Instantaneous-Rates Tagging Models Allowing for Delayed Mixing of Newly Tagged Cohorts: Partial Year Tabulation of Recaptures

Lynn Waterhouse [a] [b] & John M. Hoenig [a]

[a] Virginia Institute of Marine Science, College of William and Mary, Post Office Box 1346, Gloucester Point, Virginia, 23062, USA

[b] Department of Statistics, Pennsylvania State University, 326 Thomas Building, University Park, Pennsylvania, 16802-2111, USA

Available online: 02 Dec 2011

PLEASE SCROLL DOWN FOR ARTICLE

ARTICLE

Instantaneous-Rates Tagging Models Allowing for Delayed Mixing of Newly Tagged Cohorts: Partial Year Tabulation of Recaptures

Lynn Waterhouse[1] and John M. Hoenig*

Virginia Institute of Marine Science, College of William and Mary, Post Office Box 1346, Gloucester Point, Virginia 23062, USA

Abstract

Instantaneous-rates tagging models can be used to estimate natural and fishing mortality rates from multiyear tagging studies in which cohorts of tagged animals are released at the start of each year. The models can include additional parameters to account for a delay in the mixing of newly tagged animals with previously tagged animals. Here a new model, referred to as the partial-year tabulation model for delayed mixing (delayed PYT model), is proposed in which the year is divided into parts and tag returns are tabulated by parts of the year rather than a full year. This is beneficial when there is delayed mixing because it achieves greater precision than annual tabulation and it provides estimates of the instantaneous rate of fishing mortality in the first year, which cannot be obtained when tag returns are tabulated by full year. The new model can be used at little or no extra cost. The superiority of the delayed PYT model is demonstrated through Monte Carlo simulation.

Multiyear tagging studies are often used to obtain estimates of survival rates and instantaneous rates of fishing and natural mortality. Modern, multiyear tagging models were first described by Seber (1970) and Robson and Youngs (1971). Brownie et al. (1978, 1985) popularized the approach by describing a suite of models that enable the user to fit a variety of models with age- and year-specific survival rates; these models are now commonly referred to as Brownie models.

The biggest difficulty in conducting tagging studies is to ensure that the tagged population thoroughly mixes with the population at large so that the tagged and untagged populations experience the same mortality rates. Unfortunately, the Brownie models cannot accommodate a lack of mixing of newly tagged animals into the population (Youngs and Robson 1975). Hoenig et al. (1998a, 1998b) used additional information about the timing of the fishery within the year to reparameterize the Brownie models in terms of instantaneous rates of fishing and natural mortality. This enables one to estimate the components of mortality. Significantly, this approach can accommodate delayed mixing of newly tagged animals into the untagged population (Hoenig et al. 1998b).

The general structure of a multiyear tagging study is as follows: At the start of each of several successive time periods a sample of the population, termed a cohort, is captured, tagged, and released. Recaptures of tagged animals are tabulated by period, which is typically a year. In this paper annual tagging periodicity is assumed. The data consist of counts r_{ij} that are the realizations of (i.e., observations on) the random variables R_{ij}, that is, the r_{ij} are the observed numbers of recaptures of animals tagged in year i ($i = 1, 2, \ldots, I$) and recaptured in year j ($j = i, \ldots, J$, with $J \geq I$). The estimates of the rates of natural and fishing mortality are calculated using full years of data.

For models parameterized in terms of instantaneous rates, parameters can be included to allow for a delay in the mixing

---

of the newly tagged animals with the untagged and previously tagged population (Hoenig et al. 1998b). One of the Hoenig et al. (1998b) models allows for the newly tagged animals to be mixed after a full year at liberty, and another formulation assumes that mixing is complete after a partial year. For both models, an estimate of the normal instantaneous rate of fishing mortality during the first year of the study is not possible due to the confounding of parameters.

If newly tagged animals fully mix with the population in less than a year, the tabulation of recoveries by year results in a loss of efficiency (in terms of larger standard errors for every parameter) over the theoretical optimum model. In this paper, we present a new model, the partial year tabulation model for delayed mixing (delayed PYT model), in which recoveries are tabulated by portion of the year. The year is split into parts such that in the first part of the year, part (a), the newly tagged animals are not fully mixed with the population at large and in the second part of the year, part (b), the newly tagged animals have fully mixed with the population at large. The data consist of counts r_{ijk}, which are the realizations of the random variables R_{ijk}, the observed numbers of recaptures of animals tagged in year i ($i = 1, 2, \ldots, I$) and recaptured in part k ($k \in [a, b]$) of year j ($j = i, \ldots, J$ with $J \geq I$).

The performance of the delayed PYT model is evaluated by Monte Carlo simulation. The new model is compared with the models of Hoenig et al. (1998a, 1998b). We show that, by making better use of the available information, we can obtain an estimate of the instantaneous rate of fishing mortality in the first year and achieve an increase in the overall efficiency of estimation.

INSTANTANEOUS RATES TAGGING MODELS

Hoenig et al. (1998a) formulated the Brownie model in terms of instantaneous rates of fishing and natural mortality given information on the timing of the fishery. The Brownie models are parameterized in terms of year-specific annual survival and tag recovery rates. The survival rate, denoted S_j, is defined as the expected fraction of the population alive at the start of year j that is still alive at the end of the year. The survival rate in year j can be expressed in terms of the instantaneous rates of fishing (F_j, year^{-1}) and natural (M, year^{-1}) mortality as

$$S_j = \exp(-F_j - M). \tag{1}$$

This holds regardless of the relative timing of the fishing and natural mortality.

The tag recovery rate in year j, f_j, can be expressed as a product of its components (Pollock et al. 1991; Hoenig et al. 1998a), that is,

$$f_j = \phi u_j \lambda, \tag{2}$$

where ϕ is a combination of the short-term survival rate from tagging and the short-term probability of tag retention, u_j is the exploitation rate (the expected fraction of the population alive at the start of year j that dies due to harvesting during the year), and λ is the tag reporting rate (the probability that a tag will be reported if the fish is recaptured). The parameters ϕ and λ can be estimated internal or external to the tagging study. When estimating ϕ and λ, it is best to incorporate any external information on the reporting rate into the likelihood equation to ensure that the resulting estimates of mortality rates reflect the uncertainty in the reporting rate (Hoenig et al. 1998b). In this paper, for simplicity, we assume that ϕ and λ are known exactly.

As shown below, it is necessary to know the timing of the harvest during the year in order to relate the exploitation rate to the instantaneous rates of fishing and natural mortality. Hoenig et al. (1998a) provided a general model formulation which allows for an arbitrary fishing pattern throughout the year. Here, for expository purposes we use a simpler formulation. When fishing intensity is constant throughout the year, exploitation in year j can be expressed as

$$u_j = u_j(F_j, M) = \frac{F_j}{F_j + M}[1 - \exp(-F_j - M)]. \tag{3}$$

The notation of u followed by parentheses is used for brevity to indicate u is a function of the parameters in parentheses. It is possible to fit a model in which M varies over time, but here it is assumed to be constant both within and over years. If the year is split into two parts (a and b), the exploitation rates for year j, u_{ja} and u_{jb}, are given by the equations

$$u_{ja} = \frac{p_j F_j}{p_j F_j + \gamma M}[1 - \exp(-p_j F_j - \gamma M)], \tag{4a}$$

and

$$u_{jb} = \frac{q_j F_j}{q_j F_j + (1 - \gamma)M}\{1 - \exp[-q_j F_j - (1 - \gamma)M]\}$$
$$\times \exp(-p_j F_j - \gamma M), \tag{4b}$$

where p_j is the proportion of the fishing mortality occurring in part (a) of the year, q_j is the proportion of the fishing mortality occurring in part (b) (note that $q_j = 1 - p_j$), γ is the fraction of the year that elapses in part (a), and F and M are as defined previously. The value of p is generally determined under the assumption that the fraction of fishing mortality occurring in part (a) is equal to the fraction of fishing effort that occurs in part (a) (if determining relative effort in the two parts of the year is difficult one can use relative catch, as suggested by Hoenig et al. 1998a). This assumption is generally not critical unless the natural mortality rate is high (Hoenig et al. 1998a). The exploitation rate in year j is the sum of the exploitation rates in

parts (a) and (b), that is,

$$u_j = u_{ja} + u_{jb}. \tag{5}$$

Note that under this formulation the fishing mortality is assumed to be constant within a time period but may vary among periods.

The recapture data (r_{ij}) can be explained by modeling the expected cell counts in terms of survival rates and exploitation rates which are themselves functions of the fishing and natural mortality rates. It is then a straightforward matter to construct the likelihood function for the data in order to estimate the parameters, as shown below.

Fully Mixed Case

When the tagged cohort is fully mixed with the population at large prior to the start of fishing, the expected recovery from the cohort tagged in year i and recovered in year j, $E(R_{ij})$, can be expressed as

$$E(R_{ij}) = \begin{cases} N_i \phi \lambda u_j, & i = j \leq I \\ N_i \phi \lambda u_j \prod_{h=i}^{j-1} S_h, & i < j \leq J, \end{cases} \tag{6}$$

where N_i is the number tagged and released in year i and the other parameters are as defined previously. There is an implicit category for all animals of a cohort that are never seen again, Y_i. This can be expressed as

$$Y_i = N_i - \sum_{j=i}^{J} r_{ij}. \tag{7}$$

Maximum likelihood estimation is generally used to obtain estimates for the model parameters. The likelihood function for instantaneous-rates models can be expressed as the product of the independent multinomial distributions of tag returns over time, with each tagged cohort giving rise to a multinomial distribution. The general form of the likelihood function Λ for product multinomial models is proportional to

$$\Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} P_{ij}^{r_{ij}} \right) \left(1 - \sum_{j=i}^{J} P_{ij} \right)^{Y_i}, \tag{8}$$

where P_{ij} is the cell probability of recovering a tagged animal in year j given that it was tagged in year i,

$$P_{ij} = \frac{E(R_{ij})}{N_i}, \tag{9}$$

and the other variables are as defined previously.

Delayed Mixing Lasting a Full Year

Hoenig et al. (1998b) extended the instantaneous-rates models to allow for the newly tagged animals to not fully mix with

the previously tagged population until a full year has elapsed. As a result of this delay, the newly tagged animals in year j experience an abnormal instantaneous rate of fishing mortality, F_j^*, for the first year at liberty. Values of F_j^* are generally not of interest but are included in the model to avoid bias in the estimates of the F_j and M. For a fishery with constant fishing intensity over the year, the exploitation rate, u_i^*, for the first year a cohort is at liberty is now expressed as

$$u_i^* = u_i^*(F_i^*, M) = \frac{F_i^*}{F_i^* + M}[1 - \exp(-F_i^* - M)]. \tag{10}$$

The asterisk is used here to denote that the rate of fishing mortality is abnormal, which causes the exploitation rate to be different from that of previously tagged animals. The expected recoveries from the cohort tagged in year i and recovered in year j can be expressed as

$$E(R_{ij}) = \begin{cases} N_i \phi \lambda\, u_j^*(F_j^*, M), & j = i \\ N_i \phi \lambda\, u_j(F, M)\, S_i^*, & j = i+1 \\ N_i \phi \lambda\, u_j(F, M)\, S_i^* \prod_{h=i+1}^{j-1} S_h, & j > i+1 \end{cases}, \tag{11}$$

where S_i^* is the abnormal survival rate in year i for the cohort tagged in year i,

$$S_i^* = \exp(-M - F_i^*), \tag{12}$$

and the other parameters are as described previously.

Hoenig et al. Model with Delayed Mixing Lasting Part of the Year

As shown by Hoenig et al. (1998b), if the newly tagged animals fully mix in less than a year, the year can be split into two parts, part (a) and part (b), such that animals are fully mixed into the population at large by the end of part (a). The exploitation rates in parts (a) and (b) for the first year each cohort is at liberty are given by u_{ia}^* and u_{ib}^*, respectively, which are given by the equations

$$u_{ia}^* = u_{ia}^*(F_i^*, M) = \frac{p_i F_i^*}{p_i F_i^* + \gamma M}$$
$$\times \left[1 - \exp(-p_i F_i^* - \gamma M) \right], \tag{13a}$$

$$u_{ib}^* = u_{ib}^*(F_i^*, F_i, M) = \frac{q_i F_i}{q_i F_i + (1 - \gamma)M}$$
$$\times \{1 - \exp[-q_i F_i - (1 - \gamma)M]\}$$
$$\times \exp(-p_i F_i^* - \gamma M). \tag{13b}$$

The asterisk on u_{ib}^* occurs because the exploitation rate in part (b) of year i depends on F_i^* (the abnormal fishing mortality rate in part [a] of year i). Note that multiplying F_i^* by p_i serves only to express F_i^* on the same basis as F_i, that is, F_i^* is the mortality rate that the unmixed cohort would experience if the newly tagged cohort experienced fishing mortality F_i^* during all of the fishing activity instead of during just part of it. This formulation allows one to judge the importance of nonmixing by making a direct comparison of F_i^* with F_i.

If newly tagged animals are assumed to be fully mixed after part of the year rather than the full year, the expected recaptures change to

$$E(R_{ij}) = \begin{cases} N_i \phi \lambda [u_{ia}^*(F_i^*, M) + u_{ib}^*(F_i^*, F_i, M)], \\ \quad i = j \\ N_i \phi \lambda [u_{ja}(F_j, M) + u_{jb}(F_j, M)] S_i^*, \\ \quad j = i + 1 \\ N_i \phi \lambda [u_{ja}(F_j, M) + u_{jb}(F_j, M)] S_i^* \prod_{h=i+1}^{j-1} S_h, \\ \quad j > i + 1 \end{cases}$$

$$(14)$$

where S_i^* is the abnormal survival rate for the first year each cohort is at liberty and is given by

$$S_i^* = \exp(-p_i F_i^* - q_i F_i - M). \tag{15}$$

Note that the abnormal survival rate in the year of tagging, S_i^*, is specified differently in equations (12) and (15). For equation (12), delayed mixing lasts a full year and thus animals must survive the abnormal instantaneous rate of fishing mortality, F_i^*, for the full first year at liberty, year i. In equation (15), animals experience an abnormal instantaneous rate of fishing mortality for only part of year i and the normal rate of fishing mortality, F_i, for the remainder of the year provided the data are tabulated yearly.

In the delayed-mixing model, in which the delay in mixing lasts part of the year, for year $i = 1$ there is no separate F_1 and F_1^*; instead there is only F_1^* because the parameters are not separably estimable (Hoenig et al. 1998b).

As before, maximum likelihood estimation can be used to estimate the parameters because the likelihood under delayed mixing is of the same form as in equation (8).

New Model: Partial-Year Tabulation for Delayed Mixing Lasting Part of a Year

A new model was developed for the case in which delayed mixing occurs but the newly tagged cohort fully mixes with the untagged and previously tagged population in less than a year. As before, the year is divided into parts such that in part (a) of the year the newly tagged cohort is not fully mixed with the population at large and by part (b) of the year the cohort is fully mixed with the population at large and experiencing the same rate of fishing mortality as previously tagged cohorts. The key feature of the delayed PYT model is that tag recaptures are tabu-

lated separately for parts (a) and (b) of every year. The expected recaptures from the cohort tagged in year i and recaptured in part (a) of year j can be expressed as

$$E(R_{ija}) = \begin{cases} N_i \phi \lambda u_{ia}^*(F_i^*, M), & i = j \\ N_i \phi \lambda u_{ja}(F_j, M) S_i^*, & j = i + 1 \\ N_i \phi \lambda u_{ja}(F_j, M) S_i^* \prod_{h=i+1}^{j-1} S_h, & j > i + 1 \end{cases},$$

$$(16a)$$

and for part (b) of year j

$$E(R_{ijb}) = \begin{cases} N_i \phi \lambda u_{ib}^*(F_i^*, F_i, M), & i = j \\ N_i \phi \lambda u_{jb}(F_j, M) S_i^*, & j = i + 1 \\ N_i \phi \lambda u_{jb}(F_j, M) S_i^* \prod_{h=i+1}^{j-1} S_h, & j > i + 1 \end{cases},$$

$$(16b)$$

where u_{ia}^* and u_{ib}^* are given by equations (13a) and (13b), respectively, u_{ja} and u_{jb} are given by equations (4a) and (4b), respectively, S_i^* is given by equation (15), and the remaining parameters are as defined previously. Note that equations (16a) and (16b) differ from equation (14) in that the tabulations are done by parts of the year for the delayed PYT model and by the full year in equation (14).

The likelihood is constructed similarly to equation (8), but with the year broken into parts (a) and (b) and the tag recaptures tabulated by parts. The recapture cell representing the tagged animals that are never seen again is given by

$$W_i = N_i - \sum_{j=i}^{J} \sum_{k \in \{a,b\}} r_{ijk}. \tag{17}$$

Thus, the likelihood can be expressed as

$$\Lambda \propto \prod_{i=1}^{I} \left(\prod_{j=i}^{J} \prod_{k \in \{a,b\}} P_{ijk}^{r_{ijk}} \right) \left(1 - \sum_{j=i}^{J} \sum_{k \in \{a,b\}} P_{ijk} \right)^{W_i}, \tag{18}$$

where P_{ijk} is the cell probability of recovering a tagged animal in part k ($k \in \{a, b\}$) of year j given that it was tagged in year i, that is,

$$P_{ijk} = \frac{E(R_{ijk})}{N_i}, \tag{19}$$

and the remaining parameters are as defined previously.

SIMULATIONS

Methods

Monte Carlo simulations were performed to evaluate the performance of the delayed PYT model and to compare the results with those from the partial-year delayed-mixing model

described by Hoenig et al. (1998b) using simulated data reflecting delayed mixing lasting part of the year. The simulations consisted of three years of tagging (years 1, 2, and 3) and four full years of recaptures, that is, recaptures for periods 1a, 1b, 2a, 2b, 3a, 3b, 4a, and 4b for the new model and years 1, 2, 3, and 4 for the Hoenig et al. (1998b) delayed-mixing model. One thousand animals were tagged each year and 2,000 product multinomial data sets were simulated for each scenario. The scenarios simulated delayed mixing lasting for half a year. That is, after half a year had passed the newly tagged animals were fully mixed with the untagged population as well as with the previously tagged animals. Note that ϕ and λ are confounded and cannot be estimated separately. For the purpose of the Monte Carlo simulations the product of $\phi\lambda$ was set to 1.0, indicating the tag reporting rate was 100%, no tags were shed immediately after tagging, and there was no short-term tag-induced mortality.

The parameters used to generate the data sets are given in Tables 1 and 2. Additional scenarios (not shown here) are presented in Waterhouse (2010). For each data set, the fraction of fishing effort—and thus, by assumption, the proportion of fishing mortality—in part (a) is equal to that in part (b) ($p = q = 0.5$). One set of simulations had rates of fishing mortality ranging from 0.20 to 0.45 per year (Table 1), and another set had fishing mortality ranging from 0.22 to 0.28 per year (Table 2).

The following two models were fitted to each of the data sets: (1) the delayed mixing model described by Hoenig et al. (1998b), in which mixing is assumed after half a year and data are tabulated by year and (2) the delayed PYT model, in which mixing occurs after half a year and data are tabulated by half year.

The parameters estimated for the delayed PYT model are F_1, F_2, F_3, F_4, F_{1a}^*, F_{2a}^*, F_{3a}^*, and M. The parameters in the Hoenig et al. (1998b) model are F_2, F_3, F_4, F_1^*, F_2^*, F_3^*, and M. Although Hoenig et al. (1998b) parameterized their model in terms of F_i^*, the simulations here were parameterized in terms of F_{ia}^* for $i = 2, 3$, so that estimates could be compared between models.

Under the Hoenig et al. model, for year 1 only F_1^* can be estimated and it cannot be broken down into the temporal components of mortality. There is no F_{4a}^* in either model since tagging occurred for just 3 years.

Computations were done using the statistical language R (R Development Core Team 2008), as described below. These computations could have also been done using the program SURVIV (White 1992).

A second group of simulations was conducted to evaluate how well a model selection criterion would perform in choosing the best-fitting model. Two models were compared: the delayed PYT model and a fully mixed model in which the tag recaptures were tabulated by parts of the year (referred to as the fully mixed PYT model). The formulation of the fully mixed PYT model is similar to that of the delayed PYT model, except that anywhere an F_j^* appears it is replaced by F_j. Tabulating recaptures by parts of the year does not increase the precision of the fully mixed

TABLE 1. Simulation results for estimating instantaneous rates of fishing mortality (F_j) and natural mortality (M). Data were generated with contrasting F_j values across years and with $p = q$, that is, with 50% of F_j occurring in part (a) of every year j. Mean $\hat{SE}$ refers to the mean of the 2,000 standard errors calculated by inverting the Hessian. The SE of the estimates is the true standard error, that is, the standard deviation of the 2,000 estimates of each parameter. The RMSE is root mean square error. "Hoenig et al." refers to the delayed mixing model of Hoenig et al. (1998b) with recaptures tabulated by year; "Delayed PYT" refers to the new model in which recaptures are tabulated by parts of the year. Values of F_1 and F_{1a}^* cannot be estimated from the Hoenig et al. model. The smaller quantity (in absolute value) for each performance measure for each parameter is in bold italics.

Model fitted	Parameter	True value	Mean estimate	Bias	Mean $\hat{SE}$	SE of estimates	RMSE
Delayed PYT	F_1	0.20	0.200	0.0002	0.023	0.023	0.023
	F_2	0.30	0.300	*0.0002*	*0.023*	*0.023*	*0.023*
	F_3	0.25	0.251	*0.0006*	*0.021*	*0.021*	*0.021*
	F_4	0.45	0.453	*0.0035*	*0.048*	*0.046*	*0.046*
	F_{1a}^*	0.05	0.050	−0.0001	0.007	0.007	0.007
	F_{2a}^*	0.60	0.600	*0.0004*	*0.030*	*0.029*	*0.029*
	F_{3a}^*	0.15	0.150	*−0.0003*	*0.013*	*0.013*	*0.013*
	M	0.20	0.200	*0.0002*	*0.028*	*0.027*	*0.027*
Hoenig et al.	F_1	0.20					
	F_2	0.30	0.300	−0.0004	0.029	0.029	0.029
	F_3	0.25	0.252	0.0017	0.030	0.030	0.030
	F_4	0.45	0.456	0.0061	0.057	0.055	0.055
	F_{1a}^*	0.05					
	$F_{2a}^* = pF_2^*$	0.60	0.602	0.0019	0.036	0.035	0.035
	$F_{3a}^* = pF_3^*$	0.15	0.150	−0.0004	0.021	0.021	0.021
	M	0.20	0.202	0.0016	0.033	0.032	0.032

TABLE 2. Simulation results for estimating instantaneous rates of fishing mortality (F_j) and natural mortality (M). Data were generated with similar F_j values across years and with $p = q$; see Table 1 for further details.

Model fitted	Parameter	True value	Mean estimate	Bias	Mean $\hat{SE}$	SE of estimates	RMSE
Delayed PYT	F_1	0.22	0.220	−0.0002	0.033	0.033	0.033
	F_2	0.28	0.281	*0.0010*	*0.030*	*0.030*	*0.030*
	F_3	0.25	0.251	*0.0010*	*0.030*	*0.031*	*0.031*
	F_4	0.23	0.233	*0.0030*	*0.039*	*0.040*	*0.040*
	F_{1a}^*	0.62	0.621	0.0012	0.031	0.031	0.032
	F_{2a}^*	0.60	0.600	*0.0000*	*0.031*	*0.030*	*0.030*
	F_{3a}^*	0.68	0.680	*0.0000*	*0.034*	*0.034*	*0.034*
	M	0.20	0.199	*−0.0010*	*0.045*	*0.044*	*0.044*
Hoenig et al.	F_1	0.22					
	F_2	0.28	0.282	0.0020	0.043	0.044	0.044
	F_3	0.25	0.246	−0.0040	0.040	0.040	0.040
	F_4	0.23	0.223	−0.0070	0.042	0.043	0.043
	$F_{1a}^* = pF_1^*$	0.62					
	$F_{2a}^* = pF_2^*$	0.60	0.596	−0.0040	0.037	0.035	0.035
	$F_{3a}^* = pF_3^*$	0.68	0.676	−0.0040	0.037	0.037	0.038
	M	0.20	0.182	−0.0180	0.050	0.050	0.054

PYT model relative to that of the fully mixed model with tag recaptures tabulated by year, but it does allow for comparison between models with immediate and delayed mixing of newly tagged animals. For the model with full mixing of newly tagged animals into the population immediately after tagging, the parameters estimated when the data are tabulated by part of the year are simply F_1, F_2, F_3, F_4, and M, the apportionment of the mortality rates to portion of the year being done simply on the basis of the known fractions of the mortality occurring in each division of time.

A commonly used model selection criterion is Akaike's information criterion (AIC; see Anderson 2008), which is defined as

$$\text{AIC} = -2\log_e(\Lambda) + 2k, \quad (20)$$

where Λ is the value of the maximum for the likelihood function for the model and k is the number of parameters in the model. Models that have lower AIC values are considered better-fitting models. Additionally, one can define ΔAIC, which is the difference in AIC values between an alternative model and the best-fitting model. A general rule of thumb is that a model with a ΔAIC value of less than 2 or 3 is a plausible alternative to the best-fitting model (Anderson 2008).

For the second group of simulations, the values of fishing and natural mortality were those in Table 1. The values of the abnormal rates of fishing mortality, F_{ja}^*, relative to the normal rates of fishing mortality, F_{ja} (note that $F_{ja} = 0.5\,F_j$), were varied

to see what effect the magnitude of this difference would have on model performance and on AIC for model selection (Tables 3–5; Figure 1). Parameter estimates and standard errors were calculated for an "AIC-selected model," that is, for each simulation the estimates from the model (fully mixed PYT or delayed PYT model) with the lower AIC value were used, and then these 10,000 model-selected estimates were averaged (Tables 3–5).

The function rmultinomial (Chasalow 2005) in the R package combinat was used to generate multinomial data sets with specified sample sizes and cell probabilities. The function nlm was used to minimize the negative log likelihood function and obtain the Hessian (R Development Core Team 2008). Standard errors were estimated by inverting the Hessian matrix using the R function solve. The true standard error was determined from the variability (standard deviation) of the 2,000 estimates for each parameter. The output from the Monte Carlo simulations includes estimates of the parameters for each simulated data set and the bias, standard error, and bias of the estimated standard error. Additionally the root mean square error (RMSE) for each parameter was calculated as

$$\text{RMSE} = \sqrt{\text{bias}^2 + \text{variance}} = \sqrt{\frac{\sum_{i=1}^{T}(\hat{\theta}_i - \theta)^2}{T}}, \quad (21)$$

where T is the number of simulated data sets (2,000) and $\hat{\theta}_i$ is the ith estimate of the parameter whose true value is θ.

TABLE 3. Simulation results for estimating the instantaneous rate of fishing mortality (F_j) in years 1 through 4 and natural mortality (M). Data were generated with contrasting F_j values across years and with $p = q$, that is, with 50% of F_j occurring in part (a) of year j. Here $F_j^* = 1.0F_j$ (i.e., full mixing). The SE of the estimates refers to the true standard error, that is, the standard deviation of the 10,000 estimates of each parameter. RMSE is root mean squared error. "AIC selected" refers to the mean estimate calculated using parameter estimates for the model with the smaller Akaike information criterion (AIC) value for each simulated data set. For each parameter, the smallest values for the SE of the estimates and RMSE are in bold italics.

Model fitted	Parameter	True value	Mean estimate	Bias	SE of estimates	RMSE
Delayed PYT	F_1	0.20	0.200	0.0001	0.024	0.024
	F_2	0.30	0.300	−0.0001	0.022	0.022
	F_3	0.25	0.250	0.0000	0.020	0.020
	F_4	0.45	0.452	0.0016	0.047	0.047
	M	0.20	0.200	−0.0004	0.028	0.028
Fully mixed PYT	F_1	0.20	0.200	−0.0001	*0.016*	*0.016*
	F_2	0.30	0.300	0.0000	*0.017*	*0.017*
	F_3	0.25	0.250	0.0000	*0.015*	*0.015*
	F_4	0.45	0.452	0.0016	*0.043*	*0.043*
	M	0.20	0.200	−0.0001	*0.025*	*0.025*
AIC selected	F_1	0.20	0.200	0.0000	0.019	0.019
	F_2	0.30	0.300	0.0000	0.018	0.018
	F_3	0.25	0.250	0.0001	0.017	0.017
	F_4	0.45	0.452	0.0017	0.044	0.044
	M	0.20	0.200	−0.0002	0.026	0.026

Results

In the first set of simulations, in which the values of F_j vary greatly from year to year (range, 0.20–0.45), both the delayed PYT model and the delayed mixing model of Hoenig et al. (1998b) yield essentially unbiased estimates of the F_j ($j = 2$, ..., 4) and F_{ja}^* ($j = 2, 3$) (Table 1). For all the F_j and F_{ja}, the delayed PYT model produces estimates with smaller standard deviations and smaller RMSEs. Only the delayed PYT model yields an estimate of F_1 (mean $\hat{F}_1 = 0.200$ [$F_1 = 0.20$], with a RMSE of 0.023). Both models yield essentially unbiased estimates of the standard errors.

When the F_j are close in value to each other, both the delayed PYT model and the delayed mixing model of Hoenig et al. (1998b) yield essentially unbiased estimates of the F_j values

TABLE 4. Simulation results for estimating the instantaneous rate of fishing mortality (F_j) in years 1 through 4 and the instantaneous rate of natural mortality (M). Data were generated with contrasting F_j values across years, $p = q$, and $F_j^* = .75F_j$. See Table 3 for further details.

Model fitted	Parameter	True value	Mean estimate	Bias	SE of estimates	RMSE
Delayed PYT	F_1	0.20	0.200	0.0001	0.024	*0.024*
	F_2	0.30	0.300	−0.0001	0.022	*0.022*
	F_3	0.25	0.250	0.0000	0.020	*0.020*
	F_4	0.45	0.452	0.0016	0.047	*0.047*
	M	0.20	0.200	−0.0004	0.028	*0.028*
Fully mixed PYT	F_1	0.20	0.170	−0.0305	*0.014*	0.034
	F_2	0.30	0.265	−0.0354	*0.015*	0.038
	F_3	0.25	0.220	−0.0297	*0.014*	0.033
	F_4	0.45	0.400	−0.0505	*0.039*	0.064
	M	0.20	0.163	−0.0373	*0.026*	0.045
AIC selected	F_1	0.20	0.199	−0.0012	0.024	*0.024*
	F_2	0.30	0.299	−0.0012	0.023	0.023
	F_3	0.25	0.249	−0.0008	0.020	*0.020*
	F_4	0.45	0.450	0.0002	0.048	0.048
	M	0.20	0.199	−0.0014	0.029	0.029

TABLE 5. Simulation results for estimating the instantaneous rate of fishing mortality (F_j) in years 1 through 4 and the instantaneous rate of natural mortality (M). Data were generated with contrasting F_j values across years, $p = q$, and $F_j^* = 1.1F_j$. See Table 3 for further details.

Model fitted	Parameter	True value	Mean estimate	Bias	SE of estimates	RMSE
Delayed PYT	F_1	0.20	0.200	0.0002	0.024	0.024
	F_2	0.30	0.300	−0.0001	0.022	0.022
	F_3	0.25	0.250	0.0001	0.020	0.020
	F_4	0.45	0.452	0.0021	0.047	0.047
	M	0.20	0.200	−0.0003	0.028	0.028
Fully mixed PYT	F_1	0.20	0.212	0.0125	***0.017***	***0.021***
	F_2	0.30	0.315	0.0146	***0.017***	0.023
	F_3	0.25	0.262	0.0123	***0.016***	0.020
	F_4	0.45	0.473	0.0234	***0.045***	0.051
	M	0.20	0.214	0.0137	***0.025***	0.029
AIC selected	F_1	0.20	0.200	−0.0002	0.023	0.023
	F_2	0.30	0.300	−0.0002	0.021	***0.021***
	F_3	0.25	0.250	0.0004	0.019	***0.019***
	F_4	0.45	0.452	0.0018	***0.045***	***0.045***
	M	0.20	0.200	−0.0002	0.027	***0.027***

($j = 2, \ldots, 4$) and F_{ja}^* ($j = 2, 3$) (Table 2). The delayed PYT model produces an unbiased estimate of M, while the delayed mixing model of Hoenig et al. (1998b) is negatively biased (percent bias for $\hat{M} = -9\%$). Only the delayed PYT model produces
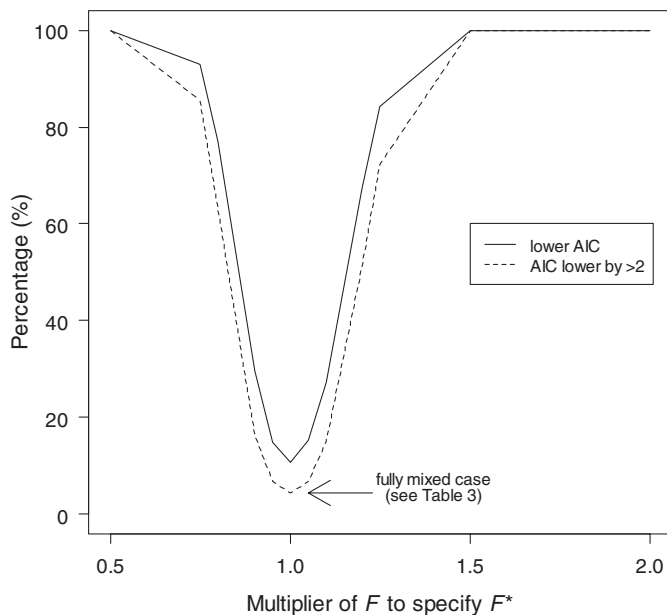

FIGURE 1. Percentage of the 10,000 simulations for which the delayed PYT model is selected as a result of having the lower AIC value (solid line) or when the AIC value of the delayed PYT model is smaller than that of the fully mixed model (dashed line) by at least 2. Both models used data with tag recaptures tabulated by parts of the year. The values of the fishing mortality in year j, F_j, can be found in Table 3. The value of F_j^* was specified as a multiple of F_j to see how this affected model performance.

an estimate of F_1. However, it is worth noting that the delayed PYT model produces a mean estimate of $\hat{F}_1$ of 0.220 ($F_1 = 0.22$) and thus is unbiased, with a RMSE of 0.033. Both models yield essentially unbiased estimates of the standard errors.

When the data are generated with the values of the F_j being closer together than in the previous simulation (Table 2), the performance of the delayed PYT model improves relative to that of the Hoenig et al. model by having lower RMSEs for the estimates of M, F_j ($j = 1, 2, 3$), and F_{ja}^* ($j = 1, 2, 3$).

For the second set of simulations, in which the tag recaptures were tabulated by parts of the year, the percentage of the 10,000 simulations in which the delayed PYT model had a lower AIC value than the fully mixed PYT model varied depending on the multiplier between F_j^* and F_j (Figure 1). When the data come from a case of full mixing (Table 3) and one uses AIC as the model selection criterion, the wrong model is only selected 10.66% of the time. Hence, uncertainty about model selection is low. In fact, bias of the estimates under the AIC-selected model is essentially zero (as with the fully mixed and delayed PYT models). The RMSE is slightly higher when model selection is used than when the fully mixed PYT model is used. This is because for 10.66% of the 10,000 simulations the less precise model (the delayed PYT model) was selected.

When the data come from a case of minor delayed mixing (i.e, $F_j^* = 1.1F_j$), the delayed PYT model is selected 27.21% of the time (Figure 1). Thus, the incorrect model (the fully mixed PYT model) is selected 72.79% of the time, yielding parameter estimates with larger biases than the delayed PYT model (Table 5). Interestingly, for four of the five parameters (F_2, F_3, F_4, and M), the AIC-selected model has the lowest RMSE. This is because 72.79% of the time a more precise model (the fully mixed PYT model) is selected and 27.21%

of the time the essentially unbiased model (the delayed PYT model) is selected, resulting in estimates that have lower RMSEs than those using either model alone.

If the data come from a stronger case of delayed mixing (i.e., $F_j^* = 0.75F_j$), the delayed PYT model is selected 93.1% of the time (Figure 1). Now the AIC-selected model is tied for the lowest RMSE for just two of the parameters (F_1 and F_3) and the delayed PYT model has the lowest RMSE for all parameters (including the two ties) (Table 4).

DISCUSSION

To determine whether it is appropriate to apply the delayed PYT model, it must first be determined whether delayed mixing is a problem. If a model for delayed mixing is applied when the newly tagged cohorts fully mix with those previously tagged, the delayed PYT model will probably still outperform the delayed mixing model of Hoenig et al. (1998b) by yielding estimates closer to those found when fitting the fully mixed model and the delayed PYT model will have smaller standard errors for each parameter (Tables 1, 2). However, the delayed PYT model will probably be less precise than the fully mixed model since the additional parameters reduce precision.

When delayed mixing does occur for part of the year, the delayed PYT model provides more precise estimates of the rates of fishing and natural mortality than the delayed mixing model of Hoenig et al. (1998b). Furthermore, if the fully mixed model is applied when delayed mixing occurs, the estimates may be biased for all parameters. In fact, whether there is full mixing or delayed mixing for part of a year, the delayed PYT model will generally outperform the delayed mixing model of Hoenig et al. (1998b) by providing estimates with smaller estimated standard errors as well as an estimate of the rate of fishing mortality in the first year (which the Hoenig et al. 1998b model does not provide).

To use the delayed PYT model, it is necessary to determine the value of p, the proportion of normal fishing mortality occurring in part (a) of the year. The accurate determination of p is not critical for the estimation of any value except F_1 (Hoenig et al. 1998b). It is critical for F_1 because the value of p determines the multiplier used to relate the fishing mortality estimated for part (b) of the year to the fishing mortality over the whole year.

Additionally, the degree to which the rate of fishing mortality varies across the years will affect the precision of the estimates of the normal and abnormal rates of fishing mortality, as shown by the simulations (compare Tables 1, 2). The greater the variability in the rate of fishing mortality, the more precise the estimates will be.

Tabulating the recaptures on a finer scale than the periodicity of tagging should not be costly. Pollock and Raveling (1982) showed that when conducting a tagging study it is important to determine the year of tag recapture in order to avoid biased estimates of survival. Thus, assuming that the advice of Pollock and Raveling is followed, determining whether the tag return is from part (a) or part (b) of the year should add little or no additional cost to the study. Even if fishers are contacted to determine the time of recapture, there may be some reporting errors and these will cause a bias (Pollock and Raveling 1982). Furthermore, by recording the date of recapture rather than just whether it is in part (a) or (b) of the year, one need not determine how to split the year into parts (a) and (b) before the start of the study.

For these simulations the product of $\varphi\lambda$ was assumed to be 1. In practice, one would either use the data to estimate $\varphi\lambda$ or use additional information to obtain an external estimate. With a large data set (i.e., high numbers of recaptures and many years of tagging), $\varphi\lambda$ can sometimes be estimated from the observed information, although the estimates of $\varphi\lambda$ tend to be highly variable. External estimates of $\varphi\lambda$ can also be obtained using methods such as high-reward tagging or planted-tag studies (Hoenig et al. 1998a; Hearn et al. 1999, 2003; Pollock et al. 2001, 2002a, 2002b), and such information should be incorporated into the likelihood if possible. Thus, the delayed PYT model can be used in a much more general setting in which there is an outside estimate of $\varphi\lambda$ or it is estimated from the observed information.

This is one of the first studies to show that tabulating tag recoveries on a finer scale than the periodicity of tagging can be advantageous. Another study, by Waterhouse (2010), described a generalization of Brownie model zero (Brownie et al. 1978, 1985) to address tag fouling occurring in the first part of the first year in which a cohort is at liberty, with the result that newly tagged animals are more visible and more likely to be reported than previously tagged ones. There are also tagging models that utilize the exact times of recapture of animals (see Leigh et al. 2006). The approach of Leigh et al. (2006) is innovative and promising, though more complicated than the models considered here. The relative performance of the two approaches remains to be seen (e.g., with respect to their robustness to failures of the assumptions). It is yet to be seen what further benefits can accrue from tabulating recapture data on a finer scale than the tagging periodicity.

REFERENCES

Anderson, D. R. 2008. Model based inference in the life sciences: a primer on evidence. Springer-Verlag, New York.

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1978. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Service Resource Publication 131.

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1985. Statistical inference from band recovery data: a handbook, 2nd edition. U.S. Fish and Wildlife Service Resource Publication 156.

Chasalow, S. 2005. Combinat: combinatorics utilities. R package version 0.0–6. Available: cran.r-project.org/web/packages/combinat/index.html. (July 2010).

Hearn, W. S., J. M. Hoenig, K. H. Pollock, and D. Hepworth. 2003. Tag reporting rate estimation. 3. Use of planted tags in one component of a multiple-component fishery. North American Journal of Fisheries Management 23:66–77.

Hearn, W. S., T. Polacheck, K. H. Pollock, and W. Whitelaw. 1999. Estimation of tag reporting rates in age-structured multicomponent fisheries where one component has observers. Canadian Journal of Fisheries and Aquatic Sciences 56:1255–1265.

Hoenig, J. M., N. J. Barrowman, W. S. Hearn, and K. H. Pollock. 1998a. Multiyear tagging studies incorporating fishing effort data. Canadian Journal of Fisheries and Aquatic Sciences 55:1466–1476.

Hoenig, J. M., N. J. Barrowman, K. H. Pollock, E. N. Brooks, W. S. Hearn, and T. Polacheck. 1998b. Models for tagging data that allow for incomplete mixing of newly tagged animals. Canadian Journal of Fisheries and Aquatic Sciences 55:1477–1483.

Leigh, G. M., W. S. Hearn, and K. H. Pollock. 2006. Time-dependent instantaneous mortality rates from multiple tagging experiments with exact times of release and recovery. Environmental and Ecological Statistics 13:89–108.

Pollock, K. H., W. S. Hearn, and T. Polacheck. 2002a. A general model for tagging on multiple component fisheries: an integration of age-dependent reporting rates and mortality estimation. Environmental and Ecological Statistics 9:57–69.

Pollock, K. H., J. M. Hoenig, W. S. Hearn, and B. Calingaert. 2001. Tag reporting rate estimation, 1. An evaluation of the high-reward tagging method. North American Journal of Fisheries Management 21:521–532.

Pollock, K. H., J. M. Hoenig, W. S. Hearn, and B. Calingaert. 2002b. Tag reporting rate estimation, 2. Use of high-reward tagging and observers in multiple-component fisheries. North American Journal of Fisheries Management 22:727–736.

Pollock, K. H., J. M. Hoenig, and C. M. Jones. 1991. Estimation of fishing and natural mortality rates when a tagging study is combined with a creel survey or port sampling. Pages 423–434 *in* D. Guthrie, J. M. Hoenig, M. Holliday, C. M. Jones, M. J. Mills, S. A. Moberly, K. H. Pollock, and D. R. Talhelm, editors. Creel and angler surveys. American Fisheries Society, Symposium 12, Bethesda, Maryland.

Pollock, K. H., and D. G. Raveling. 1982. Assumptions of modern band recovery models, with emphasis on heterogeneous survival rates. Journal of Wildlife Management 46:88–98.

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available: www.R-project.org. (July 2010).

Robson, D. S., and W. D. Youngs. 1971. Statistical analysis of reported tag recaptures in the harvest from an exploited population. Cornell University, Mimeographic Series Paper BU-369-M, Ithaca, New York.

Seber, G. A. F. 1970. Estimating time-specific survival and reporting rates for adult birds from band returns. Biometrika 57:313–318.

Waterhouse, L. 2010. Partial-year tagging models: accounting for tag fouling and delayed mixing. Master's thesis. Virginia Institute of Marine Science, College of William and Mary, Williamsburg.

White, G. C. 1992. PC SURVIV user's manual. Colorado State University, Fort Collins. Available: www.mbr-pwrc.usgs.gov/software/surviv.html. (May 2011).

Youngs, W. D., and D. S. Robson. 1975. Estimating survival rates from tag returns: model test and sample size determination. Journal of the Fisheries Research Board of Canada 32:2365–2371.