

Determination of Fish Movement Patterns from Tag Recoveries using Maximum Likelihood Estimators

Ray Hilborn

Fisheries Research Institute WH-10, University of Washington, Seattle, WA. 98195 U.S.A.

Hilborn, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. *Can. J. Fish. Aquat. Sci.* 47: 635–643.

A general method for analysis of movement data from tag returns is proposed which has four major components: (1) a population dynamics and movement model that describes how the number of tagged individuals in each spatial location changes over time; (2) an observation model which describes how the tags are recovered and reported; (3) a likelihood function that specifies the likelihood of observing a specific number of recoveries in each space/time stratum as a function of the number thought to be there under a specific set of parameters of the population dynamics, movement and observation models, and (4) a nonlinear function minimization computer algorithm. This approach is applied to movements of skipjack tuna (*Euthynnus pelamis*). When tagging and recapture take place in each spatial stratum, reliable estimates of movement rates can be obtained. The approach described is completely general and can be used in cases where movement takes place continuously, or only once in the life history. Methods for determining confidence limits and evaluation of residuals are presented and extensions that include tagging mortality, tag shedding, and size specific vulnerability are discussed.

Une méthode générale d'analyse des données portant sur les déplacements à partir des bagues retournées est proposée : cette méthode compte quatre éléments importants : (1) un modèle de mouvement et de dynamique des populations qui décrit comment la population constituée par le nombre de sujets marqués dans chaque emplacement spatial, se transforme avec le temps; (2) un modèle d'observation qui décrit comment les bagues sont récupérées et rapportées; (3) une fonction de vraisemblance qui détermine la vraisemblance de l'observation d'un nombre donné de récupérations dans chaque strate spatiale-temporelle en fonction du nombre qu'on estime être présent en vertu d'un ensemble précis de paramètres retenus pour les modèles de la dynamique des populations, des mouvements et des observations; enfin, (4) un algorithme de minimisation de la fonction non linéaire. Cette approche est appliquée aux déplacements de la thonine (*Euthynnus pelamis*) à ventre rayé. Avec le marquage et le recapture dans chaque strate spatiale, il est possible d'obtenir des évaluations fiables des taux de déplacement. L'approche décrite est on ne peut plus générale et peut s'appliquer aux cas de déplacements continus ou aux cas d'un seul déplacement dans la vie. Les méthodes utilisées pour déterminer les limites de confiance et l'évaluation des résidus sont présentées; il est question aussi de la mortalité par marquage, de la perte des bagues et de la vulnérabilité associée à chaque taille.

Received April 6, 1987

Accepted October 16, 1989

(J9227)

Reçu le 6 avril 1987

Accepté le 16 octobre 1989

Many animals are highly mobile. Individual birds, mammals, and fish have been known to move over 10 000 km and there are entire populations, such as grey whales and many arctic breeding birds, whose annual migration cover similar distances. In some cases one can track such populations, or more rarely an individual, over long distances. More commonly, however, our knowledge of long distance movements of animals comes from tagging studies.

This is particularly true of fish, where it is beyond current technology to track individuals very far. Tagging data tell us where tagged fish were released and where they were recovered. Since most large scale tagging studies rely on commercial capture of tagged fish, multiple recaptures of individuals are rare; the fish are either dead or too valuable to return to the water.

Fisheries managers and biologists want to know the movement patterns of populations or individuals to assess the interaction between fisheries in different spatial locations and to define the discreteness of stocks. Tagging studies, though, are often the only way to assess total stock size, natural and fishing

mortality, so that determination of movement often becomes a secondary goal only. Nevertheless, there is now a large body of tagging data for fisheries around the world which provides the basis of most of our current knowledge of fish movement.

Once a fisheries biologist has collected tagging data from a population, he must decide how to use this information to reveal fish movements. The most commonly used technique is to draw arrows from where fish were tagged to where they were recovered.

Occasionally more quantitative analyses are used. For instance, one can plot the distance travelled versus time since release (Schaefer et al. 1961). One obvious problem in the analysis of tag recoveries is that in areas without fishing effort no tags are recovered, or more generally, the number of recoveries is related to the fishing effort in that area. To compensate for this, recoveries per unit of fishing effort are used rather than total recoveries. The Inter-American Tropical Tuna Commission has analyzed tagging data in this way (Bayliff 1979), using measures of directional and random movement developed by Jones (1959, 1976).

A major problem with most of these methods of analysis is that no account is taken of the probability of capture between the time of release and the location of recovery. Few fish may be caught in a certain area simply because there is a high probability of capture along the way. Secondly, the methods (with the exception of Jones' method) do not state a specific hypothesis about fish movement and then attempt to estimate the parameters and validity of that hypothesis. Jones' method suffers because it allows only for random or simple directional movement.

There has been remarkably little formal statistical work on analysis of movement data. Darroch (1961) and Arnason (1972, 1973) examined spatially stratified capture recapture models, but only for the case with multiple recaptures. As noted in Schwarz and Arnason (1990), these studies assumed equal probability of capture in all areas. Further, these approaches are not applicable to most fisheries tagging studies in which individuals are recaptured once by a commercial fishery and recovered dead. Only recently has attention turned to using the traditional models and approaches of mark-recapture studies as summarized in Burnham et al. (1987) on spatially structured problems with unequal fishing effort in the spatial strata.

Ishii (1979) used a model of movement in which he specifies a Markovian movement model between geographic areas in the eastern tropical Pacific. He simulated the movement of tagged fish and used nonlinear minimization techniques to find the movement probabilities that minimize the difference between observed and expected number of recoveries in each spatial area. Ishii's model includes natural mortality and tag shedding as parameters to be estimated simultaneously with movement.

Sibert (1984) modeled natural mortality, fishing mortality, and movement between two countries, Papua New Guinea and the Solomon Islands. He used tagging data to determine the mortality rates and exchange rates between the two countries. Like Ishii, he estimated the parameters with least squares.

These two papers constitute a new approach to the analysis of animal movement data, but are unknown beyond a small circle of tuna biologists. For example, in a major review of tuna movement patterns (Hunter et al. 1986), there is no reference to Ishii's work. This is probably because Ishii's and Sibert's papers were published in rather obscure locations as specific case examples rather than general approaches. Recently, Schwarz (1988) and Schwarz and Arnason (1990) described extensions of the traditional statistical analysis of mark-recapture using explicit multinomial probability functions.

In this paper I present a general framework for the analysis of animal movement, natural mortality, and harvest from tagging data and show that Ishii's and Sibert's studies are special cases of the same approach. Their heuristic least squares approach is replaced here by a more rigorous maximum likelihood method based on the Poisson distribution. Two examples are presented and a number of potential extensions and modifications of the method are discussed. The method I present has a number of significant features: (1) it is applicable to nearly any tagging study, (2) it requires no detailed statistical derivations, but relies instead on nonlinear search techniques to minimize a simple likelihood, (3) it is easily implemented on microcomputers with a few hours programming time.

General Framework

I define the following symbols to be used in this section.

- T_{iat} = the number of tags released from tag group i , area a , time t
- \mathbf{T} = a matrix of all the T_{iat} values
- \hat{N}_{iat} = the predicted number of tagged fish of group i present in area a at time t
- R_{iat} = the number of tags recovered from tag group i , area a , time t
- \mathbf{R} = a matrix of R_{iat} values
- \hat{R}_{iat} = the predicted number of tags recovered from tag group i , area a , time t
- $\hat{\mathbf{R}}$ = a matrix of \hat{R}_{iat} values
- \mathbf{x} = a vector of parameters of the population dynamics and movement model
- \mathbf{y} = a vector of parameters of the observation model
- p_{ij} = probability of movement from area i to j
- E_{at} = fishing effort in area a , time t
- q_a = catchability coefficient in area a
- h_{at} = harvest rate in area a , time t , ($= E_{at} q_a$)

A tag group i for my analysis is defined as a group of fish tagged in the same space and time stratum, but could be extended to include size groups, sex, or whatever criteria are thought to be important in movement, survival, and probability of capture. In this analysis I write all equations in discrete space and time. The same models can easily be transformed to continuous equations and solved in exactly the same fashion.

My approach for the analysis of tagging data has four elements: (1) a population dynamics and movement model, (2) an observation model, (3) likelihood of recoveries, and (4) a nonlinear function minimization procedure.

Population Dynamics and Movement Model

A simulation model for the survival and movement of tagged individuals must be specified. This must include natural and fishing mortality as well as movement and could include tagging mortality and tag shedding. If fish vulnerability is size related, growth may need to be a specific part of this model.

If we consider the dynamics in n areas, each connected by a Markovian exchange process, the following model could be used.

$$(1) \quad \hat{N}_{iat+1} = \sum_{j=1}^n (\hat{N}_{ijt}(1 - h_{jt})p_{ja}) + T_{iat}$$

$$h_{at} = E_{at} q_a$$

Given the known tag releases \mathbf{T} , the fishing efforts E_{at} , and the p and q values, we can predict how many individuals would be alive from each tag group in each area. While equation (1) does not explicitly allow for a survival from natural mortality term, natural mortality is implicit when the $\sum_j p_{ij}$'s for an area do not sum to 1.0. No allowance for higher mortality among fish that are moving is provided, but this also could be added.

Observation Model

A model is required that specifies the relationship between the tagged fish predicted to be in an area under equation (1) and the tagged fish that are actually captured and reported to the management agency. In its simplest form this is the predicted number of tagged fish harvested, as taken from the population dynamics model, but can include nonreporting of tags, misreporting, and tag loss.

If we assume that harvest rate is proportional to fishing effort, and that all tags that are captured are returned, then equation

(2) below is the specific observation model for our n area Markovian movement model.

$$(2) \hat{R}_{iat} = \hat{N}_{iat} q_a E_{at}$$

Likelihood of Recoveries

Given the tag release data, and the population dynamics and movement model of equation (1), we know how many tagged fish we expect to be in each area at each time from each tagging group. Using equation (2) we know how many recoveries to expect from each group, in each area, at each time. We thus have a matrix of predicted recoveries $\hat{\mathbf{R}}$ for each tag release group by time and area. We also have a matrix of observed recoveries \mathbf{R} . We want to calculate the likelihood of the observed recoveries \mathbf{R} given the parameters of our population and observation models, that is

$$(3) L(\mathbf{R}|\mathbf{p}, \mathbf{q}, \mathbf{E}, \mathbf{T})$$

which using equations (1) and (2) becomes

$$(4) L(\mathbf{R}|\hat{\mathbf{R}})$$

which is the product of all the individual likelihoods for each tag group, space, and time stratum.

$$(5) \prod_i \prod_a \prod_t L(R_{iat}|\hat{R}_{iat})$$

A statistical likelihood or probability model must be developed that calculates the likelihood of the number of recoveries actually being reported if the population dynamics and observation models are true.

There are two traditions in maximum likelihood estimation in the analysis of tagging data. The use of explicit multinomial calculations dominates; almost all of the methods described in Burnham et al. (1987) rely on specifying an explicit model of tag survival and recovery, then writing the probability of recoveries as an explicit multinomial of the parameters of the model. The recoveries in each stratum are expected to be multinomially distributed. The second tradition is found primarily in the use of generalized linear models (Chapman 1948; Cormack 1981; McCullagh and Nelder 1983) in which the number of expected recoveries in each stratum is expected to be Poisson distributed. My method follows this school. Because there is a low probability of any individual tag being recovered in any space/time stratum, the multinomial can be approximated by the Poisson.

The Poisson distribution for an individual tag group, space, and time can be written as

$$(6) L(R_{iat}|\hat{R}_{iat}) = \frac{e^{-\hat{R}_{iat}} \hat{R}_{iat}^{R_{iat}}}{R_{iat}!}$$

The total likelihood for all observed recoveries given the parameters of the population dynamics and observation models, and the number of tags released (\mathbf{T}) is thus

$$(7) L(\mathbf{R}|\mathbf{p}, \mathbf{q}, \mathbf{E}, \mathbf{T}) = \prod_i \prod_a \prod_t \frac{e^{-\hat{R}_{iat}} \hat{R}_{iat}^{R_{iat}}}{R_{iat}!}$$

Nonlinear Function Minimization Procedure

A nonlinear function minimization procedure is necessary to find the values of the population dynamics and observation models that maximize the likelihood of the observations given the parameters of the model. This type of computer program is commonly available on minicomputers and growing in avail-

ability on microcomputers. The standard reference is Bard (1974) and an easily implemented BASIC computer program is described in Schnute (1982).

Comparison to Other Methods

The above approach is quite similar to the Ishii (1979) model, except it is much more general and uses maximum likelihood, rather than least squares. My approach is fundamentally an extension of the generalized linear model approach of Cormack (1981), but is extended in two ways. First, by adding a spatial dimension to the model we are able to estimate movement patterns. Second, the generalized linear model approach is quite confining in the models it can consider, and my approach breaks this confinement to allow models that cannot be written in a linear form by using a general nonlinear search algorithm instead of the computer code for linear models. However, the basic framework of predicting recoveries in each space/time stratum is identical to that used by Cormack.

My approach is closely related to the multinomial maximum likelihood estimators for analysis of tag returns. The multinomial equivalent of equation (7) is (Seber 1982 page 131):

$$(8) L(\mathbf{R}|\mathbf{p}, \mathbf{q}, \mathbf{E}, \mathbf{T}) = \prod_i \left[\frac{(\sum_a T_{iat})!}{(\prod_a R_{iat})! m_i!} \right] \times \left[\prod_a \prod_t v_{iat}^{R_{iat}} w_i^{m_i} \right]$$

where

$$v_{iat} = \frac{\hat{R}_{iat}}{\sum_a T_{iat}}$$

$$w_i = \frac{m_i}{\sum_a T_{iat}}$$

where m_i is the number of tags from release group i that were never recovered. The numerator of the left hand term is the total number of tags put out in group i , the first term in the denominator is the product of all the actual recoveries for tag group i , and the second term in the denominator is all the tags never recovered. The term v_{iat} is the proportion of tags of group i that are expected to be recovered in area a time t , and the term w_i is the proportion of all tags from group i expected to be not recovered. The multinomial likelihood is as easily implemented as the Poisson for the types of movement models described here.

Thus, one can use my methodology, but instead of following the school of Chapman, McCullagh and Nelder, and Cormack, follow the multinomial approach which fills the pages of Seber (1982) and Burnham et al. (1987). We substitute the use of the multinomial likelihood for the Poisson.

I have performed a number of Monte-Carlo experiments comparing the Poisson and multinomial approaches, and they produce effectively identical answers. The multinomial and Poisson will only differ when the probability of capture in a single space-time strata is not small. Computational convenience rather than philosophical viewpoint should govern the choice of approach.

This simulation approach represents a break with tradition. The normal approach is to algebraically derive the v and w values. My approach calculates them by simulation. The simulation approach is primarily advantageous when the models become complex, as most movement models are. Indeed, the movement models with more than two spatial areas would

TABLE 1. Estimates of catchabilities (q 's), and movement rates from the present analysis, Sibert (1984), and Kleiber et al. (1983). The other column includes both natural mortality and movement to other areas. The Kleiber et al. analysis was done on a fishery-by-fishery basis and does not estimate movement, so we call their total loss rate term 'other'. The movement parameter from PNG to PNG is $2p_{11} = 0.71$, movement from PNG to SOL is $p_{12} = 0.0001$, movement from SOL to PNG is $p_{21} = 0.0031$, and the movement from SOL to SOL is $p_{22} = 0.9$.

		Movement probabilities		
		To		
	From	PNG	SOL	Other
Present analysis				
q_{PNG}	PNG	0.7100	0.0001	0.2899
q_{SOL}	SOL	0.0031	0.9000	0.0969
Sibert (1984)				
q_{PNG}	PNG	0.5900	0.0340	0.376
q_{SOL}	SOL	0.0130	0.9280	0.059
Kleiber et al. (1983)				
q_{PNG}	PNG	0.5300		0.470
q_{SOL}	SOL		0.8400	0.160

require enormous algebraic expenditure. The complexity of the algebra increases geometrically with both the number of areas and the number of times, so that a seven area model with movement between each area for 20 time periods would not only be tedious, it would be impossible. This is much more satisfactorily replaced with a few lines of computer program.

A Two-Fishery Tag Exchange Model for Skipjack Tuna

As a first example I analyze the Papua New Guinea (PNG)-Solomon Islands (SOL) data presented in Sibert (1984). These two counties are major tuna fishing nations in the western Pacific. Fish tagged in each fishery were known to move to the other fishery, but also to different fisheries not included in the analysis. For simplicity, and unlike Sibert, we will ignore tagging mortality, tag shedding, and nonreporting of tags. The population dynamics model is thus:

$$\begin{aligned}
 \hat{N}_{iat+1} &= \hat{N}_{iat}(1 - h_{1t})p_{11} + \hat{N}_{iat}(1 - h_{2t})p_{21} + T_{iat} \\
 \hat{N}_{iat+1} &= \hat{N}_{iat}(1 - h_{2t})p_{22} + \hat{N}_{iat}(1 - h_{1t})p_{12} + T_{iat} \\
 h_{1t} &= E_{1t} q_1 \\
 h_{2t} &= E_{2t} q_2
 \end{aligned}
 \tag{9}$$

where \hat{N}_{iat+1} is the number of fish from tag group i predicted to be alive in area a at time $t + 1$, h_{1t} is the fishing mortality rate in area 1 at time t , T_{iat} is the number of tags released for tag group i in area a at time t , and the p'_{ij} 's are the proportion moving from one area to another. In this analysis there are two tag groups; one group released in Papua New Guinea, and one group released in the Solomon Islands. E_{at} is the fishing effort in area a at time t , and q_a is the catchability coefficient in area a .

Note that implicit in this model is the potential for the p 's for each area to not add to 1 because of movement to other areas and natural mortality, which are indistinguishable. If, for instance, $p_{11} + p_{12} = 0.72$ (as they do in Table 1) then 28% of the individuals either die to natural mortality or move to areas

TABLE 2. Estimated fishing mortalities, catchabilities, and movement probabilities for seven-fishery model. Fishery codes are PAL = Palau, YAP = Yap, INT = International waters between the Federated States of Micronesia (YAP, TRK, PON), Palau, PNG, and Indonesia, TRK = Truk, PON = Ponape, PNG = Papua New Guinea, SOL = Solomon Islands, OTH = loss of tags to other areas and natural mortality. For all countries except PNG and SOL the fishing mortality (F) is estimated (effort assumed constant at 1.0; for PNG and SOL the catchability coefficient q is estimated since effort data were available). If a 0 appears in the movement probability table, it means that these countries are not adjacent and no direct exchange rate was estimated unless an * appears, in which case it means the best estimate was that no exchange occurs between countries. The lower portion of the table constitutes the p matrix, with the exception of the "other" column which is calculated as one minus the sum of the other columns for each row and represents both natural mortality and movement to other areas.

	PAL	YAP	INT	TRK	PON	PNG	SOL	
<i>F</i> or <i>q</i>	.0049	.0043	.0046	.0054	.0065	.00012	.00003	
Probability of movement To								
From	PAL	YAP	INT	TRK	PON	PNG	SOL	OTH
PAL	0.677	0.047	0.107	0	0	0	0	0.168
YAP	0.008	0.707	0.032	0.041	0	0	0	0.211
INT	0.008	0.041	0.808	0.100	0.00*	0.034	0	0.008
TRK	0	0.066	0.027	0.635	0.084	0	0	0.187
PON	0	0	0	0.082	0.683	0.00*	0	0.235
PNG	0	0	0.052	0	0.00*	0.600	0.013	0.336
SOL	0	0	0	0	0	0.024	0.873	0.124

other than the Solomon Islands. There are thus six parameters to be estimated in the population dynamics model, four p 's and two q 's. These may be confounded.

The observation model is:

$$(10) \quad \hat{R}_{iat} = \hat{N}_{iat} q_a E_{at}$$

which has no additional parameters beyond those in the population dynamics model. I assume that 100% of the tags were returned, and use the Poisson likelihood function as the likelihood.

This analysis (Table 1) differs from the analyses of Sibert (1984) primarily in that my estimated loss rate in PNG is considerably lower. This is probably due to the use of the maximum likelihood estimator which assigns more weight to the few recoveries that took place more than a year after marking.

Extension to Seven Fisheries

We can expand the previous analysis to include a much larger geographic area. The skipjack survey and assessment programme of the South Pacific Commission tagged over 150 000 skipjack tuna throughout the western Pacific from 1977 to 1981 (Kearney 1983). Figure 1 shows the part of the western Pacific under consideration. Skipjack were tagged in all jurisdictions except Indonesia in Fig. 1 and most tags were recovered in this area. Tagging and recoveries from Fiji are the major data excluded from the current analysis.

The model for this analysis is the general n area Markovian model given in equation (1), and is the same as the two area model of the previous section except instead of two-fishery exchange we consider seven fisheries. Because not all fisheries are contiguous with one another, we only estimate the exchange

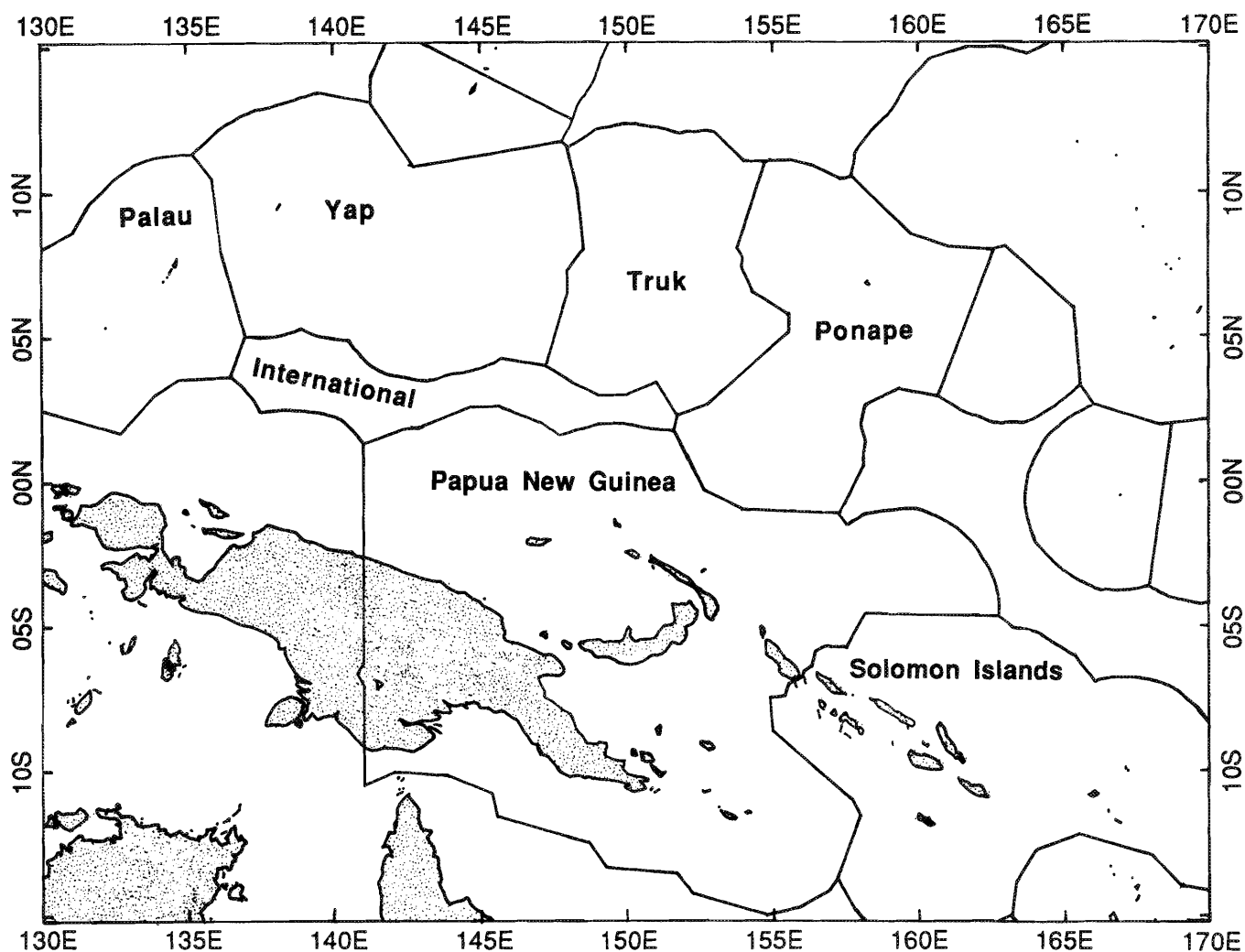


FIG. 1. Map showing major tuna fishing countries of western Pacific used in seven-fishery tag analysis model.

rate between contiguous fisheries, which reduces the number of movement parameters to be estimated from $7 \times 7 = 49$ to 25. We thus assume that it takes a minimum of 2 mo for a fish to move from one area to a noncontiguous area. No fish are known to have moved between noncontiguous areas in a 1-mo period and with the movement rates between contiguous areas generally low (< 0.10 per month), noncontiguous movement in 1 mo is unlikely enough to be safely ignored. Fishing effort data were available for PNG and SOL only, so I assumed effort was constant at 1.0 in all other areas. The q estimated for the other countries becomes, in fact, the monthly fishing mortality rate. Even in PGN and SOL, the effort data were not complete over the time period of tagging and recovery (October 1977–December 1981).

Data

I defined each tag group as fish released during a single month in a specific fishery. During the skipjack program there were 18 tag groups in the area considered in this analysis. Fish were tagged in all areas except international waters. All release, recovery, and effort data are available from the author.

We consider only recoveries from the seven countries included in this model. There were recoveries outside the

region, particularly a few tags from Palau and PNG recovered in Indonesia, but these were ignored and in effect become another form of unaccounted for disappearance that includes natural mortality and movement to other areas.

Results

The fishing mortality rates (Table 2), are about 0.005 per month, a figure consistent with previous estimates for countries other than PNG and SOL. The catchabilities for PNG and SOL are close to the previous estimates from Table 1. The movement probabilities show no great surprises, but are very interesting nevertheless. The results for PNG and SOL are essentially the same as we found in the two area model; indeed these countries seem remarkably isolated from the rest of the western Pacific. PNG has an extremely high loss rate, but the tags do not seem to go to the other countries considered in this model. Seven fish tagged in PGN were recovered in Indonesia, but these recoveries were not included in the analysis since we have no knowledge of the time pattern of fishing effort or probability of tag returns from Indonesia.

Loss to natural mortality and other areas is rather low; with the exception of PNG it is usually below 25% per month. Kleiber et al. (1983) estimated the natural mortality rate over the

entire skipjack data set at 17% per month. If one believes in a uniform natural mortality rate for skipjack, then the very low total loss rates from SOL, international waters, Truk, and Palau would suggest that the natural mortality rate must be much lower, perhaps around 10% per month.

Potential Modifications

Generalization of Population Dynamics and Observation Models

The models presented so far have been discrete Markovian movement models. The basic approach can be generalized to a very wide variety of models, which can be written as:

$$(11) \quad \hat{N}_{iat} = f(\hat{N}_{i1...m-1} | \mathbf{x}, \mathbf{T})$$

where f is the population dynamics model, the parameters of which could include natural mortality rate, vulnerability to fishing gear, and movement probabilities, as well as tagging mortality, tag shedding, and growth. In the simplest case, for a single release group in each area, the subscript i denotes area of tagging, but could include time of tagging, size at tagging, etc.

The observation model can be generally written as:

$$(12) \quad \hat{R}_{iat} = g(\hat{N}_{iat} | \mathbf{y})$$

where \mathbf{y} are the parameters of the observation model that could include tag return rates and tag reading error.

The seven-fishery model presented earlier could easily be modified to represent a general geographic square model, where any size geographic units were used. For tuna, a 5° square basis might be appropriate. Many tuna agencies, particularly the Inter-American Tropical Tuna Commission (IATTC) have databases on tagging, recaptures, and efforts that may well be amenable to this type of analysis.

A number of alternative models would be possible with this type of geographic data. The simplest would be a diffusion model with one movement parameter and one survival rate to estimate. In discrete form (and ignoring harvesting) this could be written:

$$(13) \quad N_{i,j,t+1} = N_{i,j,t}s - 4sdN_{i,j,t} + sd[N_{i+1,j,t} + N_{i-1,j,t} + N_{i,j+1,t} + N_{i,j-1,t}]$$

where d is the proportion of individuals that move from one square to an adjacent square each time period and s is survival from natural mortality. This representation is a finite difference approximation to the normal two dimensional diffusion equations described extensively in Okubo (1980) and used to model tuna movement in Hilborn (1987):

$$(14) \quad \frac{\partial N}{\partial t} = -(1-s)N + D \left[\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} \right]$$

The parameter D is the analog of d from the finite difference equations.

For discrete square models, a slightly more refined movement model would be one in which there were four movement parameters to be estimated: the proportions moving to the square to the north, the square south, the square west, and the square east. This would include the possibilities of diffusion connected with a general drift model. Note that both of these geographic movement models would require very few parameters to be

estimated relative to the more complex movement model used for the seven-fishery model.

A further extension of the geographic square model (or any movement model) would be to allow for temporal or spatial stratification. Thus, in the seven-fishery model we could have stratified the year into two periods and used different movement probabilities for each time of year. This would produce a rather large number of parameters in the seven-fishery model, but for the simpler geographic square models discussed in the preceding section, one could include seasonal variation in movements without too many parameters. This type of analysis might be most useful for truly migratory species that engage in an annual north-south movement. North Pacific albacore tuna (*Thunnus alalunga*) are a possible example.

Spatial stratification would also be possible. For instance, in the eastern Pacific, fish marked close to land appear to move less than those marked far at sea. One might use the general directional diffusion model discussed in the preceding section, but spatially stratify the geographic area into three to six areas with different movement parameters. While the calculations might remain on a 1 or 5° basis, the movement parameters would be stratified on a larger and biologically more meaningful scale.

Size-Specific Probabilities of Capture

In some fisheries, particularly gillnet fisheries, the probability of capture is size-specific. Nick Dow (Victoria Department of Fisheries, Queenscliff, Victoria, Australia) has been using similar models to those described in this paper, where he estimates simultaneously the size selectivity, growth, movement, natural and fishing mortality. This requires treating each individual fish as a tag group, since each fish has a unique size and will therefore have unique capture and growth probabilities.

Tag Mortality, Shedding, and Nonreporting and Confounding of Parameters

Three common problems with tag analysis are the potential for tagging mortality, tag shedding, and nonreporting of tags. In the models described above we have ignored these problems. However, they can be incorporated in a very straightforward manner.

Tag mortality could be considered by adding one additional parameter: the number of tags released and followed through the simulation would be the number actually released times 1.0 minus the tagging mortality rate.

Tag shedding could be considered by adding an additional parameter, yet it is normally indistinguishable from natural mortality. The method to determine tag shedding is to use double tag experiments, which could also be easily added to our analytic framework. There will then be the likelihood of recovering one and two tags from doubly tagged individuals. Nonreporting of tags could also be an additional parameter to be estimated, but this is algebraically indistinguishable from tagging mortality. Both are normally (Kleiber et al. 1983) estimated as a single parameter.

While it is conceptually and practically straightforward to add tag mortality, shedding, and nonreporting, these parameters will normally be highly or totally confounded with fishing mortality, natural mortality, and movement. Similarly it is theoretically possible to estimate temporal variation in natural mortality or catchabilities, but in practice these may prove difficult if not to impossible to identify.

TABLE 3. Parameter correlations for two-fishery model estimated by inverse Hessian method.

	PNG- q q_1	SOL- q q_2	PNG-PNG- p_{11}	PNG-SOL p_{12}	SOL-PNG p_{21}
SOL- q q_2	-0.09				
PNG-PNG p_{11}	-0.83	0.11			
PNG-SOL p_{12}	0.22	-0.07	-0.26		
SOL-PNG p_{21}	-0.09	0.20	0.01	-0.08	
SOL-SOL p_{22}	-0.09	-0.83	-0.10	-0.12	-0.25

Confidence Intervals

We would clearly like some idea of our confidence in the parameters estimated from these analyses. I can suggest two possible approaches. The first is the use of the inverse-Hessian matrix as a linear approximation to the confidence intervals (Bard 1974). This method looks at the derivative of the likelihood function with respect to the parameters. If the likelihood is insensitive to changes in the parameter, the confidence interval for that parameter is large. Table 3 shows the parameter correlations for the parameters estimated in our two-fishery model. There is reasonably little parameter confounding (high correlations), except between the q values and the proportion staying in each fishery (-0.83 in each case). The variance-covariance matrix that emerges from this approach does not directly provide confidence intervals, although it does provide a good feeling for our understanding of parameter values and a computational framework for estimating true intervals.

The second possible approach would be some form of bootstrap analysis (Efron 1982). These methods appear to be highly reliable but are computationally intense and would require a few hundred runs of the estimation procedure. This may not be practical for most of the movement analysis we would want to consider because each run involves considerable computation time (several hours on a minicomputer). However, for the two-fishery model, such methods would be practical and would provide an easily understood confidence bound that would not depend upon quadratic approximation as does the inverse-Hessian method.

Analysis of Residuals

We can examine the deviations from our model by using the "deviance" of McCullagh and Nelder (1983). The deviance in any release group, space, time stratum is defined as:

$$(15) \text{ deviance}_{iat} = -2[l(R_{iat}|\hat{R}_{iat}) - l(R_{iat}|R_{iat})]$$

where $l(x|y)$ denotes the log likelihood of x given y , and in this case is the simple Poisson likelihoods as given in equation (6). The deviance is analogous to the residuals of traditional least squares and looks like an observed minus an expected likelihood squared. A large value indicates a poor fit of the model to an individual data point.

Table 4 shows the observed, and predicted number of tags recovered. The biggest individual deviance occurred at month 7 for tagged individuals released in PNG and recovered in PNG, where only six tags were recovered yet 23.66 were expected. There appear to be no major systematic deviations between observed and expected, although months 14 to 21 for tags released and recovered in SOL show considerable variability.

Discussion

The method described in this paper is easily implemented for new data sets. Biologists working with mark-release data need not be intimidated by the apparent complexity of the statistics involved in analyzing tag movements. Almost universal access to microcomputers means that any scientist involved in the analysis of a tagging program should be in a position to use these computationally intense methods to analyze tag movements.

There are a number of obvious steps in testing and developing this approach. Monte-Carlo trials need to be used extensively to determine the performance of the methods against known answers. Secondly, the approach here needs to be compared to other possibilities, particularly explicitly derived multinomial models. The approaches used in determining confidence limits also need to be developed and tested. Particularly, bootstrapping should be tried and evaluated by testing against simulated data.

At a more general level, there is the question of experimental design required to determine exchange of fish between spatial strata. I found the estimates of movement rates to be remarkably unconfounded. The experimental design involved in analysis of movement data is normally extremely powerful. The best experimental design would involve mark and release in each area, and fishing effort data by time for each area. In the seven-country model I met all of these conditions except the last. I was unable to distinguish between natural mortality and movement outside of the seven countries considered.

In contrast, Schwarz and Arnason (1990) examined a special case, in which the tagging areas were not the same as the recapture areas, and found enormous confounding of parameters. They conclude that independent estimates of population size are required. I believe their conclusions are quite particular to their special cases and have little generality. When fish are marked and released in all strata, and we do not attempt to estimate time specific or area specific natural mortality rates, the movement and fishing mortality rate parameters are quite estimable. Time specific fishing mortality rates are estimable only if we make an assumption relating fishing effort to mortality.

It would be quite useful to have a systematic exploration of what experimental designs are necessary to estimate specific parameters. This could either be done analytically or numerically. Until such work is done, I would recommend marking and recapture in each area, and estimation of fishing effort or mortality in each area.

The analysis of skipjack movement presented here must be considered very preliminary. I have used the data as examples of what can be done, rather than as a definitive analysis of this data set. I have not considered tagging mortality, or nonreporting of tags, and any thorough analysis of this data base

TABLE 4. Observed and expected tag recoveries between PNG and SOL and the associated deviance. Top half of table are data for fish tagged in PNG, the bottom half are data for fish tagged in the Solomon Islands.

Month	Tagged in PNG					
	Recovered in PNG			Recovered in SOL		
	Observed	Expected	Deviance	Observed	Expected	Deviance
1				0	.5764	1.1528
2				0	1.0102	2.0203
3	208	179.7927	4.2109	0	1.2816	2.5633
4	100	103.4803	.1185	1	1.3675	.1090
5	58	67.2392	1.3321	0	1.4437	2.8873
6	27	30.3490	.3839	2	1.4265	.2047
7	6	23.6661	18.8648	0	1.2808	2.5616
8	3	4.7817	.7663	0	1.2424	2.4849
9	0	.0000	.0000	2	1.0691	.6434
10	0	.0000	.0000	0	.0643	.1287
11	1	3.5775	2.6056	0	.0000	.0000
12	5	4.5019	.0532	0	.1257	.2513
13	1	3.7010	2.7848	1	.5026	.3811
14	2	2.4902	.1035	0	.6236	1.2472
15	3	1.7459	.7399	2	.5902	2.0620
16	7	1.1832	13.2540	1	.5258	.3373
17	5	.7002	11.0587	1	.4884	.4101
18	0	.4032	.8065	0	.4312	.8624
19	0	.2690	.5381	1	.3860	.6759
20	0	.1235	.2469	3	.3625	7.4056
21	0	.0157	.0315	1	.2526	1.2571
22	0	.0000	.0000	0	.0000	.0000
23	0	.0435	.0870	0	.0000	.0000
24	0	.0623	.1246	0	.1035	.2069
25	0	.0484	.0968	0	.2021	.4043
26	0	.0354	.0707	0	.1982	.3964
27	1	.0301	5.0651	0	.1810	.3621
28	1	.0208	5.7860	0	.1647	.3294
29	0	.0139	.0277	0	.1468	.2935
30	0	.0089	.0178	0	.1291	.2583
31	0	.0022	.0045	0	.1157	.2313
32	0	.0004	.0009	0	.9678	.1356

Tagged in SOL						
	Recovered in SOL			Recovered in PNG		
	Observed	Expected	Deviance	Observed	Expected	Deviance
14	9	21.0244	8.7765	0	.5292	1.0584
15	28	19.8002	3.0051	0	.8709	1.7418
16	25	17.5742	2.7706	2	1.0465	.6838
17	9	16.2799	3.8911	0	.9814	1.9628
18	8	14.3465	3.3480	1	.8391	.0291
19	20	12.8238	3.4248	0	.7941	1.5883
20	16	12.0299	1.1860	0	.4999	.9997
21	18	8.3766	8.2907	1	.0085	3.1058
22	0	.0000	.0000	0	.0000	.0000
23	0	.0000	.0000	2	.3729	3.4639
24	5	3.4255	.6329	0	.6540	1.3080
25	5	6.6915	.4690	0	.6139	1.2278
26	11	6.5591	2.4931	1	.5328	.3248
27	4	5.9904	.7499	1	.5281	.3332
28	6	5.4496	.0538	0	.4161	.8321
29	2	4.8558	2.1635	0	.3093	.6186
30	0	4.2725	8.5451	0	.2166	.4332
31	0	3.8259	7.6518	0	.0058	.1164
32	1	2.2424	.8697	0	.0012	.0246

would need to do so. The lack of fishing effort for countries other than PNG and SOL may preclude a definitive analysis. Now that we know that it is possible to estimate movements from tagging data in a rigorous and reliable way, we must make sure that future tagging studies make every effort to collect the required data, which must include effort data in the recovery locations.

Acknowledgements

The general approach described here can be traced back to Ishii (1979), but has been modified and reinvented a number of times in various forms by John Sibert, (R. Hilborn), Nick Dow, Pierre Kleiber, Marc Mangel, and Jose Goularte. I thank each of them for the ideas they have provided. This work was initiated while I worked for the South Pacific Commission (SPC), and all of the tuna tagging data was in from published SPC data bases. John Schnute and John Skalski provided most helpful comments on the manuscript, and Marianne McClure assisted in manuscript preparation.

References

- ARNASON, A. N. 1972. Parameter estimation from mark-recapture experiments subject to migration and death. *Res. Pop. Ecol.* 12: 97-113.
1973. The estimation of population size, migration rates, and survival in a stratified population. *Res. Pop. Ecol.* 13: 1-8.
- BARD, Y. 1974. Nonlinear parameter estimation. Academic Press, New York. NY. 341 p.
- BAYLIFF, W. H. 1979. Migrations of yellowfin tuna in the eastern Pacific ocean as determined from tagging experiments initiated during 1968-74. *Int.-Am. Trop. Tuna Comm. Bull.* 17: 447-506.
- BURNHAM, K. P., D. R. ANDERSON, G. C. WHITE, C. BROWNIE, K. H. POLLOCK. 1987. Design and analysis methods for fish survival experiments based on release-recapture. *Am. Fish. Soc. Monogr.* No. 5. Bethesda, MD.
- CHAPMAN, D. G. 1948. A mathematical study of confidence limits of salmon populations calculated from sample tag ratios. *Int. Pac. Salmon Comm. Bull.* 2: 68-85.
- CORMACK, R. M. 1981. Loglinear models for capture-recapture experiments in open populations. p. 217-235. *In* R. W. Hiorns and D. Cooke, [ed.] *The mathematical theory of the dynamics of biological populations*. Academic Press, London.
- DARROCH, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. *Biometrika* 48: 241-260.
- EFRON, B. 1982. The jackknife, bootstrap and other resampling plans. *SIAM Review* No. 38. Society for Industrial and Applied Mathematics. 92 p.
- FINK, B. D., and W. H. BAYLIFF. 1970. Migrations of yellowfin and skipjack tuna in the eastern Pacific ocean as determined by tagging experiments, 1952-1964. *Int.-Am. Trop. Tuna Comm. Bull.* 15: 7-63.
- HILBORN, R. 1987. Spatial models of tuna dynamics in the western Pacific: is international management of tuna necessary? p. 276-286. *In* T. L. Vincent, Y. Cohen, W. J. Grantham, G. P. Kirkwood, and J. M. Skowronski [ed.] *modelling and management of resources under uncertainty*. Lecture notes in biomathematics, No. 72. Springer-Verlag, Berlin.
- HUNTER, J. R., A. W. ARGUE, W. H. BAYLIFF, A. E. DIZON, A. FONTENEAU, D. GOODMAN, and G. R. SECKEL. 1986. The dynamics of tuna movements: an evaluation of past and future research. *FAO Fish. Tech. Paper* No. 277. 78 p.
- ISHII, T. 1979. Attempt to estimate migration of fish population with survival parameters from tagging experiment data by the simulation method. *Inv. Pesq.* 43: 301-317.
- JONES, R. 1959. A method of analysis of some tagged haddock returns. *J. Cons. CIEM*, 25: 59-72.
1976. The use of marking data in fish population analysis. *FAO Fish. Tech. Paper* No. 153. 42 p.
- KEARNEY, R. E. 1983. Assessment of the skipjack and baitfish resources in the central and western tropical Pacific Ocean: a summary of the skipjack survey and assessment programme. South Pacific Commission, Noumea, New Caledonia.
- KLEIBER, P., A. W. ARGUE, and R. E. KEARNEY. 1983. Assessment of skipjack (*Katsuwonus pelamis*) resources in the central and western Pacific by estimating standing stock and components of population turnover from tagging data. Tuna and Billfish Assessment Programme, South Pacific Commission, Noumea, New Caledonia, Tech. Rep. No. 8. 38 p.
- MCCULLAGH, P., and J. A. NELDER. 1983. Generalized linear models. Chapman and Hall, London. 261 p.
- OKUBO, A. 1980. Diffusion and ecological problems: mathematical models. Springer-Verlag, Berlin. 254 p.
- SCHAEFER, M. B., B. M. CHATWIN, and G. C. BROADHEAD. 1961. Tagging and recovery of tropical tunas, 1955-1959. *Int.-Am. Trop. Tuna Comm. Bull.* 17: 447-506.
- SCHNUTE, J. 1982. A manual for easy nonlinear parameter estimation in fishery research, with interactive microcomputer programs. *Can. Tech. Rep. Fish. Aquat. Sci.* 1140: xvi + 115 p.
- SEBER, G. A. F. 1982. The estimation of animal abundance and related parameters. 2nd ed. MacMillan Publishing Co. New York, NY. 654 p.
- SIBERT, J. R. 1984. A two-fishery tag attrition model for the analysis of mortality, recruitment, and fishery interaction. Tuna and Billfish Assessment Programme, South Pacific Commission, Noumea, New Caledonia, Tech. Rep. No. 13. 27 p.
- SCHWARZ, C. J. 1988. Postrelease stratification and migration models in band-recovery and capture-recapture models. Ph.D. thesis, University of Manitoba, Winnipeg, Man. Canada.
- SCHWARZ, C. J., and A. N. ARNASON. 1990. Use of tag-recovery information in migration and movement studies. *Proceedings of the American Fisheries Society workshop on tagging* (In press).