An Operational Characterization of Stock Size Estimation of Migratory Species

by

Jose A. Goulart

Department of Mathematics
University of California
Davis, California 95616

An Operational Characterization of
Stock Size Estimation
of Migratory Species

by

Jose A. Goulart

1986

Report 86-2 on Sea Grant Project 6-F-N
"New Methods in Stock Abundance Estimation"

California Sea Grant
College Program, R/F-109

AN OPERATIONAL CHARACTERIZATION OF STOCK SIZE ESTIMATION OF MIGRATORY SPECIES

José A. Goulart
Department of Mathematics
University of California
Davis, CA 95616

Abstract

Two applications of a theory for the estimation of stock abundance using search effort are presented, based on an operational definition of the search parameter. Abundance estimates are obtained for South Pacific tuna data sets.

1. Introduction

During the last forty years different modeling strategies have led to the establishment of many stock assessment techniques in response to one of the dominant challenges posed to fisheries research which is, unquestionably, the accurate determination of stock abundance. Recently, Mangel and Beder (1985) developed a method which incorporates the time of search as an autonomous parameter in the characterization of fishing effort. This feature assumes particular relevance in the modeling of fisheries' operations of highly migratory species (i.e. tuna) since a considerable fraction of the total fishing effort is allocated to the search process.

In this paper, the maximum likelihood methods used to estimate the stock size, N, rely on the knowledge of the search parameter, ε , and the assumption that random search is a justifiable analogue for the process of school detection. Section 2 presents the operational definition of ε and some considerations about the theoretical background of the basic model.

In Section 3 we apply the theory and method proposed by Mangel and Beder (1985) to one vessel performing random search for patches of schools within a single-cell fishing area in one fishing period.

Section 4 contains an extension of the method to the search conducted in a two-cell area, during a two-week fishing season, given the migratory pattern of the fish schools.

The results, which include abundance estimates calculated from South Pacific tuna data sets and from a computer simulation of the extended method are exhibited in Section 5.

In Section 6, the concluding remarks, which include an analysis of the results, suggest the importance of investigating possible generalizations of the model, covering added levels of realism, to account for recruitment, the uncertainty associated with migration and a multi-cell fishing region.

2. Operational Parametrization of the Model

A simple substitution of the search targets in the probabilistic model proposed by Mangel and Beder (1985) provides the necessary adaptation to describing random search for patches of schools.

This minor change is introduced in order to distinguish random search occurring between detections of patches of schools from deterministic search for individual schools which fishermen normally execute within patches.

The modeling assumptions remain unchanged, except that the cohesive search units are, now, patches of schools.

The operational definition of the search parameter, ε, is (Koopman 1980)

$$\varepsilon = \frac{W v}{A} \tag{1}$$

where W is the sweep width of the searcher, v is the speed of the vessel during search and A is the area searched.

Therefore, given the existence of an initial number of patches, N, within a prescribed area, with the operational definition of ε, the search model becomes:

$$\text{Prob\{ detection of another patch in } \Delta t \mid i \text{ have been detected \}} = (N-i)\varepsilon\,\Delta t$$

and (2)

$$\text{Prob\{ no patch is detected in } \Delta t \mid i \text{ have been detected \}} = 1 - (N-i)\varepsilon\,\Delta t$$

3. The Basic Method

From an initial number of N patches of schools, the detection of n patches by a searcher defines uniquely a set of search times (times between sucessive detections), $\{T_1, \ldots, T_n\}$, where T_i, $i = 1, \ldots, n$, is the time elapsed between the (i-1)st and the ith detection.

The basic method for the estimation of stock abundance in a single-cell fishing region is established by reformulating the model (2). According to Mangel and Beder (1985), the following probabilistic function equivalently describes the search model:

$$\text{Prob\{ one searcher detects n patches in } (0,t) \text{ \}} = \binom{N}{n}(1-e^{-\varepsilon t})^{n} e^{-\varepsilon t(N-n)} \quad , \quad (n=0,1,\ldots,N)$$

$$= 0, \text{ otherwise.} \qquad (4)$$

Since equation (4) is a likelihood function for N, (considering t fixed), and given that n patches of schools were encountered, by letting t = T the likelihood function L(N;n) becomes

$$L(N;n) = \binom{N}{n}(1-e^{-\varepsilon t})^{n} e^{-\varepsilon T(N-n)} \qquad , N = n, n+1, \ldots$$

$$= 0, \text{ otherwise.} \qquad (5)$$

The MLE for N is straightforwardly calculated from (5), using the likelihood ratio, and the solution is

$$\hat{N} = \left[\frac{n}{1 - e^{-\varepsilon T}} \right] \qquad (6)$$

where [.] symbolizes the integer part of the expression within brackets.

The basic estimation method is, therefore, described by equation (6), given the number of patches, n, detected in time T and the operational value of ε.

4. Extended Method for Migratory Species

In a region $\mathcal{R}$, consisting of two adjacent and identical cells, each with area A, designated by $\mathcal{R}_1$ and $\mathcal{R}_2$, the migratory pattern of the patches of schools is defined such that the fish move first into $\mathcal{R}_1$, then across to $\mathcal{R}_2$ and, finally, leave $\mathcal{R}$ from $\mathcal{R}_2$. It is also assumed that when the fish are in $\mathcal{R}$, the whole stock is in either $\mathcal{R}_1$ or $\mathcal{R}_2$.

The search strategy is the following: Given the initial number of patches of schools, N, which migrate through $\mathcal{R}$, a single searcher conducts random search in $\mathcal{R}_1$, starting in the first day of the season. After k_s successive days of no detection the searcher leaves $\mathcal{R}_1$ and starts searching in $\mathcal{R}_2$ the following day, where it stays until the last day of the season. Consequently, at the end of the season, catch data and search effort are known for each day of the season. The search parameter is obtained from operational variables, as indicated in Section 2.

For a season with length of K days let there be m_s days of successful detections corresponding to m_1 days in $\mathcal{R}_1$ and $(m_s - m_1)$ days in $\mathcal{R}_2$, $(0 \leq m_1 \leq m_s < K)$.

If one lets m_2 be the number of days of unsuccessful search in $\mathcal{R}_1$ it follows that $(K - m_s - m_2)$ represents the number of days without detections in $\mathcal{R}_2$, $(0 \leq m_2 \leq K - m_s)$.

Also, let

S_{i,l_j} = time observed between the $(i-1)$st and the ith detection in day l_j, $(i = 1,\ldots,n_j)$

where $\{l_j\}$, $(j = 1,\ldots,m_s)$ is the set of days with successful detections (n_j represents the total number of detections recorded in day l_j). Note that the "0th" detection is defined as the beginning of the search period.

If a search leads to no detection in day k_j, $(j = 1,\ldots,K)$, let S_{k_j} represent the corresponding search time.

To model the uncertainty about the existence of fish in the cell where the vessel is searching, it is necessary to know, for each cell, the probability that the stock is present for each day of the season. To this effect, let

$p_i(t, k)$ = Prob(stock is present in $\mathcal{R}_i$ on day $(t + k)$ | present in $\mathcal{R}_i$ on day t)

such that $p_i(t,0) = 1$ and $p_i(t,k) \sim \bar{p}_i(t + k)$, where

$\bar{p}_i(t)$ = Prob(stock is present in $\mathcal{R}_i$ on day t), $i = 1,2$.

Therefore, given an initial number of patches of schools, N, the application of the random search formula (Mangel 1985) leads to the following search model, which includes depletion of the stock:

$$\text{Prob(a detection in time } S_{i,l_j}) = 1 - e^{-N(l_{j-1};i)\varepsilon S_{i,l_j}}$$

and

Prob{no detection in time S_{ℓ} | last detection occurred on day l_j } =

$$= [1 - \overline{p}_i\ (l_j)] + \overline{p}_i(l_j)\ e^{-N(l_j)\varepsilon S_{\ell}}\ , \tag{7}$$

where $\overline{p}_i$ (l) = Prob{stock is present in $\mathcal{R}_i$ on day l},

$N(l_{j-1} ; \nu)$ = number of undetected patches after the (ν - 1)st detection on day l_j , (ν = 1,...,n_j), and

$N(l_j)$ = number of undetected patches at the end of day l_j .

It follows from (7) that if n_j patches were detected on day l_j , the likelihood of the corresponding set of search times, $\{S_{i,l_j}, l_j\} \cup \{S_{l_j}\}$, is

$$\mathcal{L}\,(n_j\ ; l_j\ ) = \{\prod_{i=1}^{n_j} [N(l_{j-1}\ ; i) - i + 1]\ \varepsilon\ e^{-[N(l_{j-1};i)-i+1]\varepsilon S_{i,l_j}}\}.\ e^{-N(l_j)\varepsilon S_{l_j}}$$

Where S_{l_j} = 0 if there is a detection in the last search period of day l_j .

Similarly, if no patch was detected on day k_j , and the last detection was observed on day l_j , the likelihood of the set $\{S_{k_j}\}$ is

$$\mathcal{L}\,(k_j\ ) = 1 - \overline{p}_i\ (k_j\ ) + \overline{p}_i\ (k_j\ )e^{-N(l_j)\varepsilon\ S_{k_j}}$$

Therefore, the likelihood of the set of search times for the whole season is

$$\mathcal{L} = \mathcal{L}_1 \cdot \mathcal{L}_2\ . \tag{8}$$

where

$$\mathcal{L}_1 = \prod_{j=1}^{m_s} \left\{ \prod_{i=0}^{n_j-1} \left[\left(N-i-\sum_{l=0}^{j-1} n_l\right) \varepsilon\, e^{-\left(N-i-\sum_{l=0}^{j-1} n_l\right)\varepsilon S_{i+1,l_j}} \right] e^{-\left(N-\sum_{l=0}^{j} n_l\right)\varepsilon S_{l_j}} \right\} \tag{8.a}$$

and

$$\mathcal{L}_2 = \prod_{j=1}^{m_2} \left[1-\bar{p}_1(k_j)+\bar{p}_1(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}} \right] \cdot \prod_{j=m_2+1}^{K-m_s} \left[1-\bar{p}_2(k_j)+\bar{p}_2(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}} \right] , \tag{8.b}$$

where $l^*(j)$ = max i, such that $l_i < k_j$ $(i = 1,\ldots,m_s)$.

The application of the log-likelihood method to $\mathcal{L}$ (Appendix I), yields the following algebraic equation for the MLE for N:

$$\sum_{j=1}^{m_s} \left[\sum_{i=0}^{n_j-1} \left(\frac{1}{N-i-\sum_{l=0}^{j-1} n_l} - \varepsilon S_{i+1,l_j} \right) - \varepsilon S_{l_j} \right] = \sum_{j=1} F_1(N;j) + \sum_{j=m_2+1} F_2(N;j) \tag{9}$$

with

$$F_1(N;j) = \frac{\bar{p}_1(k_j)\varepsilon S_{k_j} e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}}}{1-\bar{p}_1(k_j)+\bar{p}_1(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}}}$$

and

$$F_2(N;j) = \frac{\bar{p}_2(k_j)\varepsilon S_{k_j} e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}}}{1-\bar{p}_2(k_j)+\bar{p}_2(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right)\varepsilon S_{k_j}}}$$

We observe that a solution of (9) is an MLE, (Appendix I) only if

$$X_0(N;i,j) > X_1(N;i,j) + X_2(N;i,j). \tag{10}$$

Note that, as expected, if $\bar{p}_i(k_j) = 1$, $j = 1,\ldots,(k - m_s)$ and $i = 1,2$ then $X_1(N;i,j) = X_2(N;i,j) = 0$ and any solution of (9) will be an MLE. Thus, the extended method, defined by the algebraic equation (9), is only valid if (10) is satisfied.

Abundance Estimation of South Pacific Tuna

A set of tuna catch data from the South Pacific, collected in December of 1979 by a scientific vessel in a tagging expedition contains, essentially, the chronological times of school detections, the coordinates of all sightings and the number of fish handled after each school encounter.

In order to apply the basic method developed in Section 3, the total search time, T, spent on a random search for distinct patches of schools within a region $\mathcal{R}$ with area A, must be extracted from the data since no explicit form of search data was available.

It also requires the evaluation of the search parameter ε, through an adequate choice of the operational variables W, v, and A, in accordance with equation (1).

The calculation of the total search time spent in $\mathcal{R}$ rests on the knowledge of the individual search times associated with every detection observed in $\mathcal{R}$. In other words, if n cohesive patches of schools were detected in $\mathcal{R}$, the set of search times $\{T_1, \ldots, T_n\}$ must be determined, where T_i is the time the vessel spends searching between the (i - 1)st and the ith school encounter. Therefore, the first problem to solve was the identification of the number of distinct patches detected in $\mathcal{R}$, using the information available in the data sets. Since the data refer to school detections, the first step was to establish a criterion to differentiate random search between patches from deterministic search for schools within a patch.

The following rule was adopted through an empirical selection of meaningful temporal and spatial threshold values, estimated from a careful analysis of the vessel's tracks and the coordinates of the detection sites:

> Two schools, discovered consecutively, are considered to belong to the same patch and, consequently, non-randomly searched if a) their separation was less than 2 n.miles or, b) their separation was less than 3 n.miles and the time elapsed between detections did not exceed 26 minutes.

This rule not only allows the counting of the random encounters in $\mathcal{R}$ but also provides the basis for the determination of the set of individual search times. In fact, since the chronological time for each sighting of a school is known, by excluding the search times corresponding to non-random search and also subtracting the set times, the result will be the value of a search time between the detections of two distinct patches, randomly searched. Namely, given t_i and t_{i-1} to be the chronological times logged, respectively, for the ith and (i - 1)st detection of two independent patches, by discounting the set time associated with the fish handling operations which took place immediately after the ith detection, the value of T will be found.

The data did not contain any information about set times, and the literature researched (Pella and Psaropulos 1975) did not cover the subject of set times for tagging operations. It became, therefore, necessary to extract from the data an approximation of the order of magnitude for the set times corresponding to different categories of quantities of fish handled. Henceforth, a certain quantity of fish handled will be designated by a volume and the objective is, thus, to devise a functional relationship between the set times and the volume of fish.

In the four data sets considered, the volumes ranged from 1 to 989 individuals and a more detailed analysis of the data revealed, for instance, that the set time corresponding to a large volume was unexpectedly short. Indeed, to a volume of 650 individuals there corresponds a set time which could not have exceeded 10 minutes, since the time elapsed between the two successive patch encounters was 40 minutes and in order to travel the respective distance ($\sim$ 5.4 n.miles) the vessel would have taken, at least, 30 minutes.

Verifying that only two volumes of fish exceeded 650 individuals (by less than 50%), three categories of volumes were defined:

Small (S) : $1 \leq V \leq 350$

Medium (M) : $350 < V \leq 650$

Large (L) : $650 < V$

where V stands for number of fish.

Then the empirical functional rule established for V and the set times is indicated below:

Table 1

V	Set Time (min)
S	5
M	10
L	15

The four data sets analyzed are designated by Set 1, Set 2, Set 3 and Set 4, corresponding to the original references MAQ2-79-12-22, MAQ2-79-12-23, MAQ2-79-12-24 and MAQ2-79-12-25, respectively.

Each data set covers one fishing day and refers to a distinctly searched region. The vessel started to search for schools at sunrise (5:45) and ended searching at sunset (18:30). Using the criterion adopted for the identification of distinct patches of schools, linear track lines were plotted (Figs. 1, 2, 3 and 4) and search times between detections, T_i, $(i = 1,\ldots,n)$, were calculated (Tables 2, 3, 4 and 5), as well as the value of s (search time of last search period leading to no detection) for each data. The average number (density) of schools per patch, d, and the average volume of individuals handled per school detected, $\bar{V}$, were evaluated and the results are presented in Table 6.

Set 1

Table 2

i	1	2	3	4	5	6	7	8	9
T (min.)	175	35	65	85	115	20	75	20	35

s = 15 (min.)

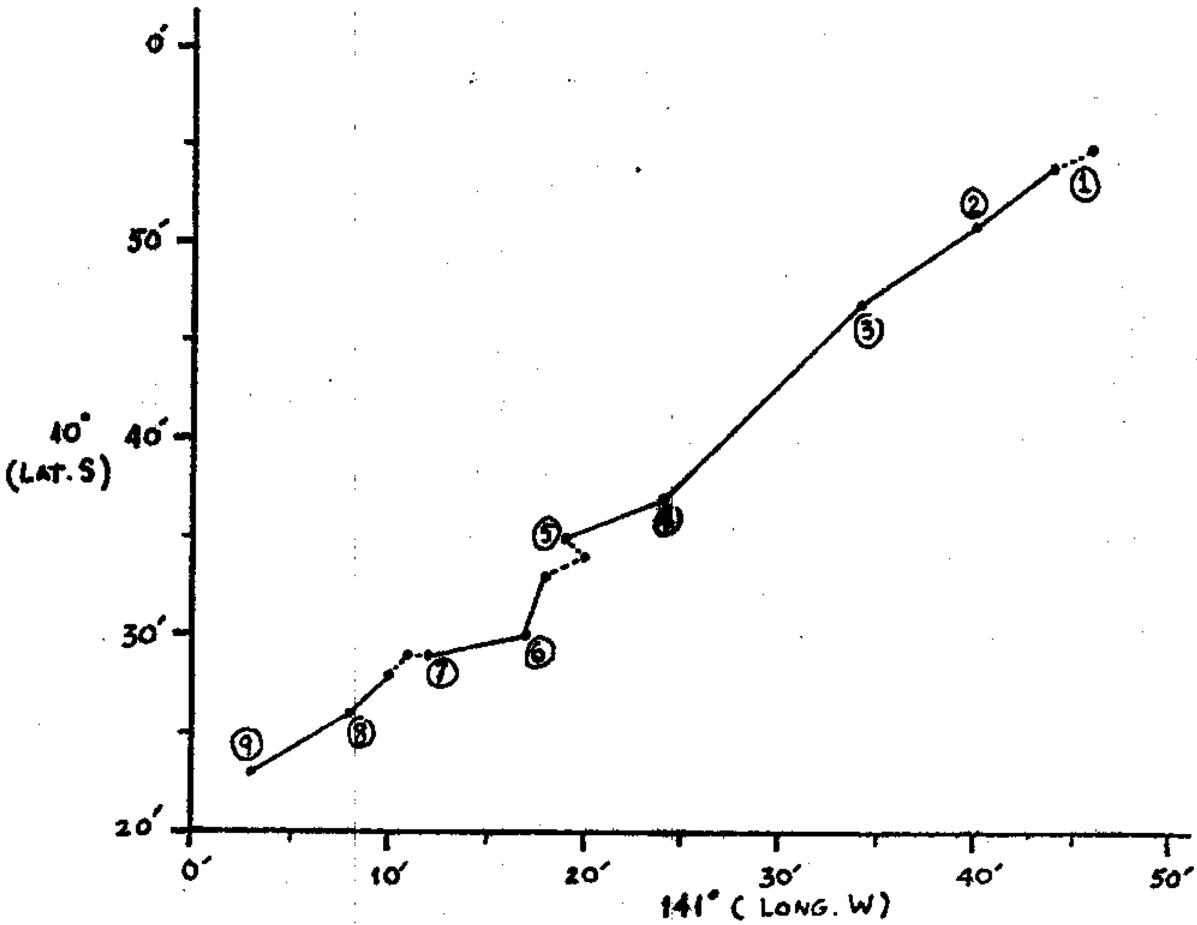


Fig. 1

Set 2

Table 3

i	1	2	3	4	5
T (min.)	10	50	35	60	25

s = 0 (min.)

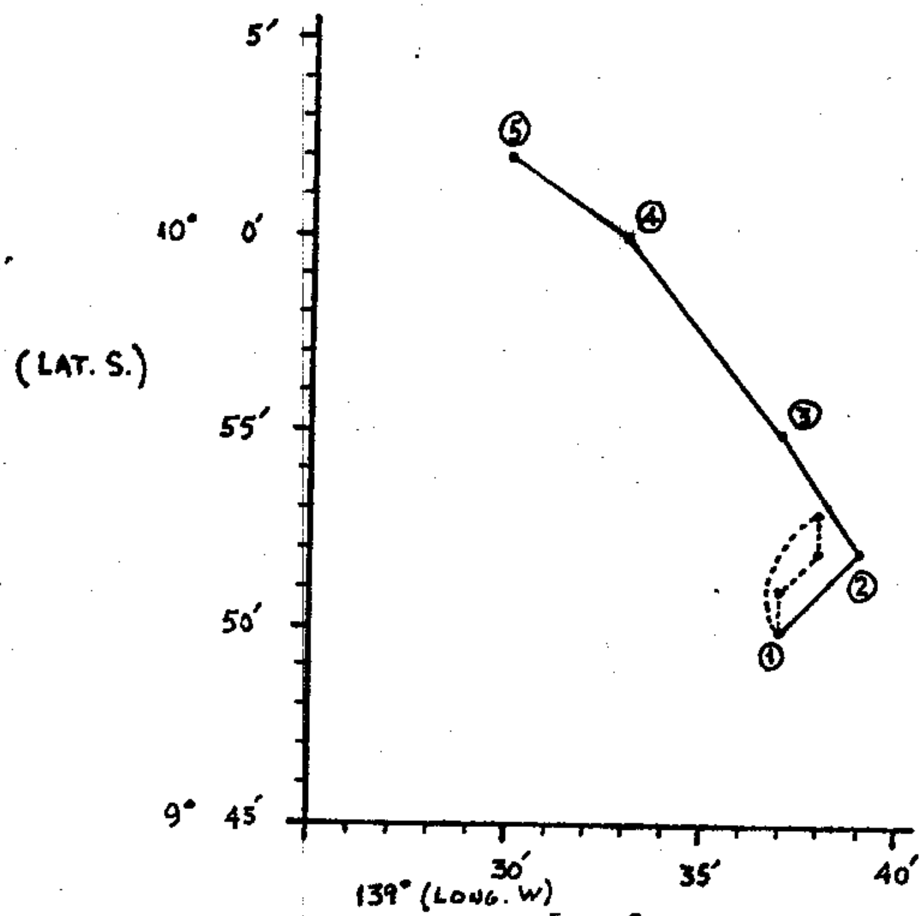


Fig. 2

Set 3

Table 4

i	1	2	3	4	5	6
T (min.)	135	90	40	35	230	45

s = 85

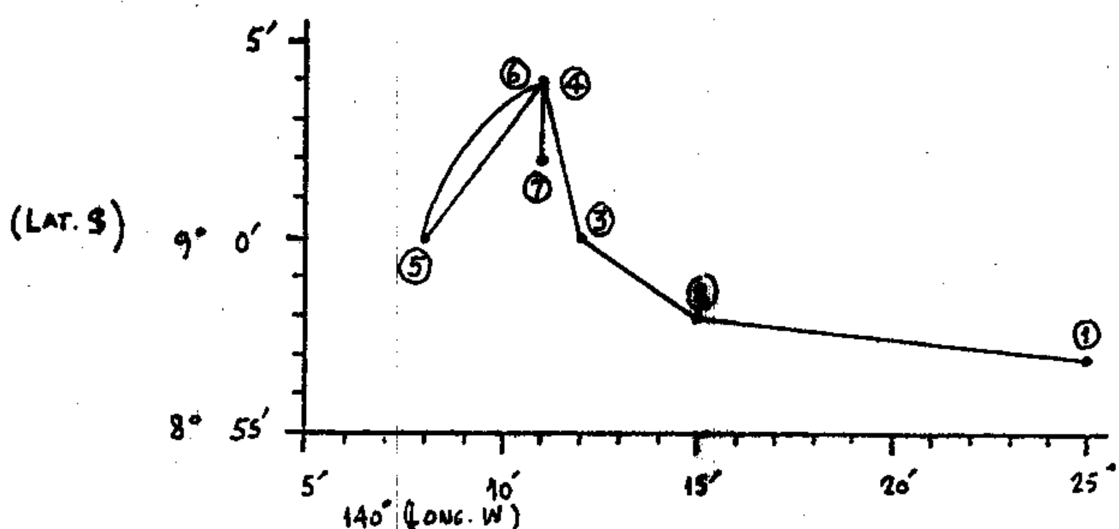


Fig. 3

Set 4

Table 5

i	1	2	3	4	5	6
T (min.)	135	145	25	125	30	20

s = 85 (min.)

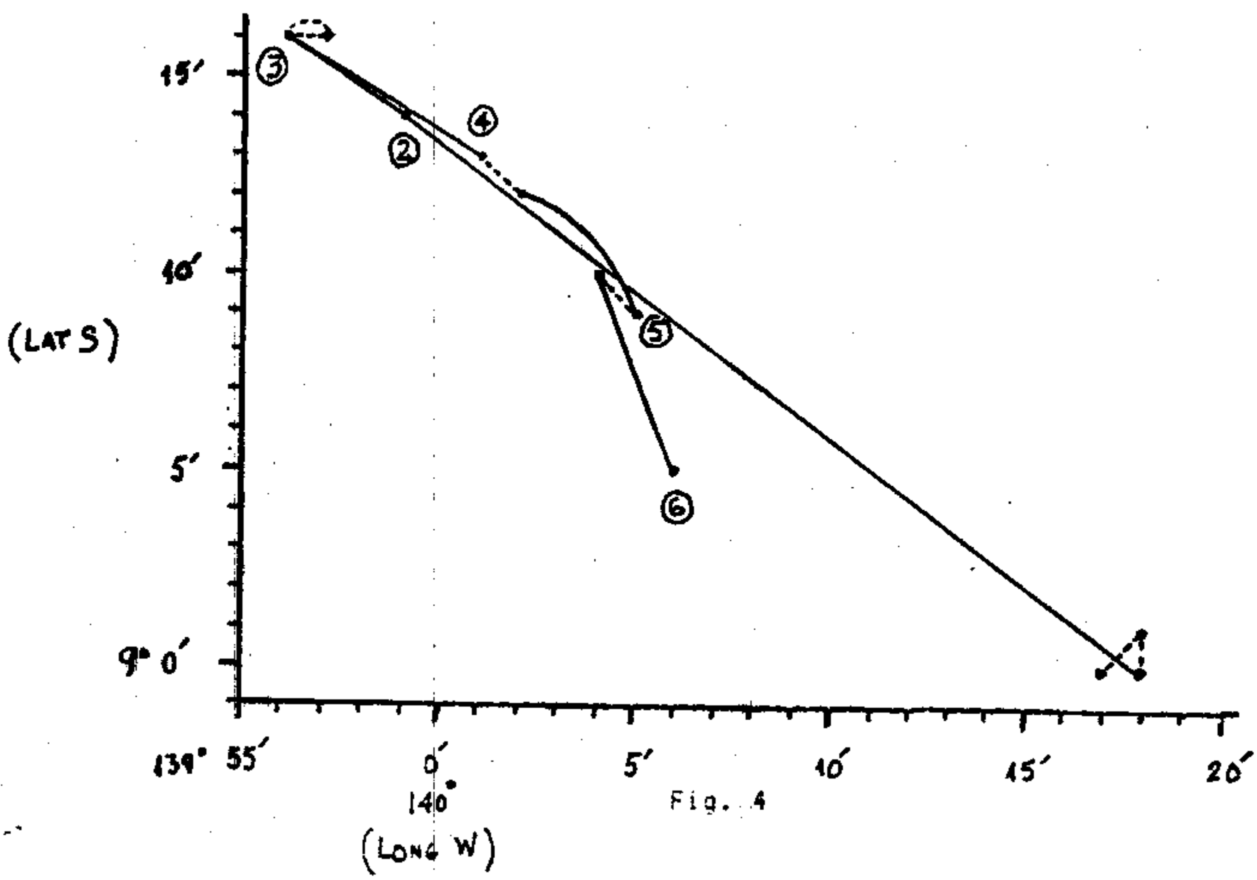


Fig. 4

Table 6

	Set 1	Set 2	Set 3	Set 4
n (Number of patches detected)	9	5	7	6
T(min) (total search time)	640	180	660	565
d (density of schools in patches)	1.6	1.8	1.17	2
V (average volume of fish per school)	28.6	147.5	232.6	182.9

Finally, in order to apply the basic method, the value of the search parameter was computed. The value of W = 15 n.miles was read from a table of sweep widths for visual search (National Search and Rescue Manual 1973), assuming a 20 n.mile meteorological visibility and a choice of target characteristics reasonably comparable to a tuna school (small ships; 500-1000 gross tons) at a height of 0 feet. The average search speed of the vessel, v, is 10 n.miles/hr (R. Hilborne, pers. comm.). The determination of the area of $\mathcal{R}$, A, for each day of fishing could not be obtained from the data. However, one observes that the maximum linearized search track length of the vessel for each data set considered was less than 84 n.miles, thus justifying an acceptable choice of $\mathcal{R}$ as a $1^\circ \times 1^\circ$ region (diagonal length $\sim$85 n.miles; A = 3600 sq.n.miles). Since the vessel conducts random search within an area enclosed by $\mathcal{R}$, this assumption will not invalidate the method.

The value of the search parameter corresponding to $W = 15$ n.miles, $v = 10$ n.miles/hr and $A = 3600$ sq.n.miles is, according to formula (1), $\varepsilon = .0007\ \mathrm{min}^{-1}$.

Proceeding to apply the values of n and T from Table 2, with $\varepsilon = .0007$, to formula (3), yields the following MLE estimates of the number of patches for each region $\mathcal{R}_1$, $\mathcal{R}_2$, $\mathcal{R}_3$ and $\mathcal{R}_4$, corresponding to the data sets Set 1, Set 2, Set 3 and Set 4, respectively:

Table 7

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
$\hat{N}$	24	42	18	18

To obtain the estimated stock, available to the fishery, $\hat{V}$, for each region, in terms of the numbers of individuals, let

$$\hat{V} = \hat{N} \cdot \bar{V} \cdot d \tag{16}$$

Where $\hat{N}$ is the MLE estimate for the number of patches, $\bar{V}$ represents the average number of fish per detected school and d is the density of schools in patches. The results are:

Table 8

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$
$\hat{V}$	1098	11,151	4,8[illegible]	[illegible],584

Stock Estimation for a Two-Cell Region with Migration

The method presented in Section 4 was implemented in a simulation program named GENMLE (see Appendix III) which contains two distinct parts, organized in the following sequence:

a) catch and search time data generation; and

b) The computation of the maximum likelihood estimate for N (using Newton's method), corresponding to the values of parameters and data generated in part a).

The data generation scheme presumed the input of the initial number of patches of schools, N, the search parameter, ε, (assumed to be operationally determined), the switching strategy, defined by the particular choice of k_s (i.e. the number of consecutive days with no detections which determine that the vessel switches from $\mathcal{R}_1$ to $\mathcal{R}_2$), the length of the fishing season, K, and the sets $\{\bar{p}_1(k)\}$ and $\{\bar{p}_2(k)\}$, containing the probability values about the presence of fish in $\mathcal{R}_1$ and $\mathcal{R}_2$, respectively, for k = 1,..., K (obtainable by time averaging the pertinent historical statistics about the fishery for the corresponding period).

It is assumed that, at the beginning of the season, the stock is in either $\mathcal{R}_1$ or $\mathcal{R}_2$. To decide in which cell the fish are on the first day of the season, an uniformly distributed random number, X, ($X \sim U[0,1]$), is drawn. If $X < \bar{p}_1(1)$ then the stock is in $\mathcal{R}_1$ on day 1; otherwise the fish are in $\mathcal{R}_2$ on day 1. After the fish "move" to $\mathcal{R}_2$, the same method is applied, such that if $X < \bar{p}_2(k)$

then the fish remain in $\mathcal{R}_2$ on day k, $(k = 2,\ldots,K)$; otherwise the remaining stock will have moved out of $\mathcal{R}$ and there is no available stock until the end of the season.

A similar scheme is followed to simulate detection of patches of schools. Letting $N(k)$ represent the remaining stock on day k and s_k be a value of search time, the probability, p_k, of a detection after the vessel searches for s_k units of time is given by $p_k = [1 - e^{-s_k \varepsilon N(k)}]$. Drawing an uniformly distributed random number in $[0,1]$, Y, if $Y < p_k$ then one patch is detected; otherwise there is no detection corresponding to the observed search time.

The daily period of searching was set constant (12 hrs) and the fishing season lasted 14 days.

The simulated data for the following choices of initial values of N, ε, and k_s, for each of five sets of $\{\bar{p}_1(k)\}$ and $\{\bar{p}_2(k)\}$, (henceforth designated by Probability Sets 1,..., 5) shown in Table 9 and pictorially represented in Figures 5 and 6, are determined:

a) $N = 50$, $N = 100$ and $N = 500$;

b) $\varepsilon = .001\ hr^{-1}$ and $\varepsilon = .005\ hr^{-1}$; and

c) $k_s = 1, 2, \ldots, 10$.

After a data set was generated the program proceeded to evaluate the corresponding MLE for N. For each value of N, ε, k_s and each probability set, catch and search time data were generated for fifty seasons, thus allowing the calculation of fifty values of $\hat{N}$ and, subsequently, the sample's average, $\hat{N}_{avg}$, and respective coefficient of variation (C.V.).

The corresponding results are contained in Figures 7 through 11, which give the values of $\hat{N}_{avg}$, as a function of ε, k_s, and each particular probability set.

Finally, for each choice of N, considering every probability set and the two values of ε, the following results are presented in Tables 10, 11 and 12:

i the value of $\hat{N}_{avg}$ which is the best approximation to N and the corresponding values of the coefficient of variation (C.V.) and k_s; and

ii the lowest value of the coefficient of variation (C.V.) (if different from the value shown in i)) and the corresponding values of $\hat{N}_{avg}$ and k_s.

Table 9

k (day)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\bar{p}_1(k)$	.9	.75	.7	.64	.58	.53	.475	.425	.37	.32	.26	.2	.15	.1
$\bar{p}_2(k)$	.1	.15	.2	.26	.32	.37	.425	.475	.53	.58	.64	.7	.75	.8

Prob. Set 1

$\bar{p}_1(k)$	1.	1.	1.	1.	1.	.8	.6	.4	.2	.0	.0	.0	.0	.0
$\bar{p}_2(k)$	.0	.0	.0	.0	.0	.2	.4	.6	.8	1.	1.	1.	1.	1.

Prob. Set 2

$\bar{p}_1(k)$	.9	.8	.7	.6	.6	.5	.5	.4	.3	.2	.1	.0	.0	.0
$\bar{p}_2(k)$	.1	.2	.3	.4	.4	.5	.5	.6	.6	.7	.7	.8	.7	.6

Prob. Set 3

$\bar{p}_1(k)$	.95	.94	.9	.85	.8	.75	.6	.2	.15	.1	.07	.05	.02	.01
$\bar{p}_2(k)$	.01	.02	.05	.07	.1	.15	.2	.6	.75	.8	.85	.9	.94	.95

Prob. Set 4

$\bar{p}_1(k)$	.5	.7	.8	.7	.7	.6	.6	.4	.3	.2	.1	.0	.0	.0
$\bar{p}_2(k)$	.0	.0	.0	.1	.2	.3	.4	.6	.6	.7	.7	.8	.7	.6

Prob. Set 5

PROBABILITY SETS

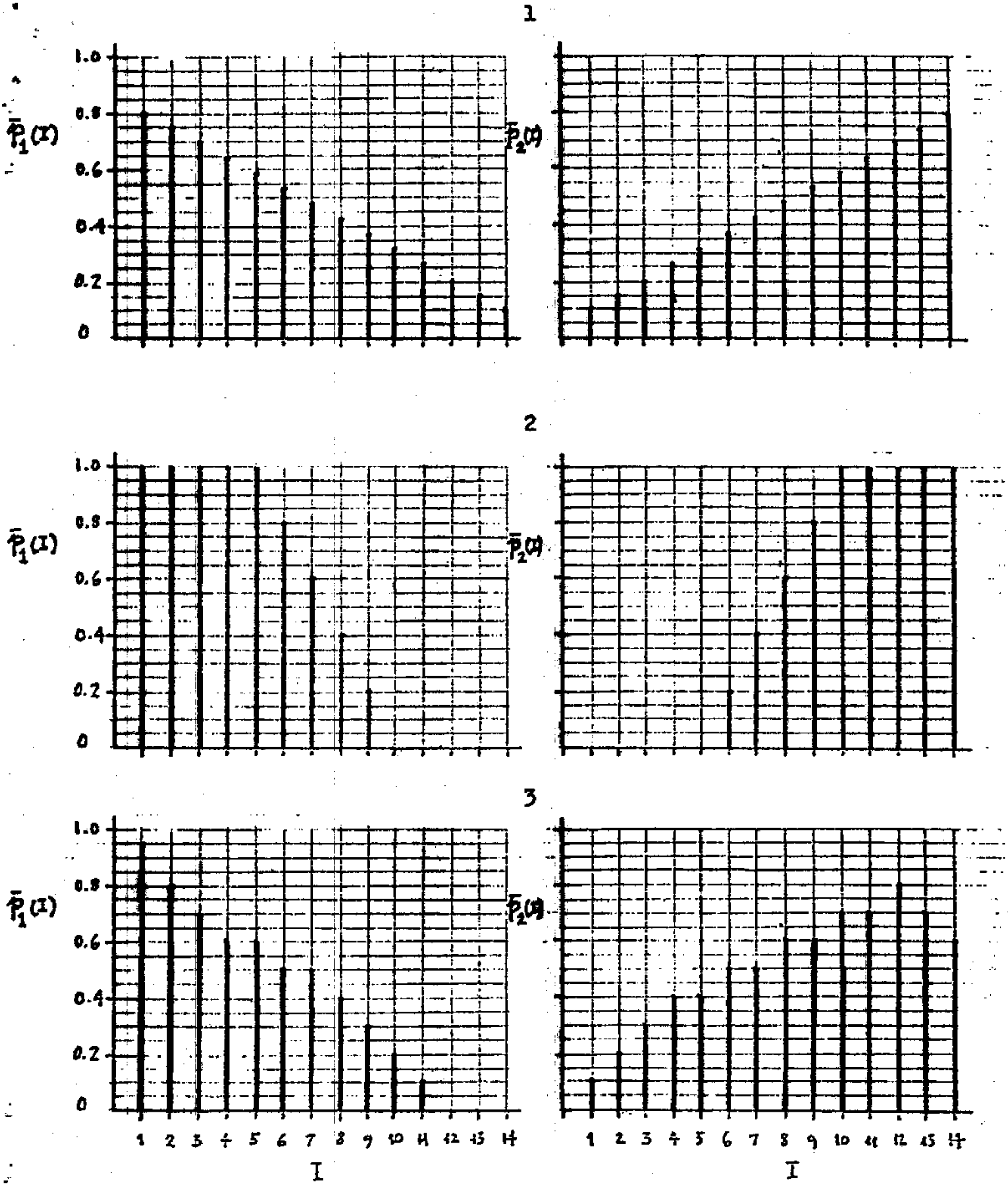


Figure 5

PROBABILITY SETS
(comt.)

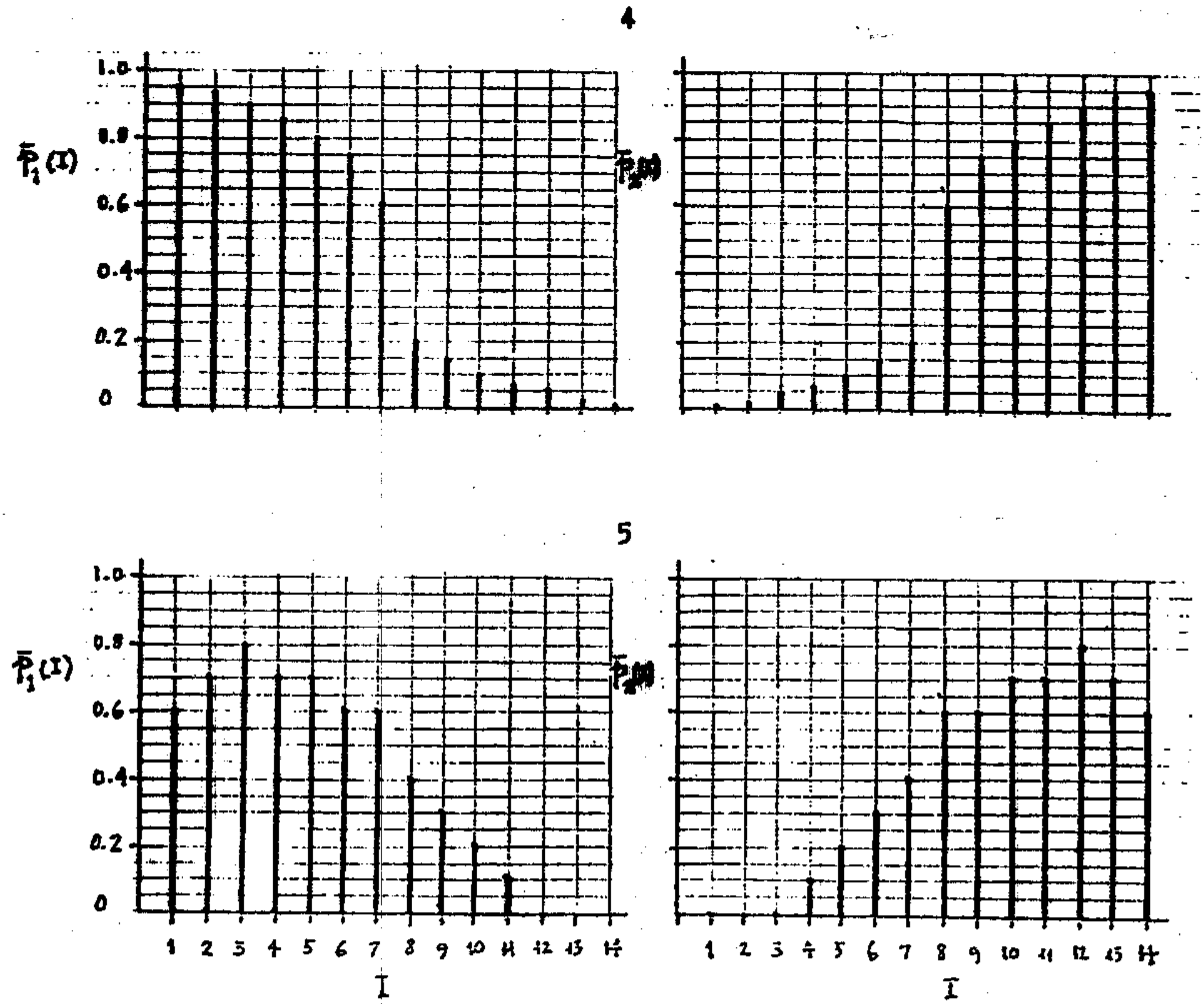
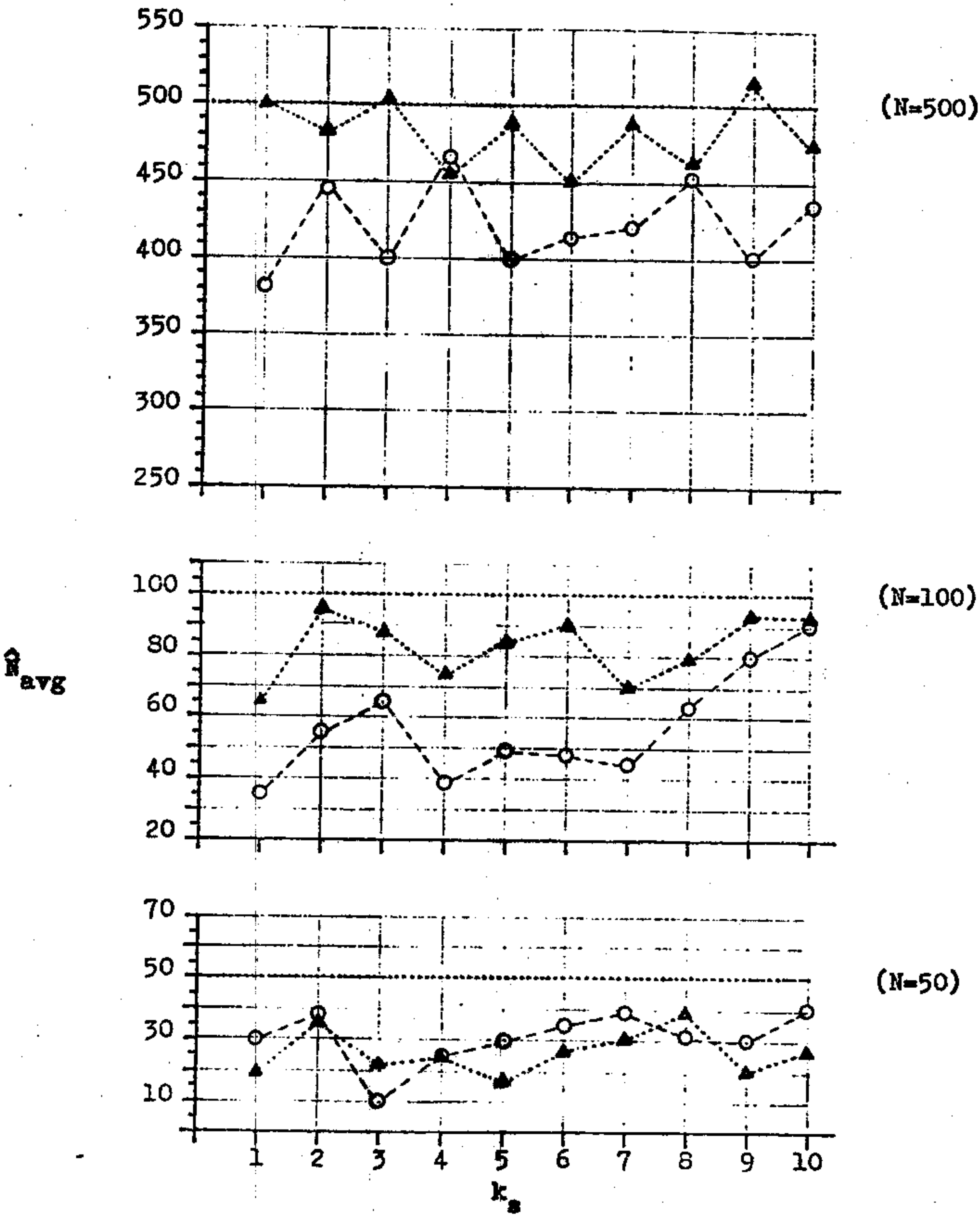
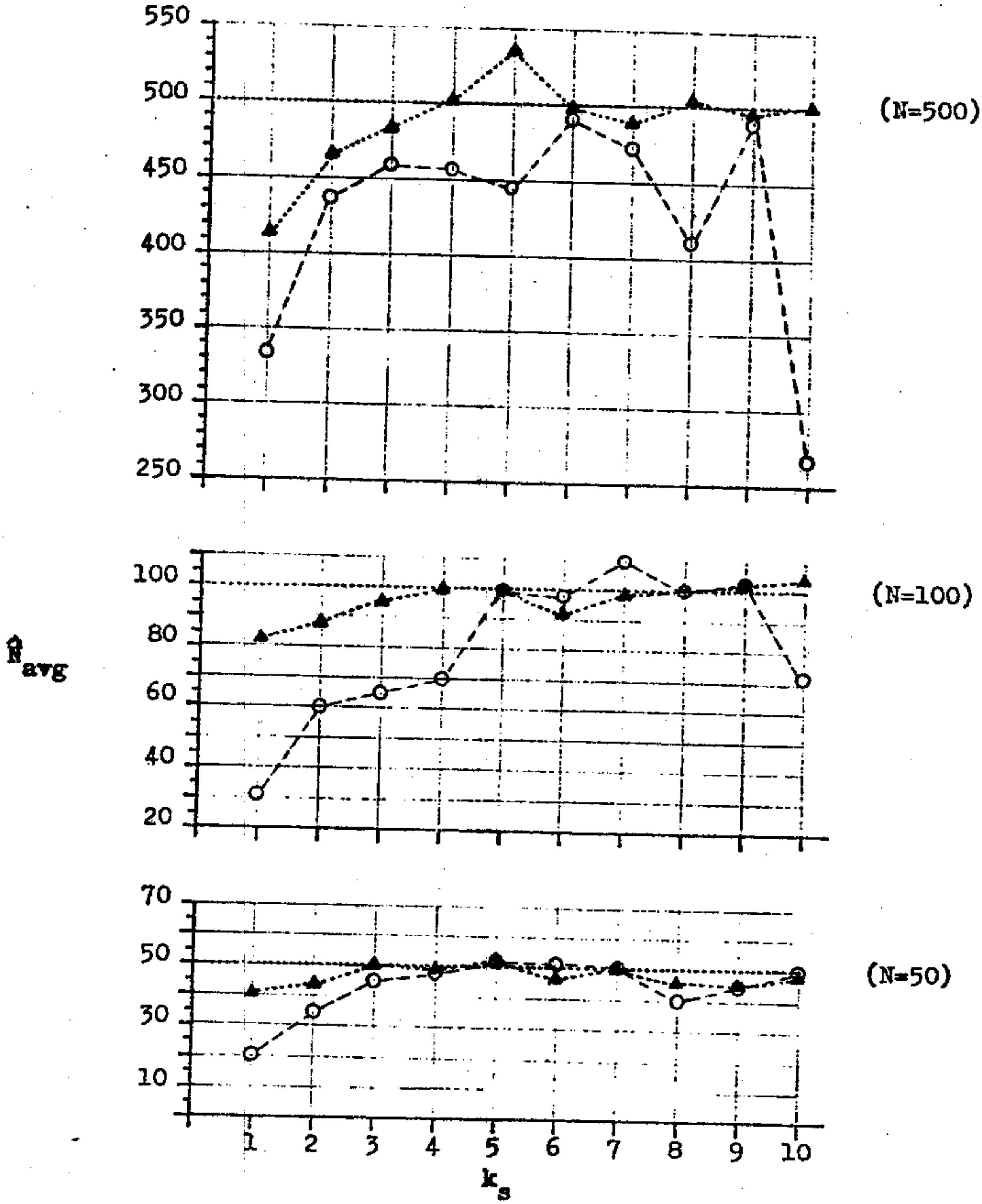


Figure 6



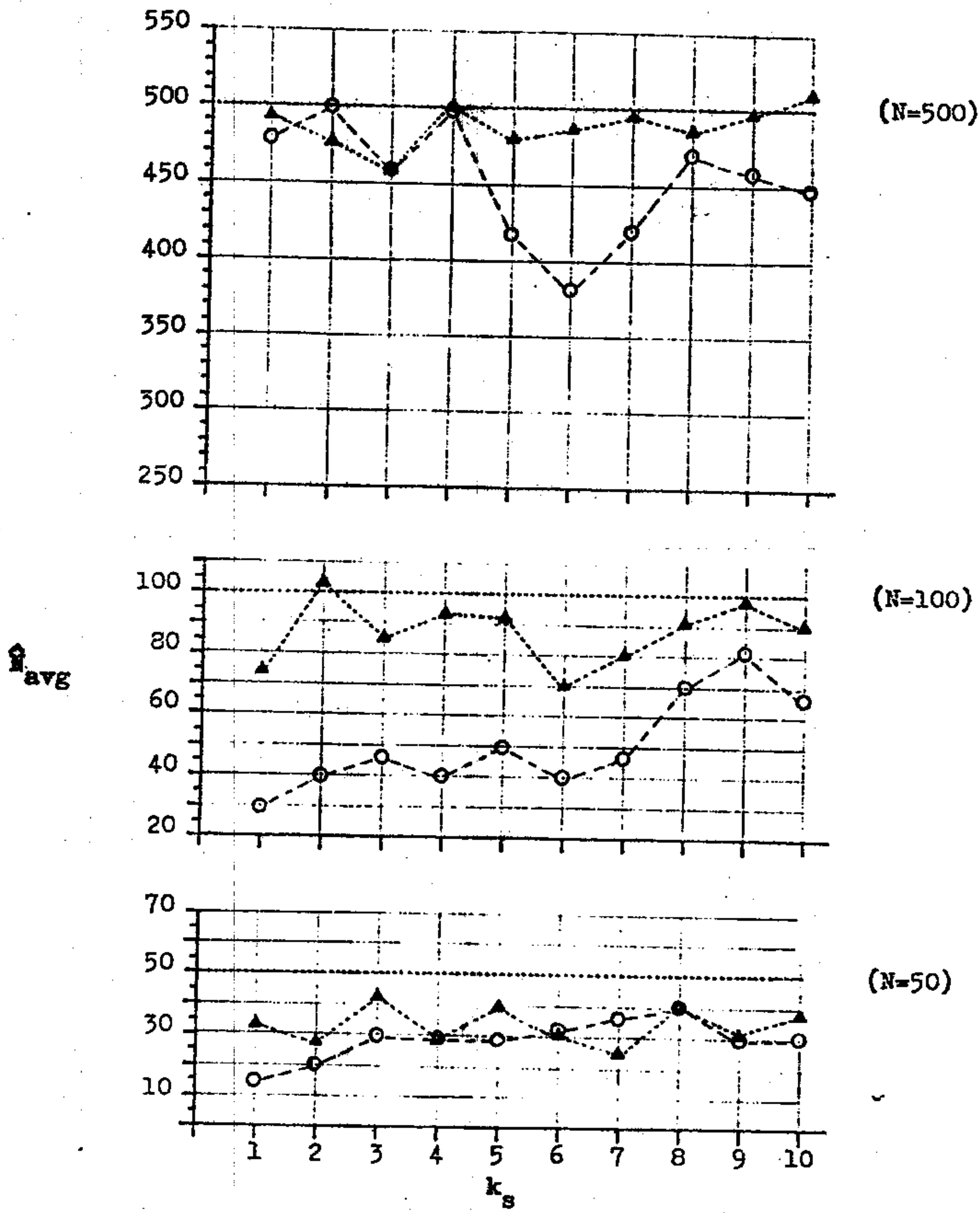
(Prob. Set 1: O : ε = .001, ▲ : ε = .005)

Figure 7



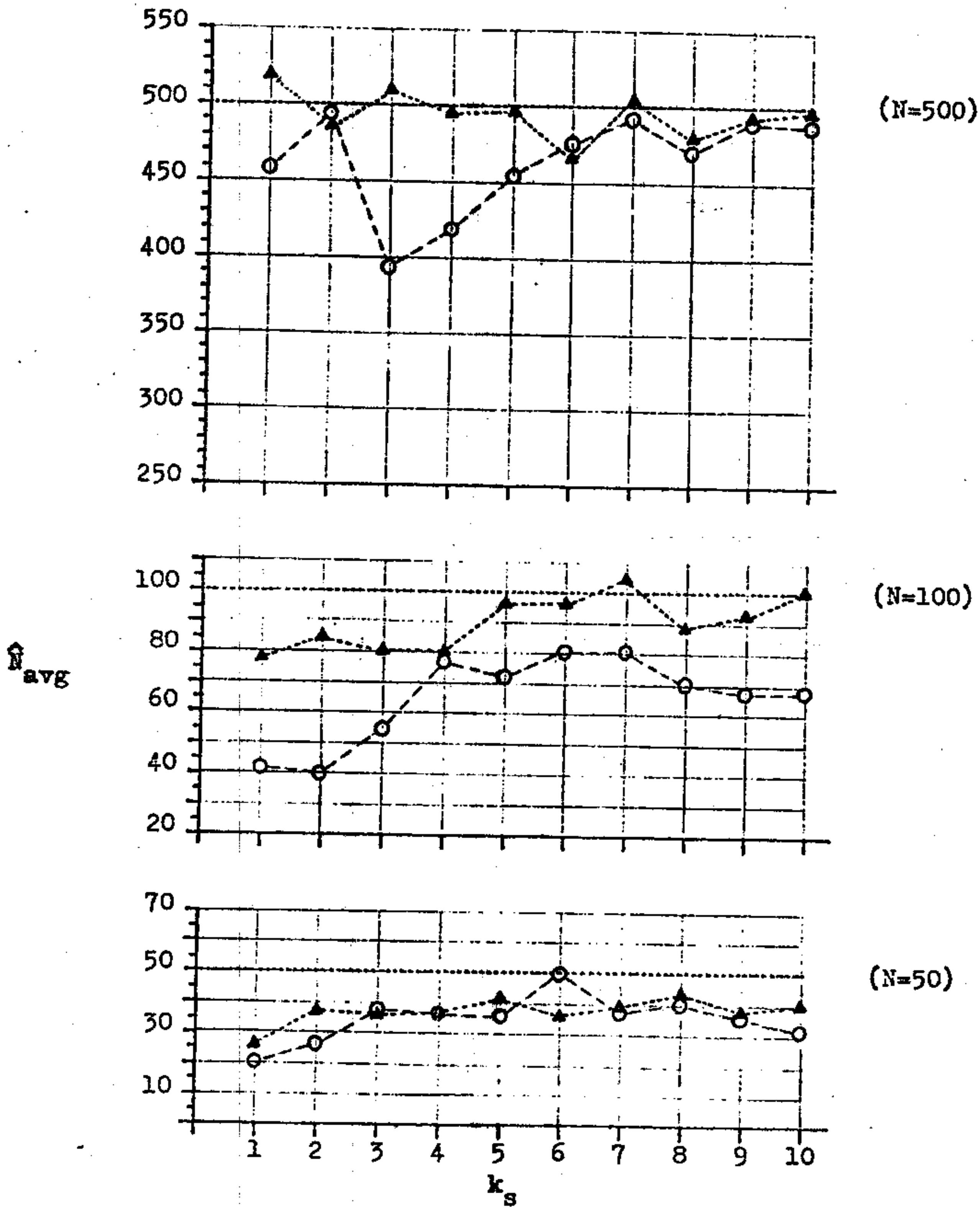
(Prob. Set 2: O : ε = .001, ▲ : ε = .005)

Figure 8



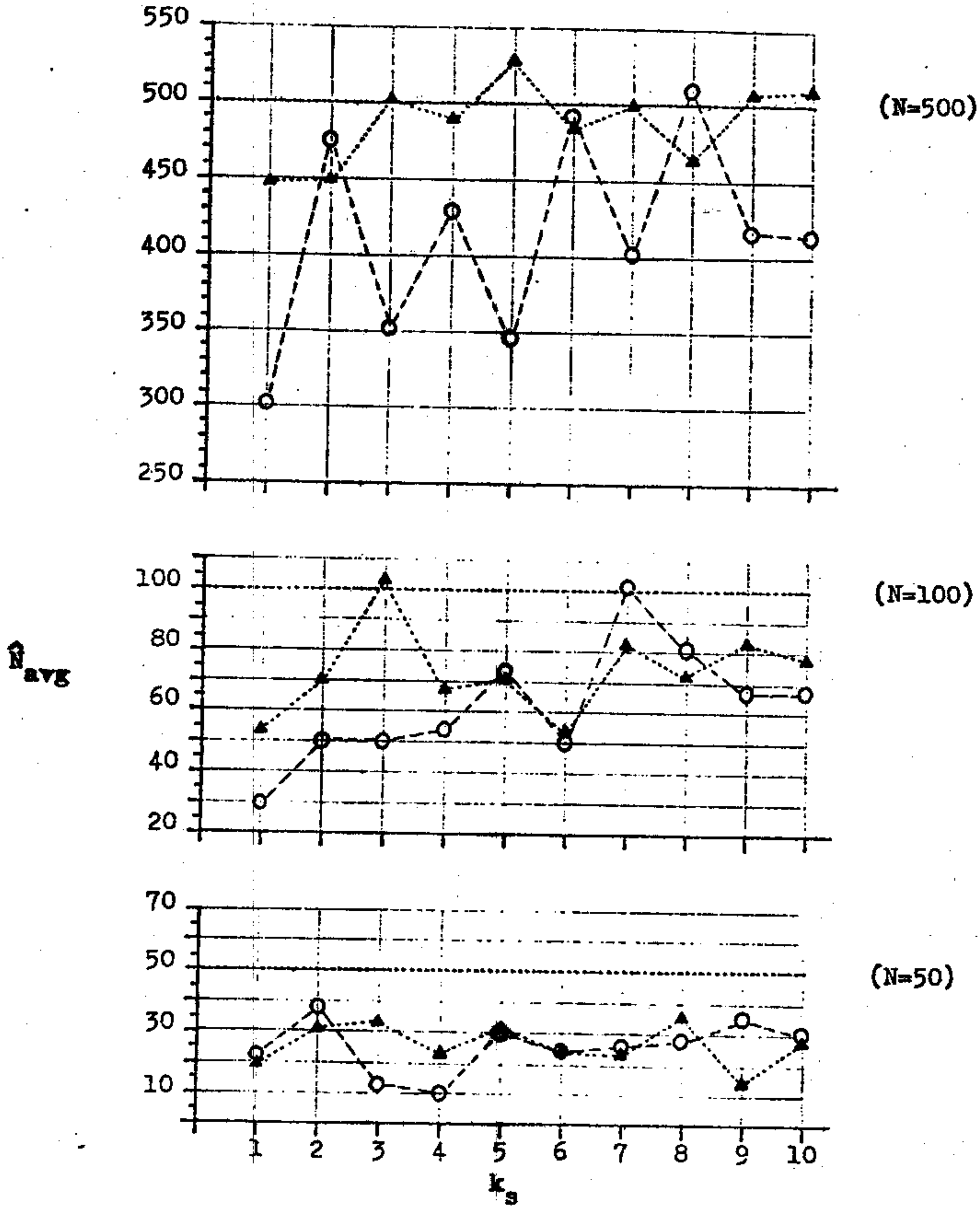
(Prob. Set I: O: ε= .001, ▲: ε= .005)

Figure 9



(Prob. Set 4; O : ε= .001, ▲ : ε= .005)

Figure 10



(Prob. Set 5: O : ε = .001, ▲ : ε = .005)

Figure 11

Table 10

Prob. Set	$\mathcal{E}$	k_s	$\hat{N}_{avg}$	C.V.
1	.001	10	41.84	.54
	.005	8	39.7	.40
2	.001	7	52.46	.33
		10	49.6	.47
	.005	9	46.72	.12
		7	49.4	.17
3	.001	8	40.58	.54
	.005	3	43.4	.39
4	.001	10	32.56	.50
		6	49.78	.57
	.005	7	39.42	.23
		8	44.82	.29
5	.001	1	22.8	.11
		2	38.16	.26
	.005	2	31.36	.27
		8	36.22	.61

(N = 50)

Table 11

P. Set.	ε	k_s	$\hat{N}_{avg}$	C.V.
1	.001	5	49.88	.49
		10	91.32	.56
	.005	2	95.2	.09
2	.001	6	97.66	.30
		5	99.9	.36
	.005	4	99.58	.06
3	.001	9	82.3	.48
	.005	9	97.78	.15
4	.001	6	81.94	.37
	.005	7	105.54	.09
		10	102.94	.14
5	.001	2	51.18	.47
		7	102.3	.69
	.005	3	104.48	.13

(N = 100)

Table 12

Prob. Set.	ε	k_s	$\hat{N}_{avg}$	C.V.
1	.001	8	455.46	.19
	.005	3	506.3	.06
		2	500.4	.09
2	.001	6	486.64	.13
	.005	4	504.36	.01
		6	499.08	.07
3	.001	3	460.4	.17
		4	497.26	.20
	.005	8	482.54	.06
		4	501.3	.08
4	.001	7	489.34	.13
	.005	8	478.44	.04
		5	496.88	.05
5	.001	2	474.36	.22
		8	510.94	.30
	.005	7	499.08	.05

(N = 500)

6. Concluding Remarks

The results exposed in Section 5, which refer to applications of the two stock estimation methods studied, merit diverse analyses due to the different nature of the data sources utilized.

The application of the basic method to the South Pacific tuna data sets required the adoption of certain empirical rules, in an attempt to deduce needed operational information from the available data or to impose a consistent structure to the system. The assumption made created a secondary source of uncertainties, which restricts the reliability of the final results.

In particular, many discrepancies can be verified in the logged detection coordinates if the vessel speed equals 10 n.miles/hr. Also, since the vessel has an effective search radius of, approximately, 3 n.miles (R. Hilborne, pers. comm.), a large error can be introduced in the assignement of detection site coordinates, considering the geographical scale of the system.

As mentioned is Section 5, some data points restrict set times to very short periods, corresponding to large volume of fish handled (e.g. a maximum of 10 min. for 550 individual fish tagged). On the contrary, it is estimated (R. Hilborne, pers. comm.) that an "average tagging operation will take, approximately, one hour for each school detected. Obviously, the data does not comport such order of magnitude. The potential inaccuracies of the data sets also threaten the precision of the rule used to identify distinct patches of schools, thus introducing another error in the determination of the total number of distinct patches of schools detected.

However, in order to establish some internal consistency to the available data, the criteria choices made permitted a coherent, albeit subjective, utilization of the data sets.

Therefore, the numerical values given in tables 7 and 8 are not to be considered as precise abundance estimates of the stock. The calculation of error bounds was not possible and the method will only perform well if the data base contains sufficient information to allow a correct identification of the number of independent patches of schools detected and the corresponding total search time, besides being provided the accurate knowledge of the operational variables which define the search parameter.

The performance of the extended method can be scrutinized by an exhaustive analysis of the simulation results shown in Figs. 7 to 11 and Tables 10, 11 and 12.

The five probability sets, were chosen to contrast the effect of variations about the uncertainty of fish location and, together with the range of values considered for the "switching parameter", k_s, (which can be classified as a control variable), set the performance characteristics of the method.

The main objective of the simulation was to investigate the predictive ability of the extended method. For the typical initial population levels chosen ($N = 50$, $N = 100$ and $N = 500$), a simulated "historical" record was created for fifty seasons. The performance indices were chosen to be the sample average, $\hat{N}_{avg}$ and the sample coefficient of variation (C.V.) for the fifty values of $\hat{N}$ evaluated from each configuration of N, ε, k_s, and probability set.

For N = 50 (Table 10), the method performs well for ε = .005 (Prob. Sets 2, 3 and 4) and certain values of k_s. The results for ε = .001 (Prob. Sets 2, 3 and 4) are also acceptable. They are consistent with the level of uncertainty about the fish location, which is lowest for Prob. Set 2 and increases in the following order: Prob. Set 4, Prob. Set 3, Prob. Set 1 and Prob. Set 5.

For N = 100 and N = 500 (Tables 11 and 12, respectively), the performance of the method improves as N increases and the predictive ability measured by $\hat{N}_{avg}$ and the corresponding value of C.V. especially indicate an excellent estimation for Prob. Sets 2, 4 and 3, except for ε = .001 in Prob. Sets 1 and 5.

The results show, therefore, that there is a switching strategy which "optimizes" the method's predictive capacity. The "optimum" values of k_s which range from 5 to 10 days may play an important role in the testing of fishing strategies and in the definition of management policies.

In particular, the results obtained suggest the inclusion of two distinct phases in the process of optimizing the economic yield of a migratory species' fishery: an information gathering operation, executed at the beginning of each season to provide the data for the estimation of the stock abundance, and an adaptive management policy (Smith and Walters 1981), (Walters 1981) to be devised in consonance with the estimate found and further refine it.

The performance of the method justifies the development of future research on the generalization of the model through the incorporation of added levels of realism. To account for the complex and real population dynamics phenomena, the population level N(t), at time t, must include the effects of recruitment, mortality, two- and three-dimensional migratory patterns and dispersion.

Mangel (1985) suggests some ways to initiate this investigative process aimed at modeling an open population which can be supported by the extended method's structure.

The next logical step is, consequently, to test the performance of the extended method in the case of a multi-cell fishing region, with adjacent cells of identical area (Appendix II).

Acknowledgements

This work is a result of research sponsored in part by NOAA, National Sea Grant College Program, Department of Commerce, under grant number NA85AA-D-SG140, project number R/F-109, through the California Sea Grant Program. The U.S. Government is authorized to reproduce and distribute for governmental purposes. Additional support was provided by the Secretaria Regional da Agricultura e Pescas, Azores Regional Government, in compliance with the Azores Tuna Stock Assessment Research Program. The data on South Pacific tuna were provided by Dr. Ray Hilborne, South Pacific Commission, Noumea, New Caledonia. I am especially indebted to Prof. Marc Mangel for his stimulating guidance and constant support.

References

KOOPMAN, B.C. 1980. Search and screening. Pergamon Press, New York, NY. 368 p.

MANGEL, M. 1985. Search models in fisheries and agriculture. Lecture Notes in Biomathematics, 61:103-138, M. Mangel (ed.). Springer-Verlag Berlin Heidelberg New York Tokyo.

MANGEL, M. 1986. Abundance estimates based on encounter rates. (To be published).

MANGEL, M. AND J. H. BEDER. 1985. Search and stock depletion: Theory and applications. Can. J. Fish. Aquat. Sci., 42:150-163.

PELLA, J.J. AND PSAROPULOS, C.T. 1975. Measures of tuna abundance from purse-seine operations in the eastern Pacific Ocean adjusted for fleet-wide evolution of increased fishing power, 1960-71. Bull. inter-Am. trop. Tuna Commn, 16:281-400.

SMITH, A.D.M. AND C.J. WALTERS. 1981. Adaptive management of stock - recruitment systems. Can. J. Fish. Aquat. Sci. 39(6):690-703.

U.S. COAST GUARD. 1973. National Search and Rescue Manual, CG 308. Superintendent of Documents, Washington D.C.

Walters, C.J. 1981. Optimum escapements in the face of alternative recruitment hypothesis. Can. J. Fish. Aquat. SCi., 38(6):678-689.

Appendix I

Derivation of the Log-Likelihood Function

From equations (8), (8a) and (8b), the logarithm of $\mathcal{L}$ is

$$\log \mathcal{L} = P_1(N;i,j) + P_2(N;i,j), \text{ where}$$

$$P_1(N;i,j) = \sum_{j=1}^{m_s} \left\{ \sum_{i=0}^{n_j-1} \left[\log\left(N-i-\sum_{\ell=0}^{j-1} n_\ell\right) + \log \varepsilon - \left(N-i-\sum_{\ell=0}^{j-1} n_\ell\right) \varepsilon S_{i+1,\ell_j} \right] - \left(N-\sum_{\ell=0}^{j} n_\ell\right) \varepsilon S_{\ell_j} \right\}$$

and

$$P_2(N;i,j) = \sum_{j=1}^{m_2} \log\left[1-\bar{p}_1(k_j)+\bar{p}_1(k_j) e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}\right] + \sum_{j=m_2+1}^{K-m_s} \log\left[1-\bar{p}_2(k_j)+\bar{p}_2(k_j) e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}\right].$$

Taking the derivative of the logarithm of $\mathcal{L}$ gives

$$\frac{\partial}{\partial N} \log \mathcal{L} = P_3(N;i,j) + P_4(N;i,j), \tag{A.1}$$

with

$$P_3(N;i,j) = \sum_{j=1}^{m_s} \left[\sum_{i=0}^{n_j-1} \left(\frac{1}{N-i-\sum_{\ell=0}^{j-1} n_\ell} - \varepsilon S_{i+1,\ell_j} \right) - \varepsilon S_{\ell_j} \right]$$

and

$$P_4(N;i,j) = -\sum_{j=1}^{m_2} \frac{\bar{p}_1(k_j)\varepsilon S_{k_j} e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}}{1-\bar{p}_1(k_j)+\bar{p}_1(k_j) e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}} - \sum_{j=m_2+1}^{K-m_s} \frac{\bar{p}_2(k_j)\varepsilon S_{k_j} e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}}{1-\bar{p}_2(k_j)+\bar{p}_2(k_j) e^{-\left(N-\sum_{\ell=0}^{\ell^*(j)} n_\ell\right)\varepsilon S_{k_j}}}$$

Setting the right side of (A.1) equal to zero yields algebraic equation (9).

The second derivation of log $\mathcal{L}$ is given by

$$\frac{\partial^2}{\partial N^2} \log \mathcal{L} = - X_0 \quad (N;i,j) + X_1 \quad (N;i,j) + X_2 \quad (N;i,j) \tag{A.2}$$

where

$$X_0 \quad (N;i,j) = \sum_{j=1}^{m_s} \sum_{i=0}^{n_j+1} \frac{1}{\left(N - i - \sum_{l=0}^{j-1} n_l\right)^2} \tag{A.3}$$

$$X_1 \quad (N;i,j) = \sum_{j=1}^{m_2} \frac{[1-\bar{p}_1(k_j)]\,\bar{p}_1(k_j)\,\varepsilon^2 S_{k_j}^2\, e^{-(N-\sum_{l=0}^{l^*(j)} n_l)\varepsilon S_{k_j}}}{\left[1-\bar{p}_1(k_j)+\bar{p}_1(k_j)\, e^{-(N-\sum_{l=0}^{l^*(j)} n_l)\varepsilon S_{k_j}}\right]^2} \tag{A.4}$$

and

$$X_2 \quad (N;i,j) = \sum_{j=m_2+1}^{k-m_s} \frac{[1-\bar{p}_2(k_j)]\,\bar{p}_2(k_j)\,\varepsilon^2 S_{k_j}^2\, e^{-(N-\sum_{l=0}^{l^*(j)} n_l)\varepsilon S_{k_j}}}{\left[1-\bar{p}_2(k_j)+\bar{p}_2(k_j)\, e^{-(N-\sum_{l=0}^{l^*(j)} n_l)\varepsilon S_{k_j}}\right]^2} \tag{A.5}$$

Appendix II

Extension of the Method to a M-cell Fishing Region

For a M-cell fishing region, the likelihood of the set of search times for the whole season (K days) is

$$\mathcal{L} = \prod_{\nu=1}^{3} \mathcal{L}_\nu \quad ,$$

where

$$\mathcal{L}_1 = \prod_{j=1}^{m_s} \left\{ \prod_{i=0}^{n_j-1} \left[\left(N-i-\sum_{l=0}^{j-1} n_l\right) \varepsilon \, e^{-\left(N-i-\sum_{l=0}^{j-1} n_l\right) \varepsilon S_{i+1,l_j}} \right] e^{-\left(N-\sum_{l=0}^{j} n_l\right) \varepsilon S_{l_j}} \right. ,$$

$$\mathcal{L}_2 = \prod_{j=1}^{m_2} \left[1-\bar{p}_1(k_j)+\bar{p}_1(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right) \varepsilon S_{k_j}} \right] \prod_{j=m_M+1}^{K-m_s} \left[1-\bar{p}_M(k_j)+\bar{p}_M(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right) \varepsilon S_{k_j}} \right]$$

and

$$\mathcal{L}_3 = \prod_{i=2}^{M-1} \left\{ \prod_{j=1+m_i}^{m_{i+1}} \left[1-\bar{p}_i(k_j)+\bar{p}_i(k_j)\, e^{-\left(N-\sum_{l=0}^{l^*(j)} n_l\right) \varepsilon S_{k_j}} \right] \right\} ,$$

where $m_2 = d_1$,

$m_{i+1} = m_i + d_i$, $2 \leq i \leq M-1$, with

d_i ≡ number of days of unsuccessful catch in cell i (i = 1,....M-1),

and, as stated previously, $l^*(j)$ = max i, such that $l_i < k_j$ (i = 1,....m_s).

and $\{\bar{p}_i(k)\}$ is the probability set corresponding to cell $\mathcal{R}_i$ (i=1,...,M).

Simulation Program for the Extended Method (GENMLE)

```
1 REM *************************GENMLE.BAS******************************
5 REM GENMLE SIMULATES CATCH DATA AND ENCOUNTER RATES FOR A TWO-CELL
6 REM FISHING REGION, GIVEN THE INITIAL STOCK SIZE (N), CATCHABILITY (E),
7 REM AND THE PROBABILITIES THAT FISH ARE PRESENT ON DAY K ON CELL 1 AND
8 REM CELL 2, P1(K) AND P2(K), RESPECTIVELY (K=1,...,14).
9 REM IT CALCULATES THE MEAN AND VARIANCE OF THE MLE ESTIMATE FOR N OF
10 REM A SAMPLE WITH SIZE OF NS FISHING SEASONS.
11 REM SWITCHING STRATEGY: AFTER NK DAYS OF NO CATCH IN CELL ONE, MOVE TO
12 REM CELL TWO AND STAY IN CELL TWO UNTIL END OF SEASON.
15 LPRINT "genmle"
20 DIM SIO(14.500),P1(14),P2(14),S(14),KT(14),KL(14)
21 INPUT "Enter N1, N2 and N3";N1,N2,N3
22 LPRINT "First stock size=";N1
23 LPRINT "Second stock size=";N2
24 LPRINT "Third stock size=";N3
25 INPUT "Enter JK, NS and Epsilon";JK,NS,E
26 LPRINT "Number of seasons=";NS
27 LPRINT "Epsilon=";E
28 LPRINT "Convergence bound=";JK
29 LPRINT
30 OPEN "I",#1,"DATP.DAT"
35 FOR I=1 TO 14
40 INPUT #1,P1(I),P2(I)
45 NEXT I
47 NK=1
48 LPRINT "k=";NK
50 N=N1
51 TOT=0!
52 SSQ=0!
55 NP=0
56 NCTX=0
57 MZ=0
70 NP=NP+1
75 NR=N
80 K=0
85 KSUM=0
90 RR=0!
95 SUM1=0!
100 FOR I=1 TO 14
105 IF RR<.5 THEN GOTO 120
110 PY=P2(I)
115 GOTO 125
120 PY=P1(I)
125 RANDOMIZE TIMER
130 Q=RND
```

```
135 IF Q<PY THEN GOTO 145
140 NR=0
145 SE=12!-SUM1
150 Y=RND
155 QX=1!-EXP(-E*NR*SE)
160 IF Y=1! THEN GOTO 170
165 IF Y<QX THEN GOTO 225
170 IF SUM1=0! THEN GOTO 180
175 GOTO 210
180 S(I)=SE
185 KT(I)=0
186 IF RR>.5 THEN GOTO 295
190 NCTX=NCTX+1
195 IF NCTX=NK THEN GOTO 270
200 GOTO 295
210 S(I)=SE
215 KL(I)=0
220 GOTO 250
225 NCTX=0
230 ST=-LOG(1.-Y)/(E*NR)
235 KL(I)=1
240 NR=NR-KL(I)
244 K=K+KL(I
245 SIO(I.K)=ST
246 KT(I)=K
247 SUM1=SUM1-ST
248 IF SUM1<12! THEN GOTO 145
250 KSUM=KSUM-K
255 K=0
260 SUM1=0!
265 GOTO 295
270 RR=RR+1!
275 NR=N-KSUM
280 L=I
285 K=0.
290 SUM1=0!
295 NEXT I
300 Z=KSUM
305 IF KSUM=0 GOTO 450
310 SUML=0!
315 SUMP=0!
320 SLGF=0!
325 SLGP=0!
330 NT=0
335 FOR I=1 TO 14
```

```
340 IF I<(L+1) THEN GOTO 355
345 PY=P2(I)
350 GOTO 365
355 PY=P1(I)
365 IF KT(I)=0 THEN GOTO 395
370 FOR M=1 TO KT(I)
371 ZY=Z-NT-M+1!
372 IF ZY<=0! THEN GOTO 512
375 SUML=SUML+1!/ZY-E*SIO(I,M)
380 SUMP=SUMP+1!/(ZY^2)
385 NEXT M
386 SUML=SUML-E*S(I)
390 NT=NT+KT(I)
391 GOTO 420
395 PK=PY*EXP(-(Z-NT)*E*S(I))
400 PKF=S(I)*PK/(1!-PY+PK)
405 PP=PK*(1!-PY)*(S(I)/(1!-PY+PK))^2
410 SLGF=SLGF-PKF
415 SLGP=SLGP+PP
420 NEXT I
425 FX=SUML+E*SLGF
430 FPX=(E^2)*SLGP-SUMP
431 IF FPX<0! THEN GOTO 435
432 NP=NP-1
433 GOTO 56
435 IF ABS(FX)<.000001 THEN GOTO 450
436 IF ABS(FPX)<.000001 THEN GOTO 512
439 MZ=MZ+1
440 Z=Z-FX/FPX
443 IF ABS(Z)>JK THEN GOTO 512
444 IF MZ=1000 THEN GOTO 512
445 GOTO 310
450 NX=INT(Z)
451 MZ=0
455 IF NX=0 THEN GOTO 432
465 TOT=TOT+NX
470 SSQ=SSQ+NX^2
475 IF NP=NS THEN GOTO 485
480 GOTO 70
485 AV=TOT/NP
490 AVSQ=SSQ/NP
495 VAR=AVSQ-AV^2
498 CV=SQR(VAR)/AV
505 LPRINT "n=";N,"avg=";AV,"var=";VAR,"c.v.=";CV
511 GOTO 665
512 LX=KSUM
513 FZ=0!
514 PIC=1!
515 PINC=1!
520 NT=0
525 FOR I=1 TO 14
```

```
530 IF I<(L+1) THEN GOTO 545
535 PY=P2(I)
340 GOTO 550
545 PY=P1(I)
550 IF KT(I)=0 THEN GOTO 600
555 PC=1!
556 FOR J=1 TO KT(I)
558 LNX=LX-NT-J+1
559 SST=-LNX*E*SIO(I,J)
561 PLC=LNX*E*EXP(S$T)
565 PC=PLC*PC
570 NEXT J
575 NT=NT+KT(I)
580 PL=EXP(-(LX-NT)*E*S(I))
585 TPL=PL*PC
590 PIC=PIC*TPL
595 GOTO 610
600 PLN=1!-PY+PY*EXP(-(LX-NT)*E*S(I))
605 PINC=PLN*PINC
610 NEXT I
615 HY=PINC*PIC
620 IF HY<FZ THEN GOTO 645
625 FZ=HY
630 LX=LX+1
635 IF LX>JK THEN GOTO 855
640 GOTO 514
645 NX=LX
650 GOTO 451
855 PRINT "no convergence"
656 MZ=0
860 GOTO 56
865 IF N=N1 THEN GOTO 880
870 IF N=N2 THEN GOTO 890
875 IF N=N3 THEN GOTO 900
880 N=N2
885 GOTO 51
890 N=N3
895 GOTO 51
900 NK=NK+1
905 IF NK=11 THEN GOTO 925
910 LPRINT
915 LPRINT "k=";NK
920 GOTO 50
925 END
```