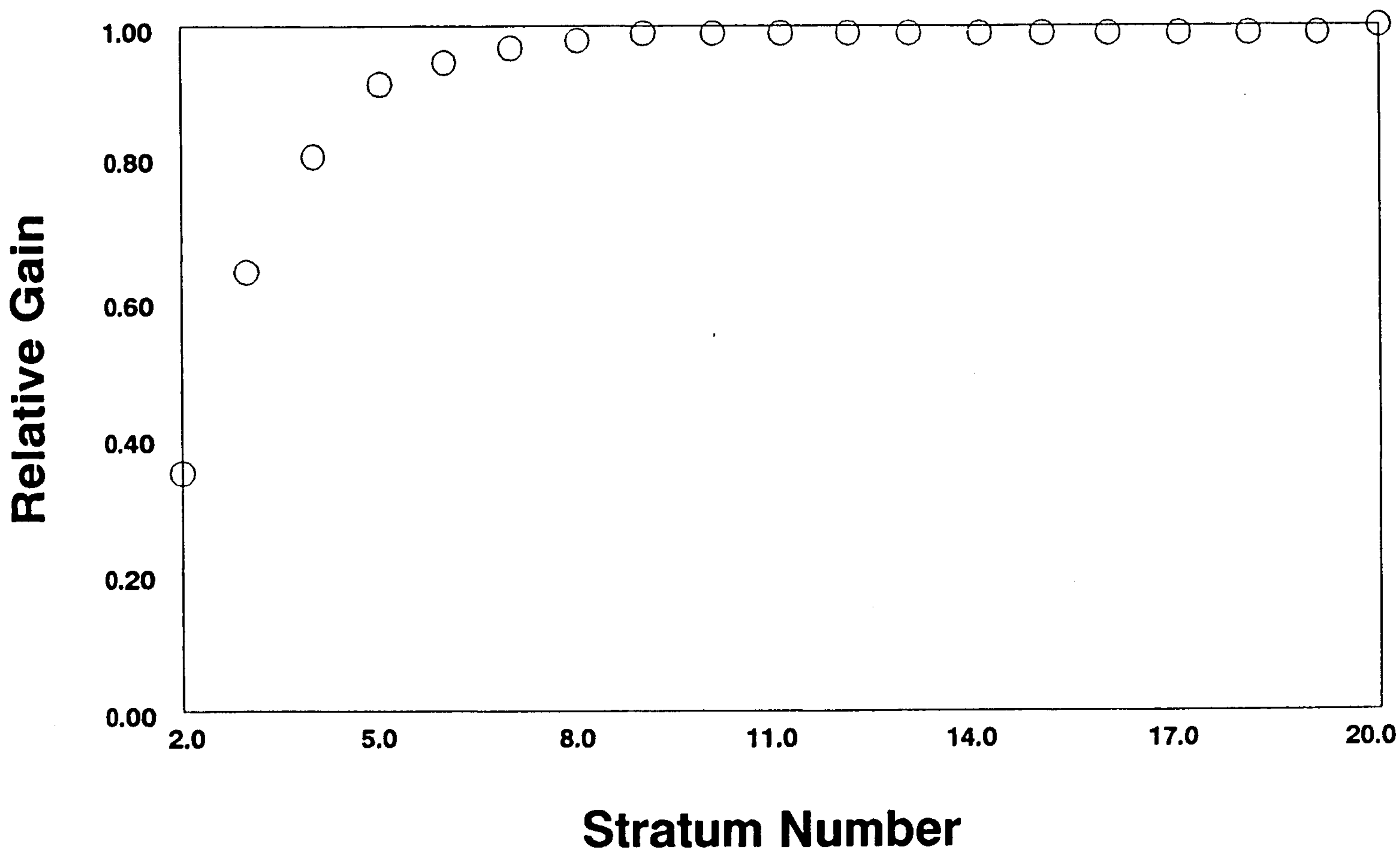
Figure 4. Location of stratum boundaries in the upper Chesapeake Bay based on historical salinity data and real-time salinity measurements.

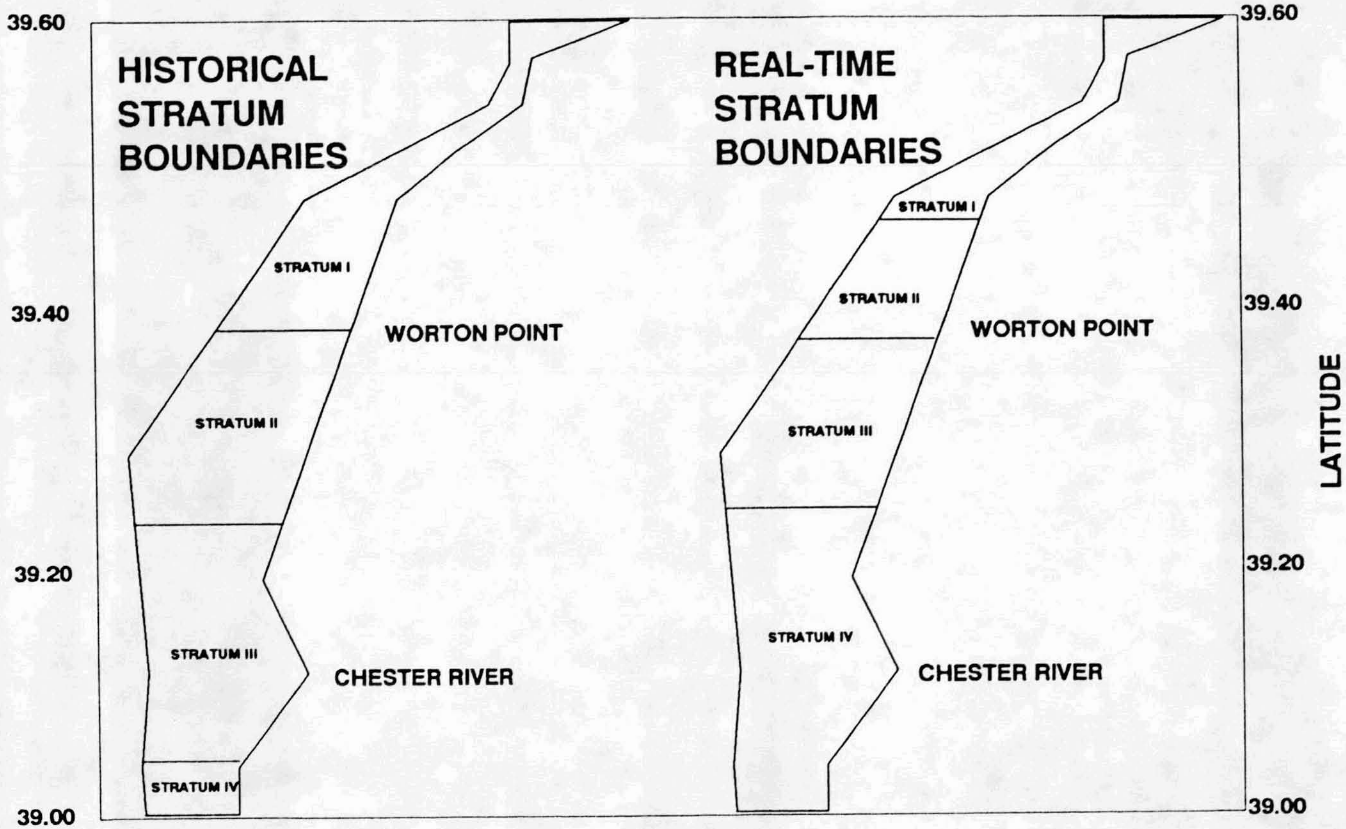
Figure 5. Location of sampling sites (*) relative to historical (A) and real-time (B) salinity stratum boundary positions.


FISHMAP
DATABASE MANAGEMENT MODULE
SAMPLING DESIGN MODULE
FIELD SAMPLING MODULE
TRAWLING CHARACTERISTIC DATABASE
CATCH DATABASE
POPULATION DYNAMICS DATABASE
STRATEGIC INFORMATION
TACTICAL MODELS
REAL-TIME SAMPLING MODELS
POPULATION ESTIMATION MODELS
STRATIFICATION SUBMODELS
SAMPLE SIZE AND ALLOCATION OF SAMPLING EFFORT SUBMODELS

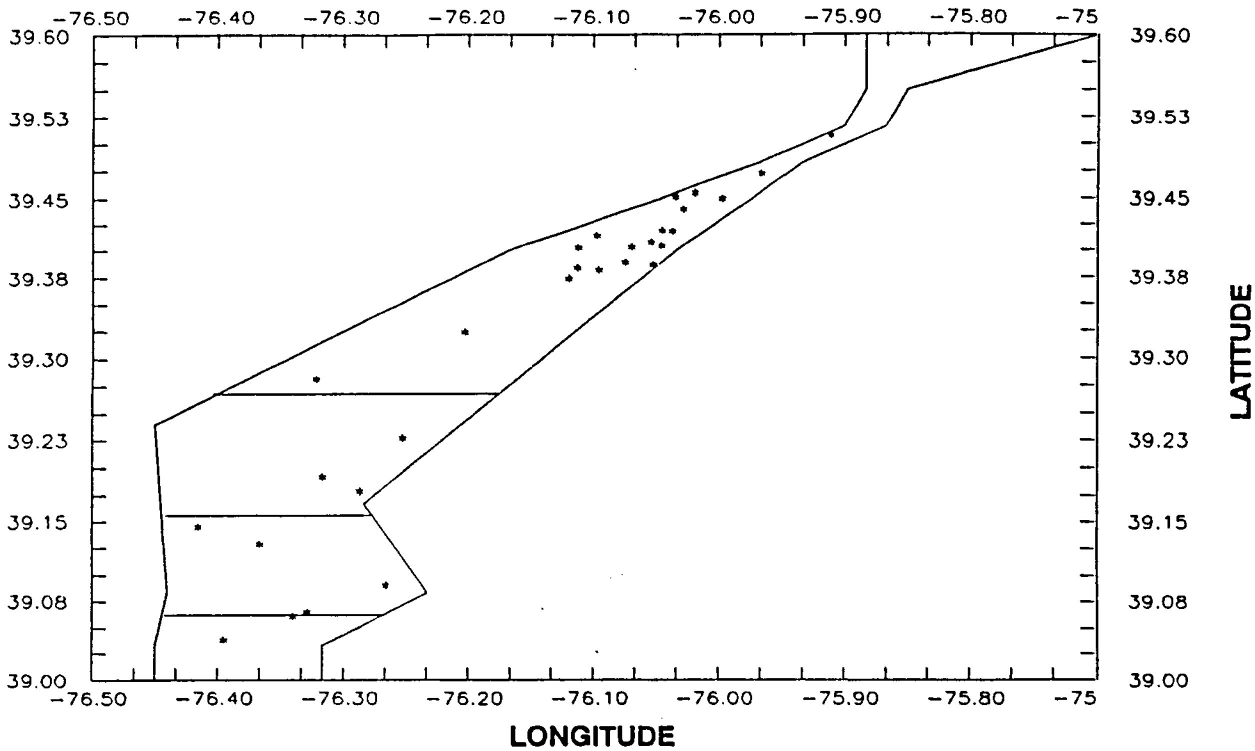

USER
TITLE SCREEN
FISHMAP MAIN MENU
SAMPLING DESIGN MODULE
FIELD SAMPLING MODULE
CHOOSE SPECIES, LOCATION, SIZE-CLASS, AND MONTH TO SAMPLE
REAL-TIME MEASUREMENTS OF STRATIFICATION VARIABLE
SYSTEM INITIALIZED AND APPROPRIATE DATA FILES ACTIVATED
REAL-TIME LOCATION OF STRATUM BOUNDARIES DETERMINED
SPATIAL BOUNDARIES OF SAMPLING DEFINED
STRATUM SAMPLE SIZE IS MODIFIED BASED ON OBSERVED CHANGES IN STRATUM SIZES
VIEW STRATEGIC INFORMATION OR ACCESS TACTICAL MODELS COMPONENT
SAMPLING SITES ARE RANDOMLY SELECTED
STRATEGIC INFORMATION
TACTICAL MODELS
TRAWLING DATA IS COLLECTED AND INPUT INTO FISHMAP
CHOOSE ATTRIBUTES OF STRATEGIC INFORMATION:
o YEAR
o MONTH
o SIZE-CLASS
DETERMINE STRATIFICATION VARIABLE
SPECIES SPECIFIC ESTIMATES OF MEAN CPUE AND VARIANCE ARE CALCULATED AND DISPLAYED
DETERMINE NUMBER OF STRATA REQUIRED
DISPLAY STRATEGIC INFORMATION
DETERMINE STRATUM BOUNDARY VALUES
DETERMINE SAMPLE SIZE
END CONSULTATION WITH MODULE

Relative Gain vs Sratum Number




HISTORICAL
STRATUM
BOUNDARIES
REAL-TIME
STRATUM
BOUNDARIES
39.60
39.40
39.20
39.00
STRATUM I
STRATUM II
STRATUM III
STRATUM IV
WORTON POINT
CHESTER RIVER
LATITUDE

HISTORICAL STRATIFICATION SCHEME



FISHMAP STRATIFICATION SCHEME

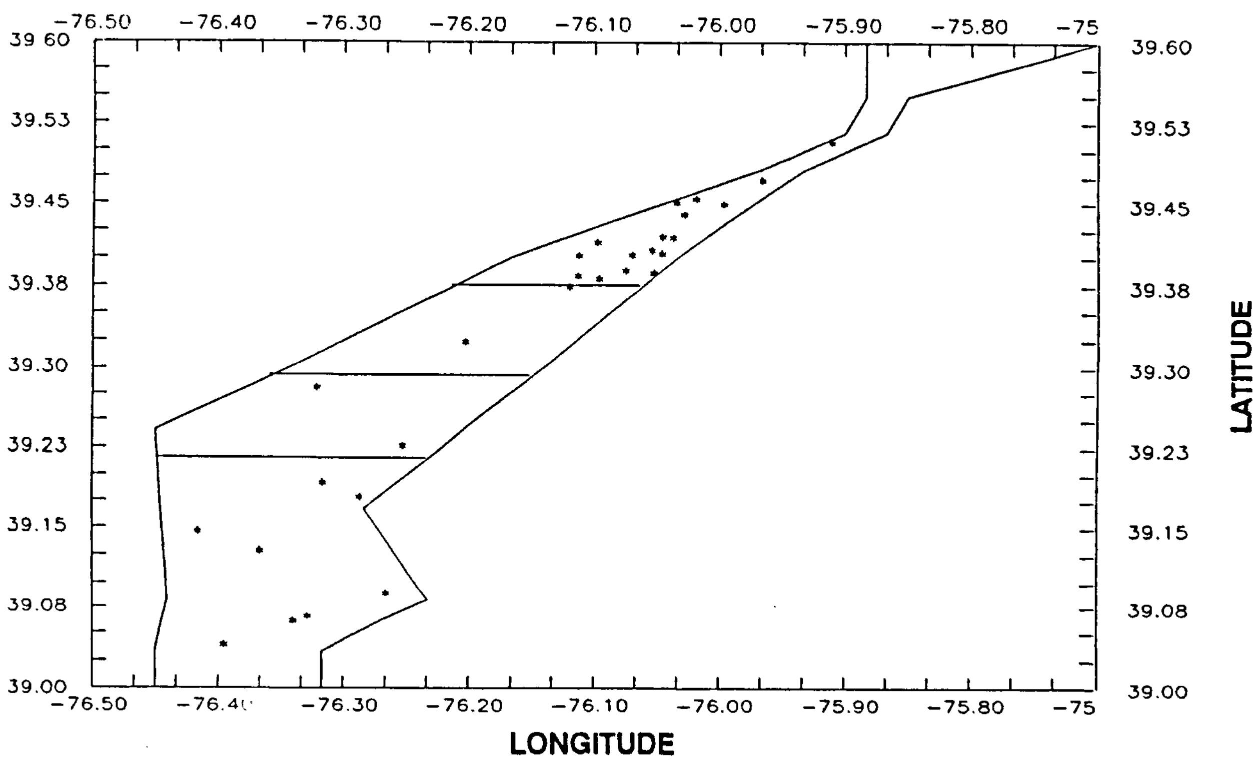


Table 1. Supporting data used to develop FISHMAP's application system.

SURVEY	YEARS	NUMBER OF TRAWLING SITES	SAMPLING DESIGN	LOCATION OF TRAWLING	SURVEY OBJECTIVE
A	1961-64	38	FIXED SITE	MAINSTEM, POTOMAC R., PATUXENT R., AND CHOPTANK R.	WINTER JUVENILE STRIPED BASS SURVEY
B	1965-67	25	FIXED SITE	MAINSTEM, POTOMAC R., CHESTER R., PATUXENT R., AND CHOPTANK R.	WINTER FLOUNDER STUDY
C	1966-68	10	FIXED SITE	UPPER CHESAPEAKE BAY	IMPACT OF DREDGING AND DREDGE SPOIL ON FISHERY RESOURCES
D	1967-75	10	FIXED SITE	ADJACENT TO CALVERT CLIFFS NGS	IMPACT ASSESSMENT OF FACILITY
E	1970-73	11	FIXED SITE	CHESAPEAKE AND DELAWARE CANAL	USE OF C & D CANAL BY STRIPED BASS
F	1973-75	68	FIXED SITE	ELEVEN TRIBUTARIES OF CHESAPEAKE BAY	ASSESS IMPORTANCE OF TRIBUTARIES TO GROUNDFISH
CHESFISH	1988	36	STRATIFIED RANDOM ALONG TRANSECTS	PATUXENT R. AND MAINSTEM ADJACENT TO SOLOMONS,MD	GEAR TESTING AND ID OF SOURCES OF VARIABILITY
CHESFISH	1989	71	STRATIFIED RANDOM IN MAINSTEM AND FIXED SITE IN TRIBUTARIES	MAINSTEM, PATUXENT R., AND CHOPTANK R.	SATISFY BAY AGREEMENT AND COLLECT DATA FOR FISHMAP DEVELOPMENT