#1966

COMPUTER SIMULATION TECHNIQUES TO ASSESS ANGLER SURVEY METHODOLOGY

David L. Wade
Applied Marine Research Laboratory
Old Dominion University
Norfolk, Virginia 23529

Cynthia M. Jones
Old Dominion University
Norfolk, Virginia 23529

Douglas S. Robson
150 MacLaren St. PH6
Ottawa, Ontario
Canada K2P OL2

Kenneth H. Pollock
Department of Statistics
North Carolina State University
Raleigh, N.C. 27695-8203

ACKNOWLEDGEMENTS

The authors extend thanks to The Maryland Department of Natural Resources for the use of survey data, and special thanks to Dr. John Hoenig, Mr. Bill Check and Mr. Scott Cone for their insight and help. This study was funded by Virginia Sea Grant through grants no. 5-29640 and 5-29702 (principals CMJ and KHP).

INTRODUCTION

Marine recreational angling is a component of fishing which is difficult to estimate. In 1985, an estimated 13.7 million anglers fished for 155.2 million days in the nation's oceans, gulfs, bays, and estuaries, and total marine recreational fishing expenditures for 1985 were 7.2 billion dollars (USFWS, 1988). In order to effectively manage these recreational fisheries it is important to have unbiased and precise estimates of catch and effort. Angler intercept surveys that contact angling parties during the process of fishing are an effective method of obtaining estimates of marine recreational catch and effort, and in some fisheries they may be the only practical means of acquiring these estimates. The absence of marine angling licensing programs in many states and concentrations of coastal residents makes marine recreational anglers difficult to identify and contact when they are not fishing.

There are two primary methods of conducting intercept surveys: the access site method and the roving method. The access method employs clerks who interview angling parties which have completed their trips and are leaving the fishery at access sites, such as: boat ramps, piers, and marina parking lots. Clerks inspect the catch and conduct interviews in order to obtain catch and effort information. Some fisheries have diffuse access sites, and angling parties in this case may complete fishing trips at homes, private docks, remote parking sites, and otherwise inaccessible areas. As efficient sampling of such sites is not possible, the access site method is not applicable to these fisheries. The roving clerk method is applicable to these diffuse access fisheries, and is conducted by using clerks who move through the fishery and intercept angling parties in the process of fishing. Catch and effort information for incomplete trips is obtained and expanded for complete trip estimates. However, the estimates obtained from a roving clerk survey have significant problems with bias and the derivation of variance components (Robson, 1961), and these problems are relatively mathematically intractable.

In order to examine the properties of bias and variance, we must know the actual total catch and effort for a population of anglers, and these true values can be obtained in one of two ways. We can completely census the fishery, intercepting every angler, or we can use a computer simulation model. A complete census of a diffuse access fishery is logistically and economically impractical. Hence, the preferred approach of simulation modeling was chosen, and implemented in a general manner so that data from a variety of fisheries can be modeled. In this initial study the model was applied to the recreational blue crab (*Callinectes sapidus*) fishery of the Mid-Atlantic Chesapeake Bay area.

The diffuse angler access, high value and lack of recreational catch and effort statistics for this recreational crab fishery make it an ideal subject in which to study the application of the roving clerk design. This diffuse access fishery is considered to comprise a single stock (Cronin et al, 1987) that is distributed in both rural and urban areas among a complex system of bays, tributaries and salt marshes. Crabs are captured recreationally with a variety of gears, including: traps, pots, trotlines, handlines, and dipnets. In 1986 commercial landings were approximately valued at 31 million dollars (Cronin et al, 1987), but few studies have been performed to determine the magnitude of the recreational blue crab fishery (Williams et al, 1982). Fisheries managers have estimated the recreational component to be on the order of 50%, but the National Marine Fisheries Service (NMFS) marine recreational fishing survey has not assessed the fishery except through special request and funding by Maryland in 1979 and 1983. In addition, Virginia has undertaken no surveys of this fishery and has no licensing for marine recreational anglers.

The objectives of this study were to develop a computer simulation model of the roving clerk survey design, and apply the model to the blue crab fishery of Chesapeake Bay. The feasibility of conducting a survey in the field was investigated by examining the bias and sampling distribution of the roving clerk design, and the model was generalized to allow modifications of the survey design and acceptance of data from other fisheries.

MATERIALS AND METHODS

The model was developed as a high level programming language, and was coded in Pascal. This language implementation provides a modular and flexible tool similar to commercial statistical analysis systems (e.g. SAS, SPSS), and the Pascal programming language provides modularity, peer acceptance, and portability from microcomputer to mainframe. The model is portable and was compiled on various IBM compatible personal computers and an IBM mainframe.

The application of the model to the blue crab fishery of the Chesapeake Bay used catch and effort data from a 1983 Maryland recreational angler survey. The data was compiled and provided by the Maryland Department of Natural Resources. For this application of the model the Maryland recreational dipnet fishery was simulated.

Various commercial graphics packages were interfaced to the model to allow detailed monitoring of its function, and statistical output was analyzed through the use of Statistical Analysis System (SAS) programs.

DESCRIPTION OF THE MODEL

The model was constructed in 2 main modules. The first component is a population module which stochastically generates an angler population with the same characteristics that are exhibited in the field (thus establishing realistic "ground truth"), and the second component is an estimation module which simulates the execution of a roving clerk survey on the angler population (thus providing estimates of total catch and total effort). The actual values of total catch and total effort from the population module are compared to the estimates produced by the estimation module, and through iteration of this process the bias and sampling distribution of the estimates are revealed.

POPULATION MODULE

Each population produced by the population module consists of an array of stochastically generated data records, each record representing a single angling party. These parties have the attributes of geographic location, starting time, effort, and catch.

The geographic location for each angling party is assigned the value of a uniformly distributed random variate. Parties are stationary once they have entered the fishery, and are distributed on a single, unitless dimension. Parameterization of population sizes for a survey day and the length of a survey day provide geographic densities similar to what is expected in the field.

The starting time for each party is generated stochastically, and is based on a distribution of starting times observed in the field. A probability density function constructed from field data is expressed in the form a cumulative density function, and when a distribution of uniformly distributed random variates is applied to the function then the resulting distribution closely resembles the probability density function observed in the field.

Effort for an angling party is the duration (hours) of its fishing trip, and for the Maryland crab data, effort could be fit as a linear model of starting time. Stochasticity is added to the simulation by fitting linear models of mean and variance of effort to the number of hours in the fishing day minus starting time:

$$\mu(\text{effort}) = (T_F - T_S)\beta$$

$$\sigma^2(\text{effort}) = (T_F - T_S)\beta'$$

(where β and β' are the OLS regression coefficients, T_S is the starting time [hrs] for an angling party, and T_F is the end of the fishing day). Day length minus starting time is used in these functions, so that effort values will decrease to zero at the end of each day. For each angling party, these values of mean effort and variance of effort describe normal distributions, and from these distributions a value of effort for the party is chosen randomly.

For the Maryland data, catch could be fit as a linear model of effort , and stochasticity is added to the model by fitting models of mean and variance of catch to effort.

$$\mu(\text{catch}) = (\text{effort})\beta$$

$$\sigma^2(\text{catch}) = \mu(\text{catch}) + \beta_2(\mu(\text{catch}))^2$$

The value of catch for each angling party is determined by the application of the mean and variance for catch to a normally distributed random number generator. In the Maryland crab fishery, catch could be fit as a linear model of effort, but for other fisheries catch rate may be a linear function of effort. For example, successful parties may fish for longer periods than unsuccessful parties. To model relationships of catch rate to effort we add a curvi-linear component to the linear model for mean of catch.

$$\mu(\text{catch}) = (\text{effort})\beta + (\text{effort}^2)\beta'$$

The relationship of catch rate to effort is specified by assigning the parameter B' to a negative or positive value.

Estimation Module

When the simulation model has generated a population of anglers with known values of total catch and total effort, a roving creel survey is conducted by the estimation module to estimate total catch and effort for that population. Total effort is estimated as the product of the duration of a clerk's transect (hrs) and total parties counted:

$$ETPH_i = \Sigma^n_{j=1} (X_{ij}) \cdot T_S \cdot m$$

where $ETPH_i$ is the estimate of total party hours for day i, X_{ij} is the value of an angling party j on day i ($X_{ij}=1$ if the party was encountered by the clerk, else $X_{ij}=0$), T_S is the duration of a clerk's transect (hrs), and m is the number of transects in a day.

Total catch is expanded from interview obtained information on party catch and trip length.

Individual survey designs are specified, and they are implemented as a series of linear equations in the cartesian coordinates of geographic location on a linear survey route and time of day.

For this series of simulations two primary survey designs were chosen, a simple roving creel design and a more elaborate design with clerk scheduling. The scheduling design was chosen in an attempt to reduce the bias inherent in the more simple survey method.

The simple roving clerk design, type one, (figure 1) assigns a survey clerk a circular, linear route through a fishery (e.g. along a shore line, around an embayment). A random starting location on this route is chosen for each survey day, and the clerk is assigned a starting time and completion time for the day. The clerk's scheduled movement through the fishery can be modeled by a line segment:

Location = Travel Speed(Time of Day) + Starting Location

Slope of this line is equal to the clerk's scheduled speed, and start and end coordinates are equal to the start and end of the survey day. Each stationary angling party can also be simulated by a line segment. Slope of this line being equal to zero, and starting and ending coordinates being equal to start and end of the fishing trip. An intersection of a clerk line segment and an angling party line segment simulates a field intercept of a party by the clerk. During a survey, the clerk proceeds at his or her scheduled speed through the fishery until an angling party is intercepted, and at this point the clerk must stop movement through space, conduct an interview of some specified length, and continue moving at the original speed. Note that this causes the clerk to depart from the original schedule (the interview caused a delay). This entire process can be simulated by the use of line segments for each component of a clerk's travel and each angling party.

The second type of survey design is a modified method that uses a series of check points to ensure that the clerk remains on schedule (figure 2). The checkpoints are distributed along a route, are chosen time intervals (e.g. every 2 hours), and can be positioned as "time-located" geographic points (e.g. piers, lighthouses, boat ramps). The clerk conducts interviews as in the simple example, and with each interview, he or she will become increasingly off schedule. However, at some point the clerk's path will intercept a linear schedule segment (with a slope representing some maximum speed of travel) that if followed will return the clerk to the next check point on the original schedule. During the time that the clerk is moving at the increased speed, he or she counts angling parties without stopping to interview them.

RESULTS AND DISCUSSION

The probability of the clerk encountering an angling party is directly proportional to the party's length of stay (effort) in the fishery. That is, anglers with longer trips are more likely to be encountered. If there is a relationship between individual catch rate and effort then this will result in a bias in total catch estimates. When interview time is greater than zero, the probability of intercepting a party becomes more complex. There is a shadowing effect on interception probabilities between angling parties along the route, and this is illustrated in figure 3. Due to the time a clerk is stationary while interviewing the first party on the route, the probability of intercepting the second party is reduced by some amount up to the length of the interview, and this shadowing is additive as the clerk conducts interviews throughout the day. This results in bias in estimates of total effort.

In the simple roving clerk survey design where the clerk does not attempt to remain on schedule, this shadowing effect can cause considerable bias in effort estimation. As shown by the simulation results (figure 4) in this example of 10,000 independently simulated survey days, the bias ranges from approximately 25 percent at interview lengths of 5 minutes to 36 percent at interview lengths of 15 minutes, and the estimator always under estimates the actual value of total effort for the population. The bias in effort estimation is variable, and it is related to the length of interviews conducted.

In an attempt to avoid this bias, survey clerks can be instructed to return to scheduled routes when interviews cause them to be off schedule, and this is simulated by the modified survey design with check points. In the modified design the bias in effort estimation was reduced to approximately 1 to 2 percent (figure 5), and under normal management conditions this might be an acceptable level of bias. However the bias is still variable, and always under estimates the actual value of total effort.

It is important to understand that bias in effort results were under the best of field conditions. Angling parties were distributed in uniform distributions on the route, and in the real world it is possible that anglers are distributed in contagious distributions. The effect of clumping may increase the shadowing affect on total effort estimation and is the subject of current study.

The sampling distributions of the estimators are important for the construction of confidence intervals about estimates, and they were determined (figure 6) by a simulation run of 10,000 independently simulated survey days. The population values of total effort for each survey day were normally distributed, but the estimated values for each survey design were non-normally distributed ($p < .01$). The consequences of these results are that standard T tables can not be used with this data, and that T distributions should be constructed to make statistical inferences with roving clerk data (Jones et al, in press).

SUMMARY

The simulation model was determined to be a useful approach to examining the roving clerk design and the statistical properties of its estimators. We found the roving clerk estimator of total fishing effort to be both biased and variable in its bias. This bias was under the best of circumstances, where angling parties were uniformly distributed along a route, and could be reduced by instructing clerk's to return to scheduled survey routes at specified check points. These results were specific to the Maryland recreational blue crab (*Callinectes sapidus*) dipnet fishery, but the model is general and can be applied to data from other fisheries. However, the bias demonstrated is inherent in the roving clerk design, and is caused by the interception probability shadowing of angling parties on one another.

Literature Cited

Cronin, L.E. (ed.) 1988. Report of the Chesapeake Bay blue crab management workshop. Chesapeake Bay Commission, MDNR, Potomac River Fisheries Commission, Virginia Marine Resources Commission, 68 p.

Robson, D.S., 1961. On the statistical theory of a roving creel census of fisherman. *Biometrics* 17, pp 415-437.

United States Fish and Wildlife Service, 1988. 1985 National survey of fishing, hunting, and wildlife associated recreation. 167 p.

Williams, J.B., H.J. Speir, S. Early, & T.P. Smith. 1982. 1979 Maryland Department of Natural Resources, Tidewater Administration TA-CRD-82-1, 100 p.

List of Figures

Estimation Module

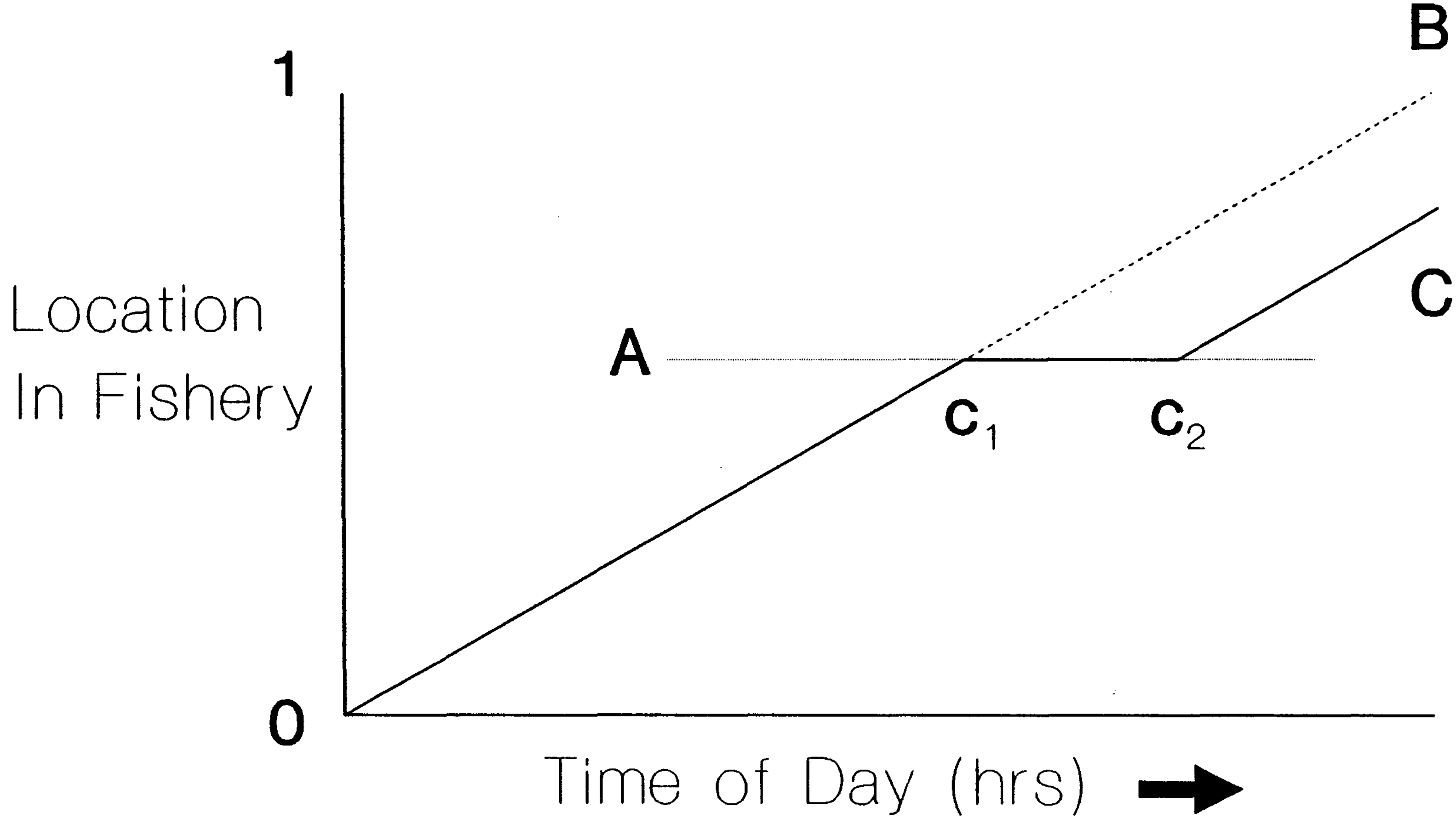

FIGURE 1

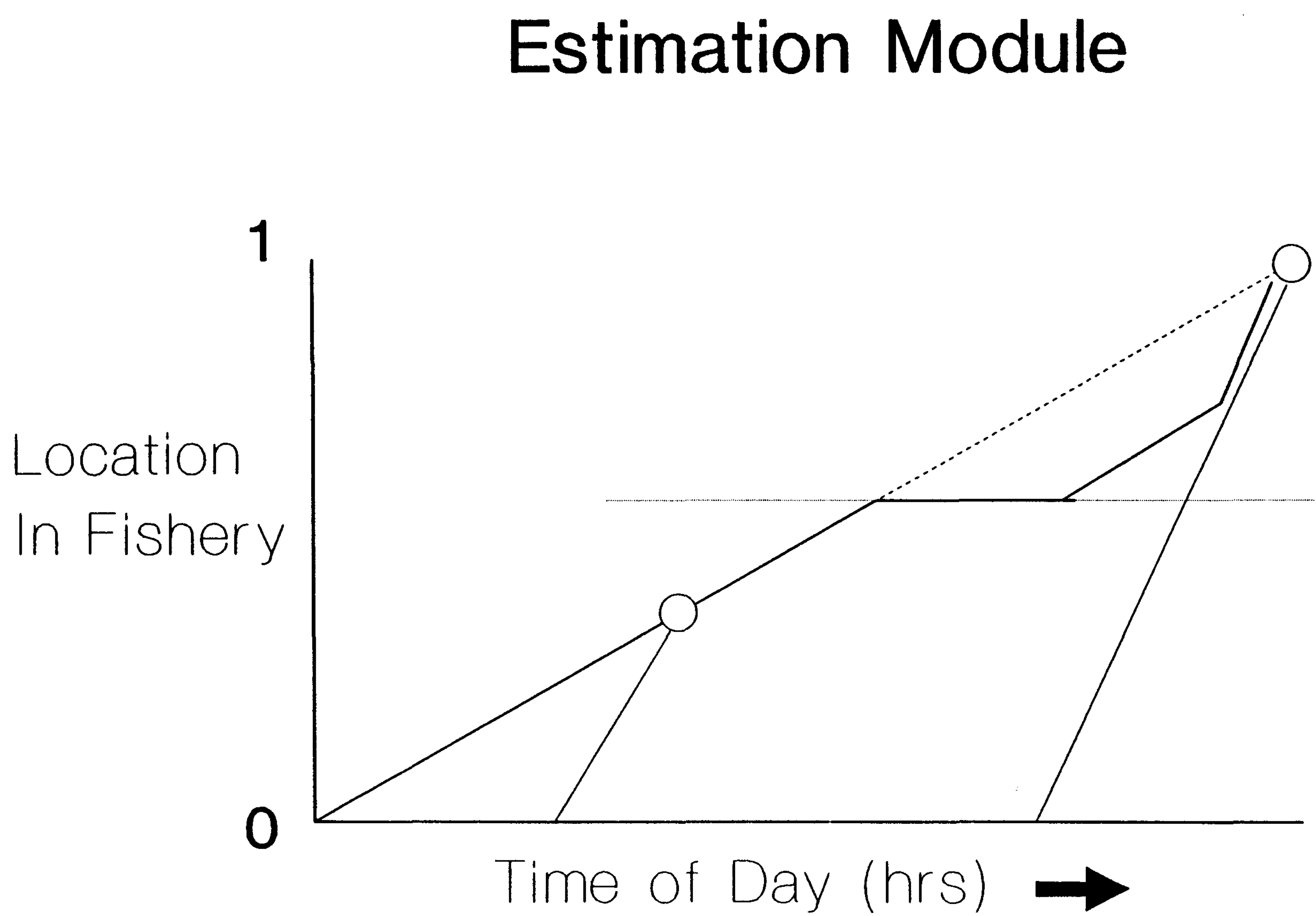

FIGURE 2

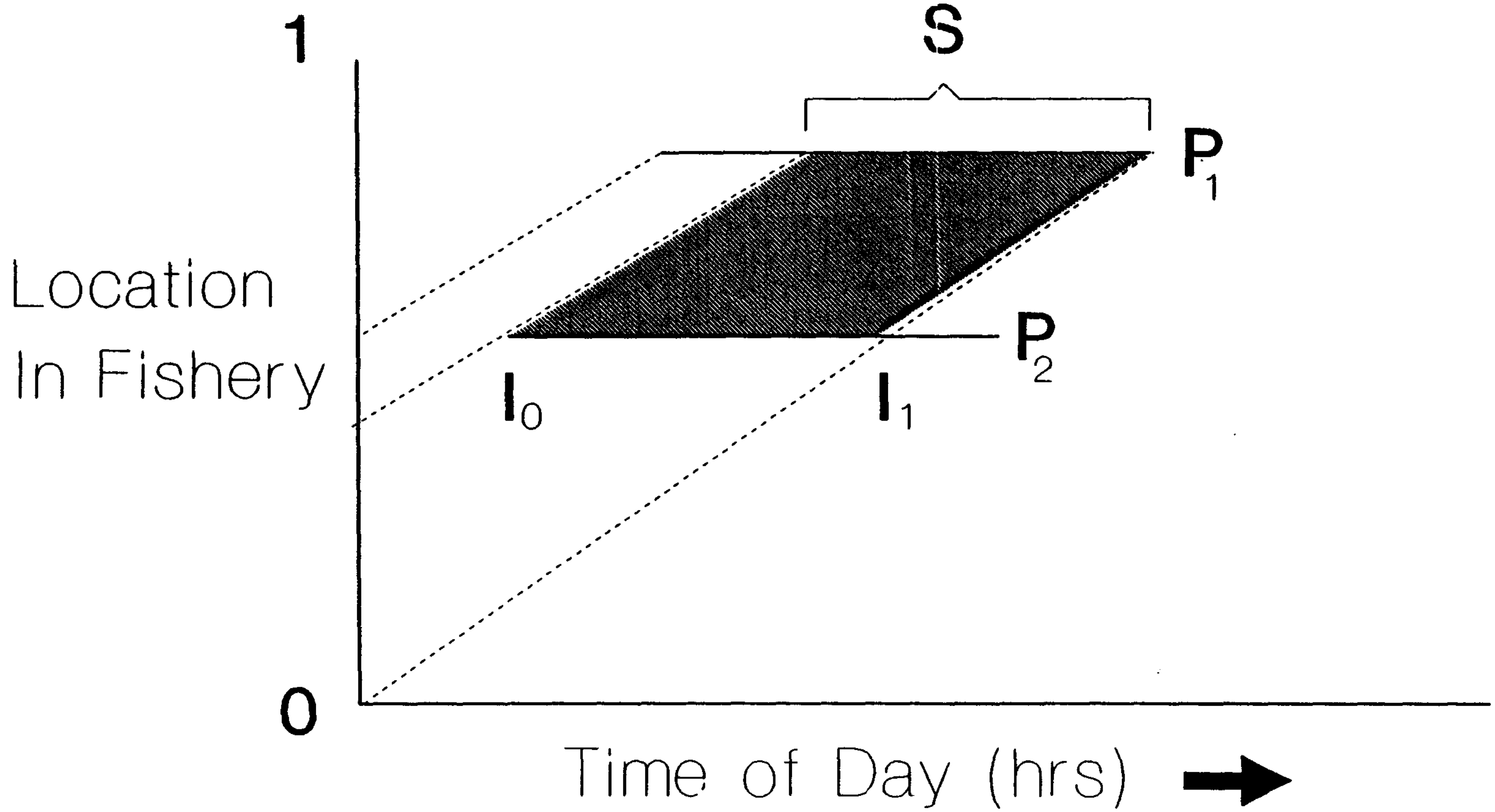

FIGURE 3

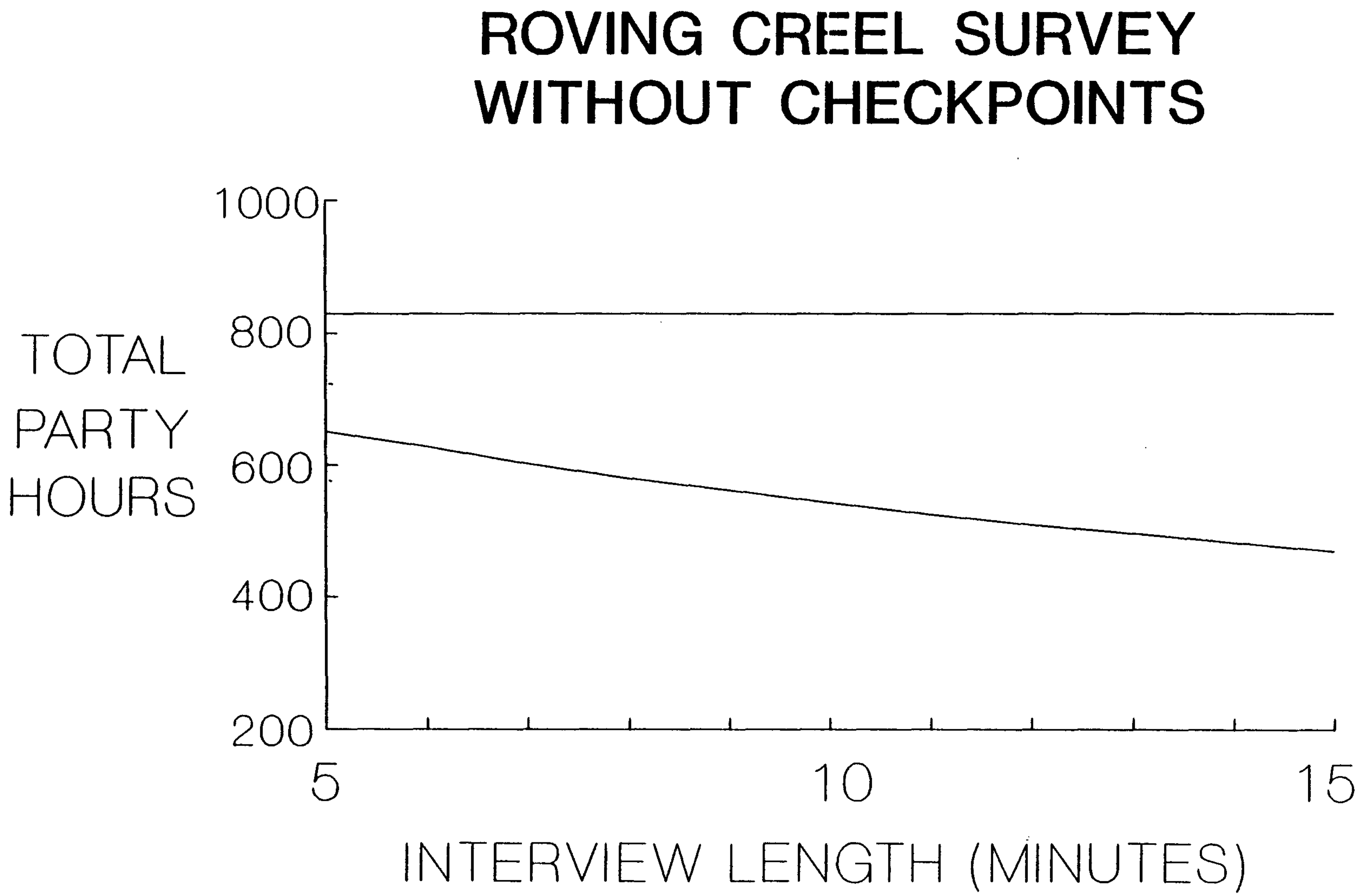
ROVING CREEL SURVEY
WITHOUT CHECKPOINTS
TOTAL
PARTY
HOURS
1000
800
600
400
200
5
10
15
INTERVIEW LENGTH (MINUTES)

FIGURE 4

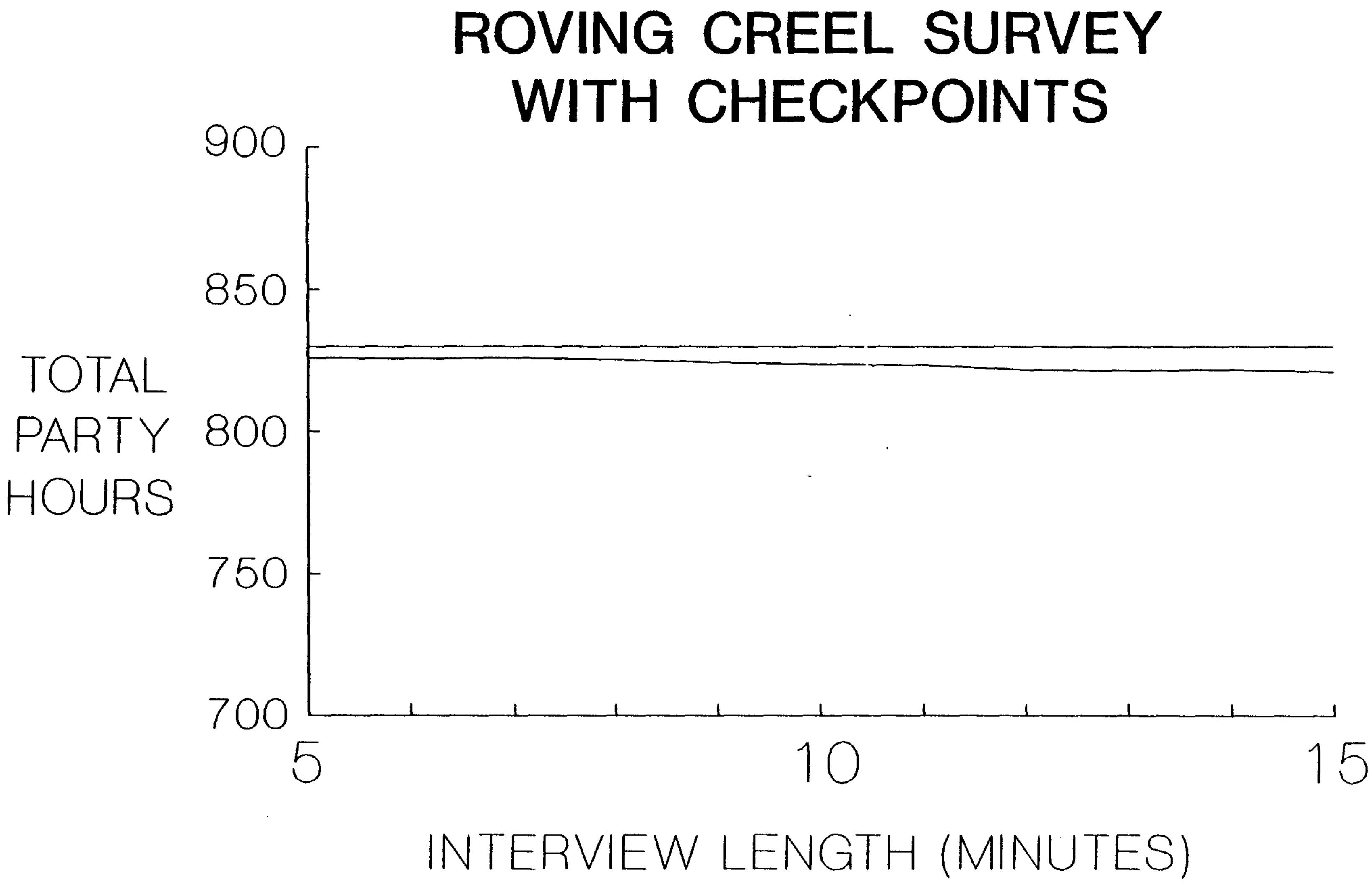
ROVING CREEL SURVEY
WITH CHECKPOINTS
TOTAL
PARTY
HOURS
900
850
800
750
700
5
10
15
INTERVIEW LENGTH (MINUTES)

FIGURE 5

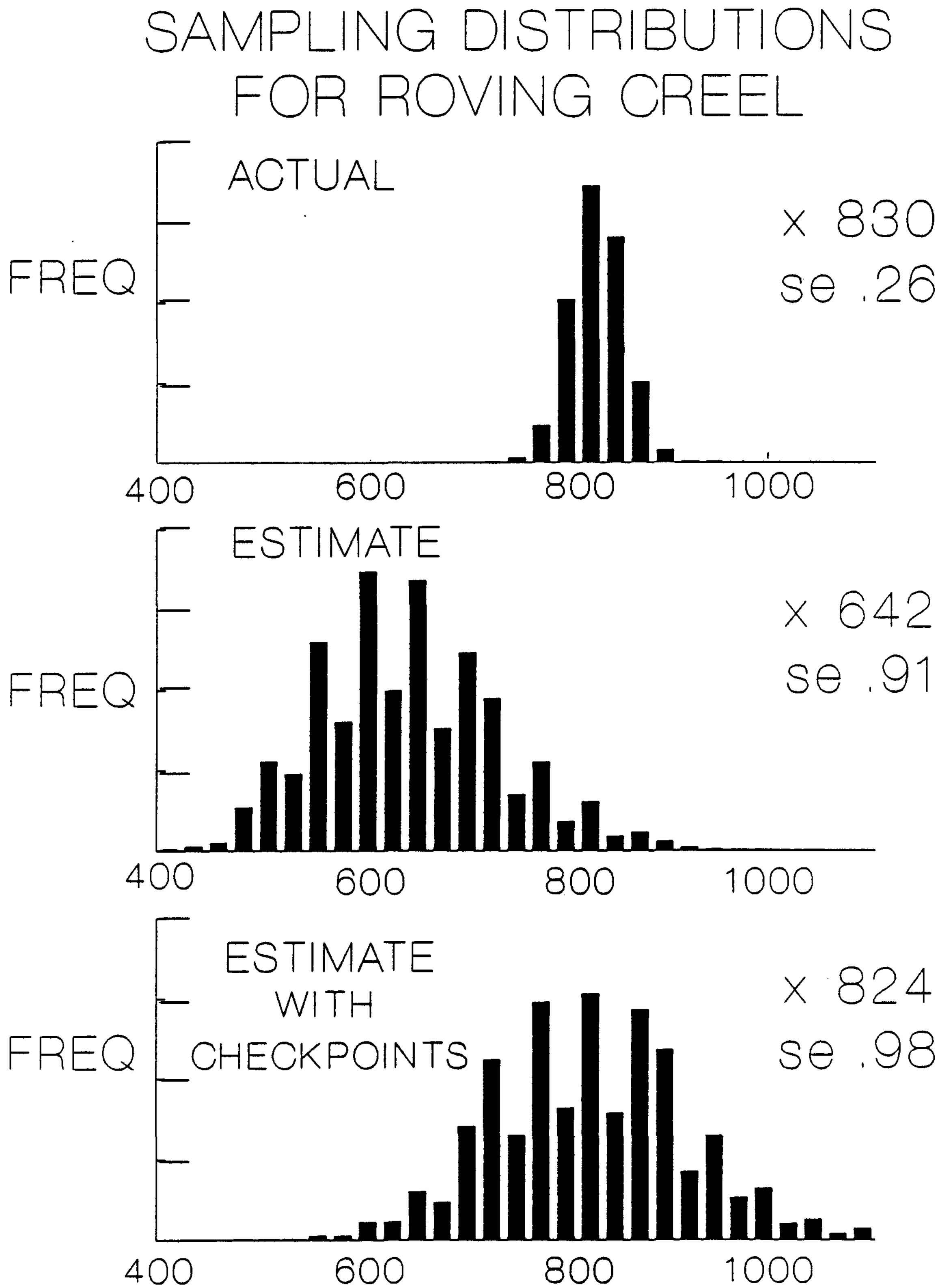
SAMPLING DISTRIBUTIONS
FOR ROVING CREEL
ACTUAL
x 830
se .26
FREQ
400
600
800
1000
ESTIMATE
x 642
se .91
FREQ
400
600
800
1000
ESTIMATE
WITH
CHECKPOINTS
x 824
se .98
FREQ
400
600
800
1000

FIGURE 6