THE USE OF AGE-DEPENDENT CAPTURE-RECAPTURE MODELS IN FISHERIES RESEARCH

K. H. POLLOCK
Technical Director Southeastern Cooperative
Fish and Game Statistics Project
Department of Statistics
North Carolina State University
Box 5457
Raleigh, North Carolina 27650 U.S.A.

R. H. K. MANN
Freshwater Biological Association
River Laboratory, East Stoke
Wareham, Dorset BH20 6BB
England

Abstract

An age-dependent generalization of the Jolly-Seber capture-recapture model for open populations is presented. The model is then extensively illustrated using data from a study on pike in Dorset. These data show evidence of differential survival between pike aged one year and those aged two or more years. This model has a useful potential in fisheries research in which the assumption that survival and capture probabilities are independent of age is often unrealistic.

The Jolly-Seber model (Jolly 1965; Seber 1965) is a stochastic capture-recapture model for an open population of animals. It embodies the assumption that the probabilities of survival and capture probabilities are independent of age. A recent paper by Pollock (1981a) described a model which allows identifiable age-classes to have different survival and capture probabilities.

In the present paper the age-dependent model is described briefly and then extensively illustrated using a data set on pike (Esox lucius L.) collected from a river in Dorset, England (Mann 1980). This is followed by a discussion on the utility of the model in fisheries research.

Age-Dependent Model

Here only a brief description of the model is given; for full details consult Pollock (1981a).

Description

We assume that there is one capture period per "year" for K "years". A "year" is used to represent the period of time an animal remains in an age-class and may not always represent a calendar year although it does in our example. There are $(\ell+1)$ distinguishable age-classes of animals ranging from 0 up to ℓ (or more) years of age and each age-class moves forward one class each "year". We further assume that each age-class has a different capture probability in the ith sample $(p_i^{(0)}, p_i^{(1)},\ldots,p_i^{(\ell)})$ and a different survival probability from the ith to the (i+1)th sample $(\phi_i^{(0)}, \phi_i^{(1)},\ldots, \phi_i^{(\ell)})$. Also immigration or emigration can occur for each age-class of the population, but births can only occur into the youngest (age 0) group. Thus, when referring to "survival", we really mean those animals which have not died or emigrated. Similarly, when referring to "recruitment", we really mean births and immigration for young animals (age = 0), but only immigration for the older animals (age > 0).

As well as the survival probabilities $(\phi_i^{(0)}, \phi_i^{(1)},\ldots, \phi_i^{(\ell)})$ described above the parameters to be estimated are the numbers of animals of each age-class in the population at each sampling time ($N_i^{(0)}, \ldots, N_i^{(\ell)}$), and the number of marked animals ($M_i^{(1)}, M_i^{(2)}, \ldots, M_i^{(\ell+1)}$), of each age-class from $1, 2, \ldots, \ell+1$. The $M_i^{(v)}$ parameters require further discussion. There are no marked young animals ($M_i^{(0)} \equiv 0$) because after one "year" they will have moved into the next age-class (age = 1). It is necessary to identify marked animals up to age ($\ell+1$) ($M_i^{(\ell+1)}$) to estimate the survival rate of the age ℓ animals ($\phi_i^{(\ell)}$) as we shall see later.

Parameter Estimation

The following statistics were used to calculate the various parameter estimates:

$m_i^{(v)}$ is the number of marked animals of age v captured in the ith sample.

$n_i^{(v)}$ is the number of animals of age v captured in the ith sample.

$R_i^{(v)}$ is the number of animals of age v released from the ith sample.

$r_i^{(v)}$ is the number of the $R_i^{(v)}$ which are captured again at least once after the ith sample.

$Z_i^{(v)}$ is the number of marked animals of age v caught before time i, not caught at time i, and caught again later. Note that this is not the usual fisheries definition for Z.

$T_i^{(v)}$ is the number of marked animals of age v which are caught at time i or after time i.

Except for some simple extensions the estimators have a structure similar to the Jolly-Seber estimators (Seber 1973).

(i) Marked population size

$$\hat{M}_i^{(v)} = m_i^{(v)} + \frac{R_i^{(v)} Z_i^{(v)}}{r_i^{(v)}} \qquad \text{for } v = 1,2,\ldots,\ell-1$$

$$\hat{M}_i^{(\ell)} + \hat{M}_i^{(\ell+1)} = m_i^{(\ell)} + \frac{R_i^{(\ell)} Z_i^{(\ell)}}{r_i^{(\ell)}}$$

$$\hat{M}_i^{(\ell)} = \frac{T_i^{(\ell)}}{T_i^{(\ell)} + T_i^{(\ell+1)}} \qquad (\hat{M}_i^{(\ell)} + \hat{M}_i^{(\ell+1)})$$

There is a similar expression for $\hat{M}_i^{(\ell+1)}$ and the estimators are defined only for $i = 2,\ldots,K-1$.

(ii) Total population size

$$\hat{N}_i^{(v)} = \frac{n_i^{(v)} \hat{M}_i^{(v)}}{m_i^{(v)}} \qquad \text{for } v = 1,2,\ldots,\ell-1$$

$$\hat{N}_i^{(\ell)} = \frac{n_i^{(\ell)} (\hat{M}_i^{(\ell)} + \hat{M}_i^{(\ell+1)})}{m_i^{(\ell)}}$$

Notice that it is not possible to estimate $N_i^{(0)}$ because there are no marked animals of age 0. All the estimators are defined only for $i = 2,\ldots,K-1$.

(iii) "Survival" rates

$$\hat{\phi}_i^{(v)} = \frac{\hat{M}_{i+1}^{(v+1)}}{\hat{M}_i^{(v)} - m_i^{(v)} + R_i^{(v)}} \qquad \text{for } v = 0,1,\ldots,\ell-1$$

$$\hat{\phi}_i^{(\ell)} = \frac{\hat{M}_{i+1}^{(\ell+1)}}{\hat{M}_i^{(\ell)} + \hat{M}_i^{(\ell+1)} - m_i^{(\ell)} + R_i^{(\ell)}}$$

Notice the need to estimate $M_{i+1}^{(\ell+1)}$ to be able to estimate $\phi_i^{(\ell)}$ as discussed earlier, and note that all estimators are defined only for $i = 1,\ldots,K-2$.

Approximate variances and covariances of the estimators are given by Pollock (1981a) and will not be presented here.

Testing Between Models

It is important to be able to test if survival and capture rates are different between age-categories. Pollock (1981a) showed that this is possible using a series of contingency table chi-square tests based on sufficient statistics.

For the first (K-1) samples we have 2 x (ℓ+1) contingency tables of the following form:

Age class	Recaptured after sample i	Not recaptured after sample i	Animals released in sample i
0	$r_i^{(0)}$	$R_i^{(0)} - r_i^{(0)}$	$R_i^{(0)}$
1	$r_i^{(1)}$	$R_i^{(1)} - r_i^{(1)}$	$R_i^{(1)}$
.	.	.	.
.	.	.	.
.	.	.	.
ℓ	$r_i^{(\ell)}$	$R_i^{(\ell)} - r_i^{(\ell)}$	$R_i^{(\ell)}$
Overall	r_i	$R_i - r_i$	R_i

Each table gives rise to an approximate chi-square statistic with ℓ degrees of freedom. In effect we are comparing the recapture rates of the (ℓ+1) different age-classes of animals.

For samples 2 to K-1 we have additional 2 x ℓ contingency tables of the following form:

Age class	Captured in sample i	Not captured in sample i	Marked animals captured in or after i
1	$m_i^{(1)}$	$T_i^{(1)} - m_i^{(1)}$	$T_i^{(1)}$
2	$m_i^{(2)}$	$T_i^{(2)} - m_i^{(2)}$	$T_i^{(2)}$
.	.	.	.
.	.	.	.
.	.	.	.
ℓ	$m_i^{(\ell)}$	$T_i^{(\ell)} + T_i^{(\ell+1)} - m_i^{(\ell)}$	$T_i^{(\ell)} + T_i^{(\ell+1)}$
Overall	m_i	$T_i - m_i$	T_i

Note that there are no marked animals of age 0, and that this series of tables is only used if there are more than two age-classes. Each table gives rise to an approximate chi-square statistic with (ℓ-1) degrees of freedom. Actually we are comparing the capture rates of known groups of marked animals of each age-class.

As these chi-square statistics are conditionally independent of each other we can obtain an overall chi-square test statistic with ℓ(K-1) + (ℓ-1) (K-2) degrees of freedom by adding the individuals chi-square statistics. In some cases it may be necessary to pool some cells in the tables in which case the degrees of freedom will be reduced.

Example

Biological Background

Here we illustrate the methodology with a capture-recapture study on pike (*Esox lucius* L.) carried out by (Mann 1980) on the River Frome, Dorset, England between the winters of 1971-2 and 1978-9. Electrofishing was carried out each winter between November and March. Each fish was individually identified by attaching a numbered, metal tag to the maxilla. Tag loss was estimated to be approximately 0.095 per year (S.E. 0.027).

In this analysis we just consider if a fish was captured at least once during a winter. Also we divide the fish into just two age categories because of the small numbers of older fish which were marked. The first age-class ($v = 0$) are fish of age 1 year and the second age-class ($v = 1$) are fish of age at least 2 years. Young of the year fish could not be marked using the metal tags and are not considered in this analysis.

As electrofishing is size-selective we would expect older fish to have higher capture rates. It also might be expected that older fish have higher survival rates. Therefore this data set should be a useful illustration of when the age-dependent capture-recapture model should be used.

Estimation of Population Parameters

There were 295 fish captured in at least one of the 8 years of the study. Each fish's capture history formed the basis of our analysis but is not presented here. The basic summary statistics necessary to calculate our estimates are given in Table 1.

(Table 1 to appear here)

The marked and total population size parameters are given in Table 2. $M_i^{(1)}$ is the marked age 2 fish, $M_i^{(2)}$ is the marked age 3 or more fish, $M_i^{(1)} + M_i^{(2)}$ is the marked age 2 or more fish and $N_i^{(1)}$ is the total population size of fish age 2 or more. It is not possible to estimate $N_i^{(0)}$, the fish of age 1 because there are no marked fish of age 1 ($M_i^{(0)} \equiv 0$).

(Table 2 to appear here)

Notice that this is a small population of fish and that the precision of the estimates is not very good. Notice also that the precision of the estimates is much worse at the beginning and end of the chain of estimates.

The estimates of survival rate estimates for the age 1 fish ($\phi_i^{(0)}$) and age 2 or more fish ($\phi_i^{(1)}$) are given in Table 3. First the model estimates are given, and then adjustment for a 9.5% tag loss (S.E. 2.7%) is made.

(Table 3 to appear here)

This is a very important table. There is a lot of variability in the estimates with those at the ends of the chain less precise than those in the middle. However, notice that for every one of the six periods where estimates are possible the older fish have a higher survival rate. Therefore, despite the variability due to the small sample sizes this population shows substantial evidence of age-dependent survival rates.

Model Testing

In Table 4 we first tested the age-dependent model versus the age-independent model (Jolly-Seber) using the contingency table approach described earlier. The overall chi-square test statistics with seven degrees of freedom is 6.99 which is not significant (p = 0.45).

(Table 4 to appear here)

This lack of significance surprised us considering the strong indication of age-dependent survival shown in Table 3. The problem is that the small sample sizes mean the test has low power. Therefore we compared the recovery rates of the age 1 fish ($r_i^{(0)}/R_i^{(0)}$) and the age 2 or more fish ($r_i^{(1)}/R_i^{(1)}$) on which the test is based. In all cases except the last (i = 7) the older fish have higher recovery rates. We carried out a one tailed test of the equality of the two recovery rates using the normal approximation to the binomial for each value of i. The approximate significance levels of the tests (p_i) are given and also an overall chi-square test $\chi^2_{14} = \sum_{i=1}^{7} \chi^2_2 = \sum_{i=1}^{7} - 2 \log_e p_i$. This test has a probability value of 0.12 which although not significant does suggest the possibility of age-dependent survival and capture probabilities. We also computed the (binomial) probability of getting a result as extreme as six values out of seven being higher for the older fish if there was no age-dependence of survival and capture probabilities. This probability is 0.060 which also tends to cast doubt on the null hypothesis that survival and capture probabilities are independent of age.

Discussion

The capture-recapture method has a wide application for estimating population parameters. The reader is referred to Seber (1973) for a detailed description and to two recent review papers by Nichols et al. (1981) and Pollock (1981b).

The authors believe that the age-dependent capture model illustrated here is potentially very useful in fisheries research where a range of age-classes can often be identified. The example we present shows that age-dependent survival can be important.

The major problem with using a more complex model like this is that there are many parameters to estimate and there may be problems with large variances for estimators (low precision). Careful thought is needed when the study is designed to try and increase the sampling effort where possible. Precision may also be increased by reducing the number of parameters in the model. For example in some cases it may be feasible to assume that the survival rate for a particular age-category is constant over time. In our example this was part of the motivation for calculating the average survival rate for the two age groups of pike.

References

JOLLY, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration-stochastic model. Biometrika 52: 225-247.

MANN, R. H. K. 1980. The numbers and production of pike (Esox lucius) in two Dorset rivers. Journal of Animal Ecology 49: 899-915.

NICHOLS, J. D., B. R. NOON, S. L. STOKES, AND J. E. HINES. 1981. Remarks on the use of mark-recapture methodology in estimating avian population size. Pages 121-136 in C. J. Ralph and J. M. Scott, editors. Estimating the numbers of terrestrial birds. Studies in Avian Biology 6.

POLLOCK, K. H. 1981a. Capture-recapture models allowing for age-dependent survival and capture rates. Biometrics 37: 521-529.

POLLOCK, K. H. 1981b. Capture-recapture models: A review of current methods, assumptions, and experimental design. Pages 426-435 in C. J. Ralph and J. M. Scott, editors. Estimating the numbers of terrestrial birds. Studies in Avian Biology 6.

SEBER, G. A. F. 1965. A note on the multiple-recapture census. Biometrika 52: 249-259.

SEBER, G. A. F. 1973. The estimation of animal abundance and related parameters. Griffin, London, England.

Table 1.— Basic statistics used in the calculation of the age-dependent capture-recapture estimates

Year	i	$R_i^{(0)}$ [b]	$r_i^{(0)}$	$T_i^{(1)}$	$R_i^{(1)}$ [b]	$r_i^{(1)}$	$T_i^{(2)}$	$m_i^{(1)}$	$Z_i^{(1)}$
71-2	1	20	6	-	29	10	-	-	-
72-3	2	29	4	6	22	5	10	12	4
73-4	3	21	3	4	28	11	9	11	2
74-5	4	29	4	3	24	5	13	14	2
75-6	5	11	3	4	25	8	7	9	1
76-7	6	47	13	3	31	11	9	11	1
77-8[a]	7	22	4	13	38	3	12	24	2

[a] 1978-9 was just a recapture year. There were no new animals marked.

[b] There were no "losses on capture" so that $n_i^{(0)} = R_i^{(0)}$ and $n_i^{(1)} = R_i^{(1)}$.

Table 2.— Estimates of the marked and total population sizes with their standard errors in parentheses.

		Marked Population Sizes			Total Population Size
Year	i	Age 2	Age 3 or more	Age 2 or more	Age 2 or more
72-3	2	10.24 (4.1)	17.09 (6.3)	27.33 (9.4)	48.35 (17.7)
73-4	3	4.87 (1.2)	10.96 (2.2)	15.83 (3.0)	38.26 (8.1)
74-5	4	4.19 (1.7)	18.14 (5.4)	22.33 (6.5)	37.21 (11.3)
75-6	5	4.32 (1.1)	7.56 (1.7)	11.88 (2.6)	30.89 (7.0)
76-7	6	3.42 (0.8)	10.25 (1.8)	13.67 (2.3)	36.45 (6.3)
77-8	7	22.62 (10.1)	20.88 (9.3)	43.50 (18.6)	67.86 (29.2)

Table 3.— Estimates of the survival rate parameters and their standard errors in parentheses. Unadjusted and then adjusted for tag loss.

		Age of Fish		Age of Fish	
Year	i	1[a]	2 or more[a]	1[b]	2 or more[b]
71-2	1	0.51[c] (0.21)	0.59 (0.22)	0.56 (0.23)	0.65 (0.24)
72-3	2	0.17 (0.04)	0.28 (0.09)	0.19 (0.04)	0.31 (0.10)
73-4	3	0.20 (0.08)	0.55 (0.17)	0.22 (0.09)	0.61 (0.19)
74-5	4	0.15 (0.04)	0.22 (0.07)	0.17 (0.05)	0.24 (0.08)
75-6	5	0.31 (0.07)	0.36 (0.07)	0.34 (0.08)	0.40 (0.08)
76-7	6	0.48 (0.21)	0.62 (0.28)	0.53 (0.23)	0.68 (0.31)
77-8	7	-	-	-	-
Mean[d]		0.30 (0.05)	0.44 (0.07)	0.33 (0.06)	0.49 (0.08)

[a]These estimates are from the age-dependent model.

[b]These estimates are adjusted for a 9.5% annual tag loss. The standard errors are found using an approximation in Pollock (1981a).

[c]This survival rate is from Year 1 to Year 2.

[d]Arithmetic mean of yearly survival rates (unweighted).

Table 4.— Testing for Age-dependence of survival and capture probabilities

2 x 2 table							
Age 1 / Age 2	$r_i^{(0)}$ / $r_i^{(1)}$	$R_i^{(0)} - r_i^{(0)}$ / $R_i^{(1)} - r_i^{(1)}$	χ_1^2 Statistics	$r_i^{(0)}/R_i^{(0)}$ [a] / $r_i^{(1)}/R_i^{(1)}$	Approx. Normal Test Statistics	Sig. level (p_i)	$\chi_2^2 = -2\log_e p_i$
i = 1	6	14		0.30			
	10	19	0.11	0.34	0.33	0.3707	1.98
i = 2	4	25		0.14			
	5	17	0.69	0.23	0.83	0.2033	3.19
i = 3	3	18		0.14			
	11	17	3.68	0.39	1.92	0.0274	7.19
i = 4	4	25		0.14			
	5	19	0.46	0.21	0.68	0.2483	2.79
i = 5	3	8		0.27			
	8	17	0.08	0.32	0.28	0.3897	1.88
i = 6	13	34		0.28			
	11	20	0.54	0.35	0.73	0.2327	2.92
i = 7	4	18		0.18			
	3	35	1.43	0.08	-1.20	0.8849	0.24
	Overall test: $\chi_7^2 = 6.99$ (p=0.45)				Overall test: $\chi_{14}^2 = 20.19$ (p=0.12)		

[a]The (Binomial) probability of getting 6 values out of 7 higher for the older fish is 0.06 if there was no age-dependence of survival and capture probabilities.