Abstract—Tagging experiments are a useful tool in fisheries for estimating mortality rates and abundance of fish. Unfortunately, nonreporting of recovered tags is a common problem in commercial fisheries which, if unaccounted for, can render these estimates meaningless. Observers are often employed to monitor a portion of the catches as a means of estimating reporting rates. In our study, observer data were incorporated into an integrated model for multiyear tagging and catch data to provide joint estimates of mortality rates (natural and fishing), abundance, and reporting rates. Simulations were used to explore model performance under a range of scenarios (e.g., different parameter values, parameter constraints, and numbers of release and recapture years). Overall, results indicated that all parameters can be estimated with reasonable accuracy, but that fishing mortality, reporting rates, and abundance can be estimated with much higher precision than natural mortality. An example of how the model can be applied to provide guidance on experimental design for a large-scale tagging study is presented. Such guidance can contribute to the successful and cost-effective management of tagging programs for commercial fisheries.

Manuscript submitted 10 December 2006 to the Scientific Editor's Office.

Manuscript approved for publication 24 May 2007 by the Scientific Editor.

Fish. Bull. 105:493–508 (2007).

Incorporating fishery observer data into an integrated catch-at-age and multiyear tagging model for estimating mortality rates and abundance

J. Paige Eveson (contact author)[1]

Tom Polacheck[1]

Geoff M. Laslett[2]

Email for J. Paige Eveson: paige.eveson@csiro.au

[1] Commonwealth Scientific and Industrial Research Organization (CSIRO)
Marine and Atmospheric Research
Castray Esplanade
Hobart, Tasmania 7000, Australia
Postal address: GPO Box 1538
Hobart, Tasmania 7001, Australia

[2] CSIRO Mathematical and Information Sciences
Private Bag 33
Clayton South, Victoria 3169, Australia

Tagging experiments are becoming increasingly important in large pelagic fisheries as a means of providing estimates of stock abundance and fishing mortality rates that are independent of catch-rate data (Polacheck and Hearn, 2003). In Polacheck et al. (2006), we developed a maximum likelihood model that combines two traditional, but fundamentally different, approaches for analyzing tagging data with a single, terminal recapture (note that we refer to this as "tag-recapture" data, but the term "tag-recovery" data is often used in the literature). The first approach, generally referred to as a Brownie model (Brownie et al., 1985), uses tag-recapture data from multiple years of tagging to provide annual estimates of mortality rates by comparing return rates over time from the releases in consecutive years. Only the numbers of tag releases and returns by year are required, not the number of animals examined for tags. The standard Brownie model is formulated in terms of rates of survival and tag recovery, but can also be expressed in terms of instantaneous rates of natural mortality and exploitation (Pollock et al., 1991; Hoenig et al., 1998a). This latter formulation is particularly useful in fishery applications (e.g., Hampton, 2000; Frusher and Hoenig, 2001; Polacheck et al., 2006). The second approach, known as a Petersen model (e.g., Seber, 1982), uses data from a single release event to provide an estimate of population size at the time of tagging based on the ratio of the number of tags returned from a sample of the population to the total number of tags in the population. In fishery applications, commercial catch data usually constitute the sample from which tags are returned.

The model developed by Polacheck et al. (2006) integrates catch data with data from a multiyear tagging experiment and, in essence, incorporates a Petersen estimator into a Brownie model; we will refer to it as the Brownie-Petersen (BP) model. The BP model involves a likelihood for the tag-recapture data and a likelihood for the catch data, which can be jointly maximized to provide estimates of natural mortality rates, fishing mortality rates, and abundance. The addition of catch data to the traditional Brownie model not only allows for the population size at the time of first tagging to be estimated but also improves the precision of the mortality-rate estimates (Polacheck et al., 2006). For readers familiar with multiple-recapture tagging models, the BP model has similarities with an age-structured Jolly-Seber (JS) model (Jolly, 1965; Seber, 1965; Pollock, 1981). Both the BP and JS models have a likelihood component for the recapture data, from which survival rates can be estimated. For the BP model, this component equates to a Brownie model, and for the JS model, it equates to a Cormack-Jolly-Seber (CJS) model (Cormack, 1964). Additionally, they both have a likelihood component involving the total number of animals sampled, from which abundance can be estimated (i.e., the Petersen component). However, for the JS model there is an unresolved problem with the Petersen component regarding how the information from unmarked animals should be integrated into the likelihood, and a variety of approaches have been developed to address this problem (see section 4.3 of Schwarz and Seber, 1999, and references therein). The Petersen component of the BP model (i.e., the catch component) is more general and integration into the likelihood is more straightforward.

A recognized problem with applying tagging experiments in fishery situations is that of nonreporting. When recapture information comes from commercial fisheries, it is unlikely that all recaptured tags will be reported, or that the rate of reporting will be known. Although Brownie models can provide estimates of total mortality rates when reporting rates are unknown (Brownie et al., 1985), the separation of natural mortality from fishing mortality generally requires that reporting rates are either known or estimable (Pollock et al., 1991; Hoenig et al., 1998a). Petersen models also require reporting rates to determine abundance estimates.

A number of methods exist for estimating reporting rates (see Pollock et al., 2001). For some methods, such as planted (also called "seeded") tag experiments, the data are independent of the tag-recapture and related catch data from the primary tagging study. In developing the BP model, we assumed that independent reporting rate data were available; therefore a likelihood could be constructed for these data and simply multiplied to the likelihoods for the tag-recapture and catch data. Another common method for estimating reporting rates is to have observers monitor a portion of the catches. Under the assumption that 100% of tags will be returned (i.e., reported) from the observed catches, the reporting rate for the unobserved catches can be estimated by using the relative return rate of tags from the unobserved versus observed catches (Hearn et al., 1999). In the case of longline fisheries, where fish are not brought into port for processing, the use of observers to estimate reporting rates is probably the most viable approach. Unlike data from a planted tag experiment, observer data cannot be considered independent of the tagging or related catch data and therefore incorporating the estimation of reporting rates into the BP model is more complicated.

Pollock et al. (2002) showed how a standard Brownie model can be modified to include the estimation of reporting rates when one component of a multicomponent fishery has 100% reporting rates (e.g., one component has observers). This modification required that supplementary catch data be brought into the model to assist in the estimation of reporting rates. Pollock et al. (2002) acknowledged that uncertainty in the catch data was not accounted for in their model, and also that it would be resourceful to take advantage of the extra information provided by the catch data to estimate population size. As a topic of future research, they advocated the development of an integrated analysis that estimates all parameters (fishing mortality, natural mortality, population size, and reporting rates) within a single likelihood.

In this article, the BP model is extended to include the estimation of reporting rates by using observer data. We will refer to this extended model as the BPO model, short for the Brownie-Petersen model with observers. The BPO model fulfills the goal of Pollock et al. (2002) for an integrated likelihood that can provide joint estimates of mortality rates, abundance, and reporting rates, and it also directly incorporates uncertainty in the catch data. Results from applying the model to simulated data are presented which demonstrate the accuracy and precision that can be achieved in the parameter estimates under various scenarios (e.g., different parameter values, different numbers of release and recapture years, different parameter constraints). Tag-recapture data and catch data from most field studies will exhibit more variability than the model predicts (i.e., will be overdispersed). Thus, extra variability was included in the simulated data sets to investigate the consequences of applying the model to overdispersed data. Finally, a practical illustration is given of how the model can be used to evaluate the trade-off between releasing more tags and increasing the level of observer coverage in terms of the accuracy and precision of the parameter estimates. Polacheck and Hearn (2003) investigated this issue using a much simpler model (e.g., only one release event; only fishing mortality rates estimated) and making many simplifying assumptions (e.g., natural mortality known; no uncertainty in the catch data; no overdispersion in the data). The BPO model provides a much more comprehensive framework for evaluating such trade-offs and can thereby make an important contribution to the successful and cost-effective management of tagging programs for commercial fisheries.

Materials and methods

Model description

Consider a multiyear tagging study in which a single cohort of fish is tagged in A consecutive years starting at age 1 (i.e., at age 1 in year 1, age 2 in year 2, up to age A in year A). Fish from this cohort are subsequently caught in a fishery over years, or ages, 1 to I ($I \geq A$), and a percentage of the tags that are recaptured each year are reported. Observers monitor a portion of the catches, and 100% of recaptured tags are reported from the observed component of the fishery. Furthermore, all fish caught in the observed component of the fishery are sampled for length or age, but no fish from the unobserved component are sampled. The catch monitored by observers is assumed to be representative of the total catch (i.e., catches from the observed and unobserved components have the same expected age distribution). If the expected catch-at-age distribution differed between the two components, then separate age information would need to be available for each component and the catch likelihood presented below would need to be modified.

The basic assumptions common to all multiyear tagging models, as summarized in Pollock et al. (1991), are also required for the BPO model. The most important of these are: 1) tagged and untagged fish are thoroughly mixed throughout the population of interest, 2) the fate of each fish is independent of the fate of other fish, 3) all fish of a given age class have the same survival and capture probabilities, and 4) there is no tag shedding or tag-induced mortality. If tag shedding or tag-induced mortality, or both, exist at non-trivial levels (i.e., assumption 4 is not met), then additional parameters and potentially additional data need to be introduced to account for them. Failing to do so will lead to biased parameter estimates and overly optimistic estimates of their precision. If any of assumptions 1 to 3 is violated, then the variance of the tag return counts will be underestimated by the model. Similarly, if assumption 2 or 3 is violated, the variance of the catch numbers will be underestimated. Extra variability, or overdispersion, in the tag return and catch data is discussed in the next section.

Assumption 1 implies that newly tagged fish are mixed throughout the population immediately after tagging. This mixing can be difficult to achieve in practice, especially when the population has a widespread geographical distribution or tagging occurs in a limited area of its distribution. Hoenig et al. (1998b) showed how delayed mixing of newly tagged fish can be incorporated into a Brownie model by allowing these fish to have a different fishing mortality rate in the year of tagging than that of previously tagged fish. In our application of the BP model to southern bluefin tuna (SBT, *Thunnus maccoyii*) data in Polacheck et al. (2006), we allowed for initial nonmixing with this approach. Only the tag-recapture component of the model needed to be modified. It would be straightforward to modify the tag-recapture component of the BPO model in an analogous manner in situations where modification was considered necessary.

Before proceeding, we introduce the notation that will be used throughout this study. The data required by the model are

N_a = the number of tag releases of age a fish from a particular cohort;
$R^o_{a,i}$ = the number of tag returns from fish that were tagged at age a and recaptured at age i in the observed (o) component of the fishery;
$R^u_{a,i}$ = the number of tag returns from fish that were tagged at age a and recaptured at age i in the unobserved (u) component of the fishery; and
C^o_i = the estimated number of age i fish from the cohort of interest caught in the observed (o) component of the fishery.

The model parameters assumed to be known are

δ_i = the proportion of fish from the cohort of interest caught in the observed component of the fishery in year i;
η^2_i = the variance of the aging error for C^o_i.

The model parameters to be estimated from the data are

M_i = the instantaneous natural mortality rate for age i fish;
F_i = the instantaneous fishing mortality rate for age i fish;
P_1 = the population size of the tagged cohort at the age of first tagging (assumed to be age 1 for convenience); and
λ_i = the tag reporting rate for fish captured at age i in the unobserved component of the fishery.

In addition, the annual survival rate (S_i) and exploitation rate (u_i), respectively, of an age i fish, are defined to be

$$S_i = \exp\left(-\left(F_i + M_i\right)\right);$$

$$u_i = \frac{F_i}{F_i + M_i}\left(1 - S_i\right).$$

Note that because only a single cohort of fish is being considered, age and year can be used interchangeably in the above definitions. If more cohorts were added to the model, it would then be important to distinguish whether the parameters vary by year, by age, or both. For example, λ may vary by year, M by age, and F by both. Because the age distribution of the catch is assumed to be the same for the observed and unobserved components, δ would vary with year, not age, when there is more than one cohort (i.e., the probability of a fish being caught in the observed component of the fishery in year i would be the same for all ages within the year). If the age distribution of the catch was allowed to differ between the observed and unobserved components, then δ would need to vary with both year and age, but it would not be estimable unless information was available about the age distribution of the unobserved catches.

First consider the tag-recapture component of the model. The probability of a fish, tagged at age a, being caught in the observed component of the fishery at age i, and having its tag returned, is

$$p^o_{a,i} = \begin{cases} \delta_i u_i & i = a \\ \delta_i S_a \cdots S_{i-1} u_i & i > a \end{cases}. \quad (1)$$

Similarly, the probability of a fish, tagged at age a, being caught in the unobserved component of the fishery at age i, and having its tag returned, is

$$p^u_{a,i} = \begin{cases} (1-\delta_i)u_i\lambda_i & i=a \\ (1-\delta_i)S_a\cdots S_{i-1}u_i\lambda_i & i>a \end{cases}. \quad (2)$$

Thus, the probability of a fish, tagged at age a, not being recaptured by age I from either component is $p'_a = 1 - p^o_{a,i} - p^u_{a,i}$. Here, and below, a dot in the subscript denotes summation over the index it replaces.

For tags released at age a, the numbers of returns at ages a to I from the observed component ($R^o_{a,i}$, $i=a, \ldots, I$) and unobserved component ($R^u_{a,i}$, $i=a, \ldots, I$), plus the number not returned by age I from either component ($R'_a = N_a - R^o_{a,i} - R^u_{a,i}$), are multinomial with probabilities given by Equations 1, 2 and p'_a, respectively. Thus, the likelihood equation for the returns from tags released at all ages is the product of multinomials, given by

$$L_R = \gamma \prod_{a=1}^{A} \left\{ (p'_a)^{R'_a} \prod_{i \geq a}^{I} (p^o_{a,i})^{R^o_{a,i}} (p^u_{a,i})^{R^u_{a,i}} \right\}, \quad (3)$$

where

$$\gamma = \prod_{a=1}^{A} \frac{N_a!}{R'_a! \prod_{i \geq a} \left(R^o_{a,i}!\, R^u_{a,i}! \right)}.$$

Note that γ is a constant that can be left out when maximizing the liklehood.

Next, consider the catch component of the model. Recall that no age information is obtained for the unobserved catches; therefore only catch-at-age data from the observed component are available for inclusion in the model. The probability of an age-1 fish from the cohort of interest subsequently being caught at age i in the observed component of the fishery is

$$\pi^o_i = \begin{cases} \delta_i u_i & i=1 \\ \delta_i S_1 \cdots S_{i-1} u_i & i>1 \end{cases}. \quad (4)$$

If the numbers of fish from the cohort of interest that are caught at ages 1 to I in the observed component of the fishery (C^o_i, i=1, . . . , I) are known accurately, then these numbers, along with the number of fish from the cohort not caught by age I, are multinomial and have probabilities given by Equation 4. Usually, however, the numbers of fish caught at each age are not known precisely because the ages are estimated either from lengths or from annuli in hard parts (the estimates will be more accurate in the latter case, but will still contain uncertainty). We assume the aging error of the age i catch has a Gaussian distribution with mean 0 (i.e., no bias) and a variance η^2_i.

Rather than modeling the catch data with both multinomial process error and Gaussian aging error, which would require a fairly complex approach, we approximated the distribution of the catch of age i fish in the observed component, C^o_i, as Gaussian with overall variance $\sigma^2_i = \eta^2_i + \tau^2_i$, where $\tau^2_i = P_1 \pi^o_i (1-\pi^o_i)$, is the multinomial variance component. The aging error, unless negligible, will tend to dominate the process error when the cohort size is reasonably large (≥100,000 individuals), as would be expected in most commercial fishery situations. For example, if the coefficient of variation (CV) of the aging error is 0.10, the cohort size is 100,000 and the probability of catching an age i fish (in either the observed or unobserved component of the fishery) is 0.10, then the ratio of the aging error variance to the process error variance is ~10 when the proportion of the catch in the observer component is 0.10, and it is ~50 when the proportion of the catch in the observer component is 0.50.

Thus, assuming that the C^o_i's are independent between ages, the likelihood for the observer catch data is

$$L_C = \prod_{i=1}^{I} \frac{1}{\sqrt{2\pi\sigma^2_i}} \exp\left(-\frac{1}{2\sigma^2_i} \left(C^o_i - E(C^o_i) \right)^2 \right), \quad (5)$$

where $E(C^o_i) = P_1 \pi^o_{\text{i}}$.

When only a single cohort of fish is being modeled, the assumption that the catch data are independent between ages (i.e., years) should be reasonable in most situations. First, the correlation in the multinomial errors will be close to zero when the size of the cohort is much larger than the size of the catch (as would be expected in most fisheries). Second, the aging errors should be uncorrelated between years provided sampling and aging data are collected each year. However, in some situations, particularly where age is being estimated from a growth curve, covariance in the estimates between years may exist and should be accounted for. Furthermore, if more than one cohort is being modeled, then catch data from multiple ages within the same year will enter the model, and aging errors within a year will be correlated across ages. The level of correlation, and thus the degree to which the independence assumption is violated, will depend on the specifics of the situation, such as how many age classes are present in the year's catch. When the correlation is strong, a more sophisticated approach for modeling the catch data may be required.

The overall likelihood for the combined recapture and catch data can be obtained by multiplying likelihoods (Eqs. 3 and 5) together:

$$L = L_R \times L_C. \quad (6)$$

In a tagging experiment with A consecutive release years, estimates can be obtained, at most, for A–1 natural mortality-rate parameters (regardless of the number of recapture years) because information for estimating M_i comes from the differential between the expected returns at age i+1 of fish released at age i and those released at age i+1. One option is to assume that $M_i = M_{A-1}$ for $i \geq A$, but other options, such as assuming M is constant or linear with age, are also possible. Furthermore, there is not enough information in the current formulation to estimate the proportion of fish caught each year in the observed component of the fishery (i.e., the δ_i's). To estimate this proportion would require knowing the total observer catch in each year, as well as the total overall catch in each year. Rather than bringing these data into the model, we assumed that the total catches are known well enough that the δ_i's can be treated as known without error. Lastly, the aging error variance parameters for the observer catches (i.e., the η_i^2's) cannot be estimated reliably and therefore they are assumed to be known without error. In Polacheck et al. (2006), we gave a detailed explanation of why the catch variance cannot be estimated reliably in the BP model, and the same argument applies here. We found, however, that the model results were fairly insensitive to the value used for the catch variance so long as it was in the right ballpark (e.g., within ~40% of the true value). The parameters that can be estimated by maximizing Equation 6 are F_i and λ_i for i = 1 to I, M_i for i = 1 to A–1, and P_1.

As is true when combining any sources of information, it is important to check that the tag-recapture data and the catch data are consistent. This can be done by maximizing the tag-recapture likelihood (Eq. 3) alone and comparing the mortality-rate estimates with those obtained from the joint likelihood (Eq. 6) (note that the catch likelihood alone is insufficient to yield parameter estimates). If the estimates are significantly different, this result would indicate that the tag-recapture and catch data are inconsistent and should not be combined; doing so would yield average values with little biological meaning. Instead, the source of the inconsistency should be investigated (i.e., does it stem from problems with the data or with the applicability of the assumptions in the model?).

Overdispersion in the recapture and catch data

In the model a multinomial distribution is assumed for the tag-recapture data. If one (or more) of model assumptions 1 to 3 is violated, then the observed return counts are expected to be more variable than predicted by a multinomial distribution; i.e., to be overdispersed in relation to multinomial data. Polacheck et al. (2006) provided a thorough discussion of possible sources of overdispersion and ways in which it can be accounted for. When overdispersion exists in the return counts, the parameter estimates obtained by using a multinomial likelihood should still be unbiased, but their standard errors, as estimated from traditional likelihood methods (i.e., from the inverse Hessian matrix), would be too small. A number of possible methods for obtaining more realistic standard errors are discussed in Polacheck et al. (2006) and Pollock et al. (2001), one of which is to use bootstrap procedures.

If overdispersion exists in the recapture data as a result of model assumptions 2 or 3 being violated, then it will also exist in the catch data. That is, the component of the variance in the catch-at-age numbers due to process error will be underestimated by a multinomial distribution. As asserted previously, aging error will generally dominate the multinomial process error in the catch data. This will often still be true when the process error is overdispersed. For example, assume that the process error variance is φ times that of multinomial variance; i.e., $\tau_i^2 = \varphi P_1 \pi_i^o (1-\pi_i^o)$. Then, in the example that was given above for multinomial process error, if φ=3, the ratio of the aging error variance to the process error variance would still be 3.3 (=10/3) and 17 (=50/3) when the proportion of catch in the observer component is 0.10 and 0.50, respectively. In situations where the aging error dominates, not accounting for overdispersion in the catch data should have little effect on the standard error estimates of the parameter estimates.

The degree to which the likelihood-based estimates of the standard errors are underestimated by not accounting for overdispersion in the tag-recapture data and catch data was investigated through simulations, as described below.

Simulation methods

Model performance To evaluate how the model performs in terms of the accuracy and precision of the parameter estimates, a series of Monte-Carlo simulations were conducted. The first scenario considered, which we will refer to as scenario 1, involved a single cohort of fish being tagged in five consecutive years starting at age 1 (i.e., at age i in year i for i=1, . . . , 5), and recaptured over the same five years. The number of tag releases was set to be 1000 at each age. Corresponding to the releases at each age, tag returns were generated from the observed and unobserved fishery components by using a Dirichlet-multinomial (D-M) distribution (Mosimann, 1962). The D-M distribution allows for overdispersion in the return counts by modeling the return probabilities as random Dirichlet variables (see Appendix A of Polacheck et al., 2006). It can be parameterized in terms of the return probabilities and an overdispersion factor, φ, that specifies the amount of extra variation in relation to multinomial data. For scenario 1, φ was set to be 3 (i.e., three times greater variance than a multinomial distribution). Other parameters were set as follows: F_i=0.15, M_i=0.2, λ_i=0.75, and δ_i=0.10, for i=1, ... , 5. Catch-at-age numbers (ages 1 to 5) for the observer component of the fishery were generated by using, first, a D-M distribution with P_1 = 100,000 and the same φ, δ_i, M_i, and F_i values as for the tag-recapture data. To these catch-at-age numbers, additional Gaussian aging error was added by using a constant CV of ν = 0.10 for all ages (i.e., $\eta_i = \nu \mathrm{E}(C_i^o)$).

The BPO model was fitted to the simulated tag-recapture and catch data by maximizing Equation 6. For this process, φ and ν were assumed to be known without error, and natural mortality was constrained to be the same at ages 4 and above (i.e., M_4=M_5; recall that only four natural mortality parameters can

Table 1

Description of simulation scenarios. In all scenarios, data were generated for a single cohort of fish tagged at ages 1 to A and recaptured at ages 1 to I, by using an age 1 cohort size of P_1, a constant natural mortality rate of M, a constant fishing mortality rate of F, a constant reporting rate in the unobserved fishery component of λ, a constant coefficient of variation for the catch aging error of ν, and an overdispersion factor of φ (values specified in table). Unless otherwise stated, data were generated by using N=1000 releases per year and a constant proportion of observer coverage of δ=0.20. In fitting the model, φ and ν were assumed to be known, and natural mortality was constrained to be the same for ages A–1 to I. Additional parameter constraints for each scenario are specified in the table. Bold text for a given scenario indicates a difference from scenario 1.

	Parameter values used to generate data								Parameter constraints imposed for model fitting
Scenario	A	I	P_1	M	F	λ	υ	φ	
1	5	5	100,000	0.20	0.15	0.75	0.10	3	λ constant
2	5	5	100,000	**0.40**	0.15	0.75	0.10	3	λ constant
3	5	5	100,000	0.20	**0.30**	0.75	0.10	3	λ constant
4	5	5	100,000	0.20	0.15	**0.50**	0.10	3	λ constant
5	5	5	100,000	0.20	0.15	**0.90**	0.10	3	λ constant
6	5	5	100,000	0.20	0.15	0.75	0.10	**1**	λ constant
7	5	5	100,000	0.20	0.15	0.75	0.10	**9**	λ constant
8	5	5	100,000	0.20	0.15	0.75	0.10	3	λ constant; ***M* constant**
9	5	5	100,000	0.20	0.15	0.75	0.10	3	λ constant; ***F* linear**[1]
10	5	5	100,000	0.20	0.15	0.75	0.10	3	**none**
11	**3**	**3**	100,000	0.20	0.15	0.75	0.10	3	λ constant
12	**3**	5	100,000	0.20	0.15	0.75	0.10	3	λ constant

[1] The line is parameterized in terms of F_1 and F_5; i.e., $F_i = F_1 + (i-1)\times(F_5-F_1)/4$ for i = 1,...,5.

be estimated with five release years). The unobserved tag reporting rate was also constrained to be constant over all recapture years (i.e., $\lambda_i=\lambda$ for i=1, ... , 5). It seems reasonable that the tag reporting rate would be constant, or at least similar, over the course of the experiment. Exceptions would occur if there was a significant change in the fishery or in tag-return promotional activities during this time, or if the fishery involves multiple fleets with different reporting rates so that the overall reporting rate would change if the distribution of catches among fleets changed. To account for such situations, year-specific reporting rates were allowed for in a later scenario (see next paragraph). The only other constraints imposed were simple bound constraints to keep all parameters positive and to keep the reporting rate from exceeding one. Thus, the parameters estimated were F_i (i=1, ... , 5), M_i (i=1, ... , 4), P_1, and λ.

Model performance will be affected by a large number of factors, including the following: 1) the parameter values used for the mortality rates, cohort size, reporting rates, catch aging error, and overdispersion factor; 2) model parameterization (i.e., whether parameters are assumed to vary with age, year, or both, or to have a particular functional form); and 3) the design of the tagging experiment (e.g., number of release and recapture years; number of releases per year; level of observer coverage). There are endless possibilities with regard to these factors; therefore we have chosen a number of scenarios that we feel are most illustrative for which to present results (Table 1). All of these scenarios use scenario 1 as a base but include a variation on one of the factors.

Scenarios 2 through 7 investigate changes to the parameter values (factor 1). In particular, scenario 2 increases the natural mortality rate, scenario 3 increases the fishing mortality rate, scenarios 4 and 5 decrease and increase the reporting rate, respectively, and scenarios 6 and 7 decrease and increase the overdispersion factor, respectively. Changes to the cohort size and variance of the catch aging error had less impact on the results and are therefore not included here.

Model parameterization (factor 2) can have a large effect on how well parameters can be estimated. For example, if natural mortality can be assumed to be constant across ages (this is a fairly common assumption in fishery models, at least over a limited range of age classes), then the precision and accuracy of the natural mortality-rate estimate should improve, which in turn may lead to improvements in other parameter estimates. Scenario 8 explores the benefits of having a constant natural mortality rate. Another standard way of reducing the number of parameters in fishery models is to model fishing mortality as a function of age by using an appropriate selectivity curve. Scenario 9 considers the situation where fishing mortality is constrained to be a linear function of age. Note that we parameterized the line in terms of F_1 and F_5 (i.e., $F_i=F_1+(i-1)*(F_5-F_1)/4$), because this made it easy to constrain the fishing mortality rates to be positive. Instead of imposing additional constraints, scenario 10 relaxes the assumption of a constant reporting rate and allows reporting rates to differ across years. This scenario has the maximum number of parameters that can be estimated by the model (i.e., the model is saturated).

In terms of experimental design (factor 3), the effect of varying the number of tag releases and the proportion of observer coverage is investigated in detail in the next section; therefore only variations to the numbers of release and recapture years are considered here. In particular, scenario 11 reduces the number of release and recapture years from five to three, whereas scenario 12 still has five recapture years but only three release years. For both scenarios, natural mortality was constrained to be equal at ages 2 and above, because only two natural mortality parameters can be estimated with three release years. For scenario 12, this meant constraining natural mortality to be equal at ages 2 to 5. In such a case, alternative constraints may be preferable, such as assuming natural mortality is a linear function of age. This was the approach taken in our application of the BP model to SBT data in Polacheck et al. (2006).

For each scenario in Table 1, 1000 sets of data were generated, as described above for scenario 1, and fitted by using the BPO model. For each parameter estimated, the percent median bias and the CV of the 1000 estimates were calculated, where percent median bias is defined as (median−true)/true×100% and CV is defined as SD/true (where SD denotes standard deviation). The median was used instead of the mean in calculating the bias because many of the parameter estimates had a skewed distribution, making the median a better measure of centrality (see "Results" section). The SDs of the parameter estimates obtained from the 1000 simulation runs (which approximate the true standard errors of the estimates) were compared with the standard error estimates obtained from the inverse Hessian matrix (these are obtained for every run; therefore we averaged the standard errors over the 1000 runs). The purpose was to see how much the Hessian-based standard errors were underestimated by applying a model that does not account for overdispersion in the data.

Trade-off between number of releases and observer coverage Of the factors that affect model performance, only the experimental design can be directly controlled by the researcher in a real application. Although model parameterization is superficially in the researcher's control, it is the true parameter values that will determine whether any parameter constraints are advantageous (i.e., imposing constraints on parameters that do not represent the true situation will lead to poorer model performance, not improved performance). Thus, in designing a tagging experiment and deciding how best to distribute resources, it would be very useful for the researcher to know the level of performance that can be achieved under different designs, as well as which design elements have the most influence on the results. Here, we illustrate how the BPO model can be used to provide such information. In particular, simulations are used to evaluate how well the parameters are estimated with different numbers of tag releases and different proportions of observer coverage, and to evaluate the trade-off between releasing more tags versus increasing observer coverage (i.e., to evaluate which leads to larger improvements in accuracy and precision of the parameter estimates).

Initially, simulations were carried out under scenario 1. For simplicity, the number of releases was kept the same for all release ages (i.e., $N_a=N$ for all a) and the proportion of observer coverage was the same over all recapture years (i.e., $\delta_i=\delta$ for all i). N was varied from 250 to 2500, and δ from 0.05 to 0.50. For each combination of N and δ, 1000 tag-recapture and corresponding catch data sets were generated, as described in the previous section, and fitted by using the BPO model. The results were used to evaluate how the percent median bias and CV of the parameter estimates changed as the number of releases and level of observer coverage changed.

For a true field study, the researcher would need to carry out such simulations using parameter values and model constraints that roughly represent the population and fishery dynamics for their situation. Our purpose was not to provide guidance on appropriate numbers of releases and observer coverage for any specific situation, but to illustrate how the model could be used to this end. Nevertheless, it is of interest to know whether the general findings using scenario 1 are likely to remain similar under other scenarios. The absolute levels of accuracy and precision that can be achieved will clearly depend on the scenario, but it is less clear whether the relative changes in these measures from increasing tag releases or increasing observer coverage will be highly scenario dependent. To investigate, we repeated the trade-off simulations using a subset of the other scenarios (4, 6, 8, 10, 11, and 12).

Results

Model performance

The biases in the medians of the parameter estimates were small for almost all parameters and scenarios (Table 2). A few of the natural mortality estimates had negative biases of greater than 5%, but this result more likely reflects the large variability and non-normality of these estimates (see next paragraph) than true biases.

Histograms of the parameter estimates revealed features that are important for evaluating biases. In particular, the natural mortality estimates often hit the lower bound of zero, and the proportion that did so was highest when the variability was largest (e.g., scenario 7, which has a high amount of overdispersion; Fig. 1). This feature makes it difficult to assess bias for these parameters and explains why the median biases seen in some of the natural mortality-rate estimates, such as scenario 7, are not likely to be meaningful. In scenarios 5, 7, and 10, the reporting-rate estimates often hit their

Table 2

Percent median bias (i.e., (median–true)/true×100%) of the parameter estimates for each scenario listed in Table 1. Results were based on 1000 simulation runs per scenario. M_i= natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ_i= tag reporting rate for fish captured at age i in the unobserved component of the fishery.

Scenario	M_1	M_2	M_3	M_4	F_1	F_2	F_3	F_4	F_5	P_1	λ_1	λ_2	λ_3	λ_4	λ_5
1	–1.0	–3.0	–1.5	–2.0	–1.3	–1.3	–0.7	–0.7	0.7	1.0	0.5	—	—	—	—
2	0.3	–1.3	–1.8	0.0	–0.7	–1.3	–0.7	0.0	0.7	1.0	0.0	—	—	—	—
3	0.0	–2.0	–1.5	0.5	–0.7	0.0	–0.7	–1.0	0.3	0.0	–0.1	—	—	—	—
4	1.5	–9.5	–4.0	–3.0	0.0	0.0	–0.7	1.3	1.3	0.0	–0.8	—	—	—	—
5	–2.0	–5.5	1.5	–0.5	0.0	–0.7	0.0	0.7	0.0	0.0	0.6	—	—	—	—
6	3.5	–1.5	–2.0	1.5	0.0	0.0	–1.3	–0.7	0.0	–1.0	0.4	—	—	—	—
7	–0.5	–13.0	–5.0	–16.0	–4.0	–2.0	–2.0	0.0	2.7	6.0	1.7	—	—	—	—
8	0.5	—	—	—	–2.0	–0.7	–0.7	–1.3	–0.7	1.0	0.4	—	—	—	—
9	1.5	–0.5	3.5	–0.5	–1.3	—	—	—	–0.7	1.0	0.5	—	—	—	—
10	–2.5	–3.0	–2.5	1.0	0.0	–1.3	–0.7	–0.7	1.3	0.0	–2.3	1.2	0.1	0.3	0.4
11	–5.5	–6.5	—	—	–0.7	–1.3	–3.3	—	—	1.0	1.6	—	—	—	—
12	1.5	–3.0	—	—	–0.7	–1.3	–0.7	–2.0	–2.0	1.0	0.4	—	—	—	—

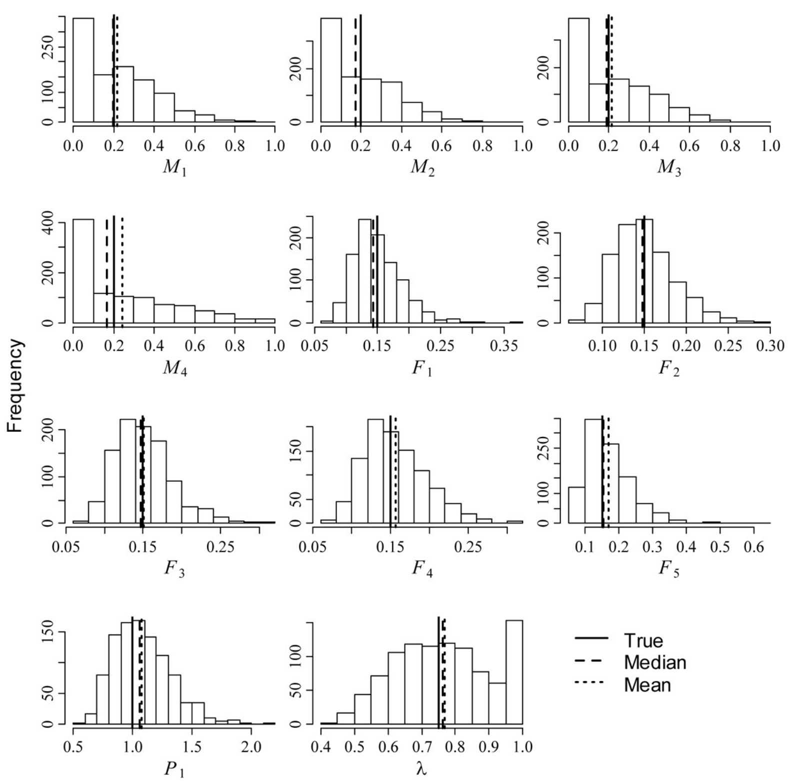

Figure 1

Histograms of the 1000 estimates obtained for each parameter under scenario 7 (see Table 1). The true parameter value, median, and mean of the estimates are indicated by vertical lines. M_i = natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1 = population size of tagged cohort at age 1 (in 100,000s); λ = tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 7).

upper bound of 1.0, and the frequency was greatest in scenario 5 where the true value was 0.90. Nevertheless, the median biases were still small. The fishing mortality estimates generally had distributions that were right skewed, and the degree of skewness became more pronounced at older ages. The skewness was usually small enough that the mean and median were still similar (e.g., scenario 1; Fig. 2). However, this was not always true. For example, in scenario 7 (Fig. 1), the median bias for F_5 was 2.7%, whereas the mean bias was 13.3%.

In Polacheck et al. (2006), we used mean bias to summarize simulation results obtained with the BP model. This meant that positive biases in the fishing mortality estimates that increased with age were reported, as well as positive biases in the natural mortality estimates. Had median bias been used instead, the bias results would have been similar to those presented here (i.e., negative biases in the natural mortality estimates, and only small biases in any of the fishing mortality estimates). In retrospect, we believe that the median provides a more reliable measure of bias. This is especially true in cases where the estimates have a skewed distribution, but should also be true in cases where a large proportion of the estimates fall on a bound.

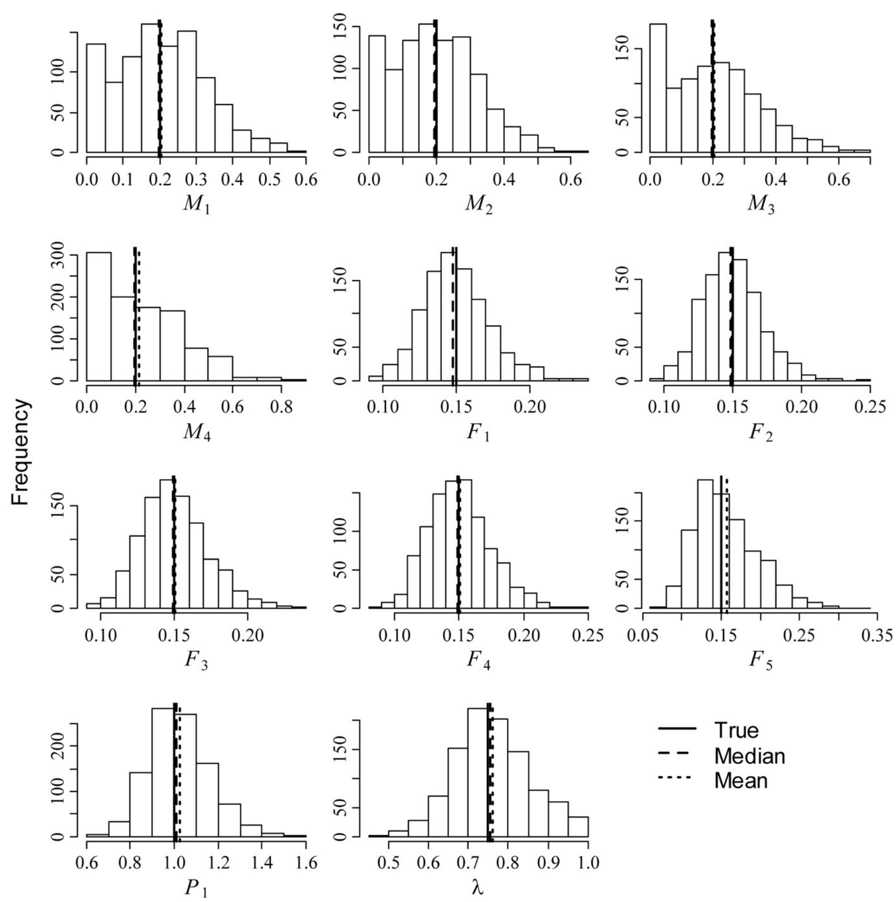

Figure 2

Histograms of the 1000 estimates obtained for each parameter under scenario 1 (see Table 1). The true parameter value, median, and mean of the estimates are indicated by vertical lines. M_i= natural mortality rate for age i fish; F_i = fishing mortality rate for age i fish; P_1 = population size of tagged cohort at age 1 (in 100,000s); λ= tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 1).

In regard to precision, we estimated fishing mortality rates, cohort size, and reporting rates with much greater precision (CVs generally in the range of 0.10 to 0.20) than that for the natural mortality estimates (CVs often exceeding 0.50) across all scenarios (Table 3). Only when natural mortality was constrained to be constant across ages (scenario 8) was reasonable precision achieved for this parameter (CV of 0.22). Of the fishing mortality parameters, the estimates for the oldest age of recapture (i.e., F_3 in scenario 11, F_5 in all other scenarios) always had the highest CV, and usually notably so.

Comparing the CVs for a specific scenario with those for scenario 1, we found that the results were generally predictable, at least in terms of direction (Table 3). For example, increasing the value used for the fishing mortality rate (scenario 3) or for the reporting rate (scenario 5) resulted in greater precision (i.e., lower CVs) for all parameters, because these changes lead to more tag returns. The results for scenario 2 were not instantly as intuitive. We would expect increasing natural mortality to give higher CVs because more fish would die naturally, leaving fewer tagged fish to be caught. Although small increases were observed in the CVs for the other parameters, large decreases were observed for the natural mortality estimates. This serves as a reminder that the CV is calculated in relation to the true parameter value, and therefore direct comparisons for parameters whose true values have been changed are more complicated. When the SDs of the natural mortality estimates from scenario 2 were compared with those from scenario 1 instead of the CVs, they did in fact increase (although this may in part be due to the fact that fewer estimates are truncated at their lower bound of zero when the value used for natural mortality is higher).

Table 3

Coefficient of variation (CV) of the parameter estimates for each scenario listed in Table 1. Results were based on 1000 simulation runs per scenario. M_i= natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ_i= tag reporting rate for fish captured at age i in the unobserved component of the fishery.

Scenario	M_1	M_2	M_3	M_4	F_1	F_2	F_3	F_4	F_5	P_1	λ_1	λ_2	λ_3	λ_4	λ_5
1	0.62	0.62	0.72	0.87	0.15	0.14	0.15	0.16	0.27	0.14	0.13	—	—	—	—
2	0.40	0.43	0.47	0.55	0.17	0.17	0.17	0.19	0.30	0.16	0.14	—	—	—	—
3	0.46	0.51	0.55	0.69	0.12	0.11	0.12	0.12	0.21	0.11	0.10	—	—	—	—
4	0.72	0.74	0.79	0.98	0.15	0.15	0.15	0.17	0.32	0.15	0.13	—	—	—	—
5	0.56	0.59	0.66	0.82	0.14	0.13	0.13	0.15	0.25	0.12	0.10	—	—	—	—
6	0.39	0.42	0.46	0.58	0.09	0.09	0.09	0.10	0.15	0.09	0.07	—	—	—	—
7	0.94	0.90	0.99	1.28	0.25	0.24	0.25	0.27	0.47	0.23	0.19	—	—	—	—
8	0.22	—	—	—	0.15	0.14	0.15	0.15	0.17	0.13	0.13	—	—	—	—
9	0.62	0.59	0.64	0.56	0.13	—	—	—	0.19	0.14	0.13	—	—	—	—
10	0.66	0.64	0.73	0.88	0.15	0.15	0.16	0.17	0.29	0.13	0.22	0.20	0.18	0.18	0.18
11	0.72	0.85	—	—	0.18	0.19	0.27	—	—	0.18	0.17	—	—	—	—
12	0.61	0.63	—	—	0.16	0.16	0.21	0.37	0.75	0.15	0.14	—	—	—	—

Several other points are worth noting. Changing the level of overdispersion in the data (scenarios 6 and 7) had the greatest influence on the CVs across all parameters. The only exception was that constraining natural mortality to be constant (scenario 8) had a greater effect on the CV of the natural mortality-rate estimation. Constraining natural mortality to be constant (scenario 8) not only reduced the CV of the M estimate substantially, but also the CV of the F_5 estimate. Similarly, constraining fishing mortality to be linear (scenario 9) substantially reduced the CV of the F_5 estimate, but also the M_4 estimate, and to a lesser degree the M_3 estimate. Interestingly, however, neither of these constraints affected the CVs of the P_1 and λ estimates. Also of interest is that allowing reporting rates to vary across years (scenario 10) had only a small effect on the precision of the mortality rate and abundance estimates. The reporting-rate estimates themselves were less precise and had a high tendency to hit the upper bound of one, but usually these parameters are not the ones of primary interest. Lastly, we note that having five recapture years but only three release years (scenario 12) resulted in much higher CVs for the F_3, F_4, and F_5 estimates, and increasingly so with age (with a CV of 0.75 for F_5). Thus, having more recapture years allows for more years of fishing mortality rates to be estimated, but these estimates quickly become uninformative unless the number of release years is also increased.

High correlations were present between many of the parameter estimates (Table 4; results are shown for scenario 1, but the patterns are very similar for all scenarios). Given the nature of the model, high correlations were expected, and have already been documented and discussed for the BP model in Polacheck et al. (2006). For example, to yield the same number of tag returns in a particular year, a higher estimate of fishing mortality for that year could be compensated by a higher estimate of natural mortality for the previous year, so that estimates of F_i and M_{i-1} tend to be positively correlated. Alternatively, it could be compensated by a higher estimate of the reporting rate; hence estimates of F_i and λ tend to be negatively correlated. When two parameters have highly correlated estimates, a large CV for one of these parameters will tend to mean a large CV for the other parameter. This may explain some of the results observed above. For example, in scenarios 1–7, 9, and 10, estimates of F_5 and M_4 were highly correlated; therefore the high uncertainty in F_5 is likely due to the very high uncertainty in M_4. An analogous statement can be made about F_3 and M_2 in scenario 11. The high correlation between estimates of F_5 and M_4 also explains why, in scenarios 8 and 9, constraints that improved the precision of one of these parameters also improved the precision of the other.

The Hessian-based standard error estimates in relation to the standard errors derived from the simulations are presented in Table 5. In all of the scenarios with φ=3, the Hessian-based standard errors were underestimated by a factor close to $\sqrt{3}$ = 1.73, and had a mean across all parameters and scenarios of 1.67 (ranging from 1.45 to 1.91). In the scenario with φ=1 (i.e., multinomial data), the Hessian-based and simulation-based standard errors were very similar, as expected. In the scenario with φ=9, the Hessian-based estimates were underestimated by a factor reasonably close to $\sqrt{9}$=3.0—the largest exception being a factor of 2.49 for P_1. Nevertheless, these results indicate that if φ can be estimated after fitting the model (e.g., from the residuals), then multiplying the Hessian-based standard error estimates by $\sqrt{\varphi}$ can provide improved, and perhaps adequate, estimates of the true standard errors. Further investigation of additional scenarios

Table 4

Correlation matrix for the parameter estimates obtained using scenario 1 (see Table 1). Results were based on 1000 simulation runs. M_i= natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ = tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 1).

	M_1	M_2	M_3	M_4	F_1	F_2	F_3	F_4	F_5	P_1	λ
M_1	1.00	–0.40	0.00	0.01	0.01	0.10	–0.07	–0.06	–0.03	0.55	0.14
M_2		1.00	–0.38	–0.06	–0.27	0.03	0.14	–0.06	–0.07	0.12	0.11
M_3			1.00	–0.35	–0.10	–0.25	0.08	0.26	–0.12	0.12	0.06
M_4				1.00	0.00	–0.04	–0.19	0.19	0.77	–0.01	0.06
F_1					1.00	0.47	0.40	0.37	0.27	–0.56	–0.61
F_2						1.00	0.49	0.38	0.27	–0.41	–0.63
F_3							1.00	0.52	0.25	–0.39	–0.63
F_4								1.00	0.59	–0.37	–0.59
F_5									1.00	–0.27	–0.39
P_1										1.00	0.67
λ											1.00

Table 5

Simulation-based standard error divided by Hessian-based standard error of the parameter estimates for each scenario listed in Table 1. Results were based on 1000 simulation runs per scenario. Recall that an overdispersion factor of φ = 3 was used in all scenarios except scenario 6, for which φ = 1, and scenario 7, for which φ = 9. M_i = natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ_i = tag reporting rate for fish captured at age i in the unobserved component of the fishery.

Scenario	M_1	M_2	M_3	M_4	F_1	F_2	F_3	F_4	F_5	P_1	λ_1	λ_2	λ_3	λ_4	λ_5
1	1.71	1.64	1.76	1.75	1.60	1.60	1.69	1.70	1.73	1.54	1.76	—	—	—	—
2	1.65	1.74	1.74	1.70	1.62	1.66	1.67	1.74	1.75	1.54	1.69	—	—	—	—
3	1.68	1.74	1.71	1.78	1.61	1.64	1.73	1.65	1.77	1.47	1.68	—	—	—	—
4	1.73	1.77	1.72	1.74	1.56	1.59	1.62	1.66	1.78	1.55	1.75	—	—	—	—
5	1.65	1.68	1.69	1.70	1.54	1.61	1.64	1.66	1.73	1.48	1.82	—	—	—	—
6	1.01	1.02	0.99	0.99	0.99	0.99	0.97	1.03	0.98	1.00	0.97	—	—	—	—
7	3.07	2.97	3.00	3.20	2.67	2.73	2.86	2.86	3.05	2.49	3.08	—	—	—	—
8	1.72	—	—	—	1.62	1.63	1.73	1.71	1.76	1.53	1.76	—	—	—	—
9	1.70	1.64	1.72	1.58	1.62	—	—	—	1.71	1.53	1.76	—	—	—	—
10	1.74	1.66	1.77	1.76	1.45	1.47	1.57	1.58	1.68	1.50	1.62	1.65	1.59	1.62	1.68
11	1.69	1.70	—	—	1.64	1.64	1.70	—	—	1.56	1.70	—	—	—	—
12	1.69	1.65	—	—	1.61	1.66	1.70	1.73	1.91	1.56	1.73	—	—	—	—

may allow for more accurate correction factors to be developed.

Trade-off between number of tag releases and observer coverage

We first concentrated on the results for scenario 1, and how changes in N and δ affected the accuracy of the parameter estimates. For all parameters, biases in the median of the estimates decreased rapidly as N increased, especially between 250 and 1000 releases (Fig. 3). Biases also tended to decrease as δ increased, especially for P_1, λ, and F_1 to F_3. In any case, only the biases in natural mortality estimates at the lowest release numbers ($N \leq 500$) were large enough to be of concern, and further investigation showed they were the result of a large proportion of the estimates falling on the lower bound of zero.

As seen in the previous section, evaluating biases could be complicated in some scenarios because of natural mortality estimates hitting a lower bound of zero, reporting-rate estimates hitting an upper bound of one, and fishing morality estimates having right-skewed distributions, especially at older ages. These problems became more pronounced as N and δ decreased, such that with N=250 and δ=0.05 the median and mean differed significantly for many parameters (e.g., M_4 had a median bias of –23% but a mean bias of +23%;

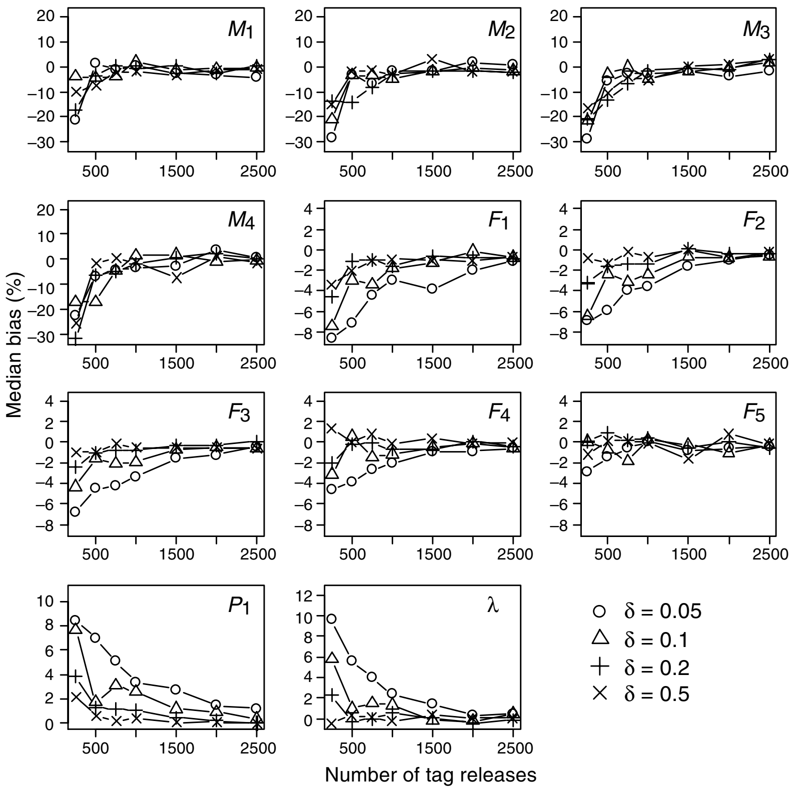

Figure 3

Effect of changing the number of releases at various proportions of observer coverage (δ) on the percent median bias (i.e., (median−true)/true×100%) of the parameter estimates. Results are shown for scenario 1 (see Table 1) and are based on 1000 simulation runs per combination of tag releases and observer coverage. M_i= natural mortality rate for age i fish; F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ = tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 1).

F_5 had a median bias of only −3% but a mean bias of +16%). Thus, with small N and δ, the real issue was not with biases, but with the non-normality and very high variability (as seen next) of the estimates.

We now consider how changes in N and δ affected the precision of parameter estimates. For a given value of δ, increasing N reduced the CVs of all estimates in an exponential fashion (Fig. 4). For the fishing mortality estimates, the rate of decline became greater with age, and was particularly notable for F_5. The CVs of the fishing mortality, abundance, and reporting rate estimates all decreased as δ increased; however, the natural mortality estimates did not change much. Overall, larger gains were achieved in the precision of the fishing mortality, abundance, and reporting rate estimates by increasing δ from 0.05 to 0.50 than by increasing N from 500 to 2500 (note that going from 250 to 500 releases led to significant decreases in the CVs of most parameter estimates). On the contrary, much larger gains were achieved in the precision of the natural mortality estimates by increasing N than by increasing δ.

As a specific example of using such simulation results to aid in the design of a tagging study, suppose a researcher's goal was to achieve a CV of 0.20 or lower in the estimate of abundance. This could be accomplished under scenario 1 with the following: N=250 and δ= 0.50; N=500 and δ=0.20; N=1000 and δ=0.10; or, N= 2000 and δ=0.05. If, in addition, the researcher's goal was to achieve a CV of 0.30 or lower in all of the fishing mortality estimates, then only the latter two of these options would still be acceptable.

Although the magnitude of the CVs varied significantly between scenarios (as seen in Table 3), the relative changes that resulted from increasing N or δ were very similar to those seen for scenario 1. The most significant difference came from constraining natural mortality to be constant (scenario 8), in which case the precision of the natural mortality parameter became influenced by changes in δ (Fig. 5).

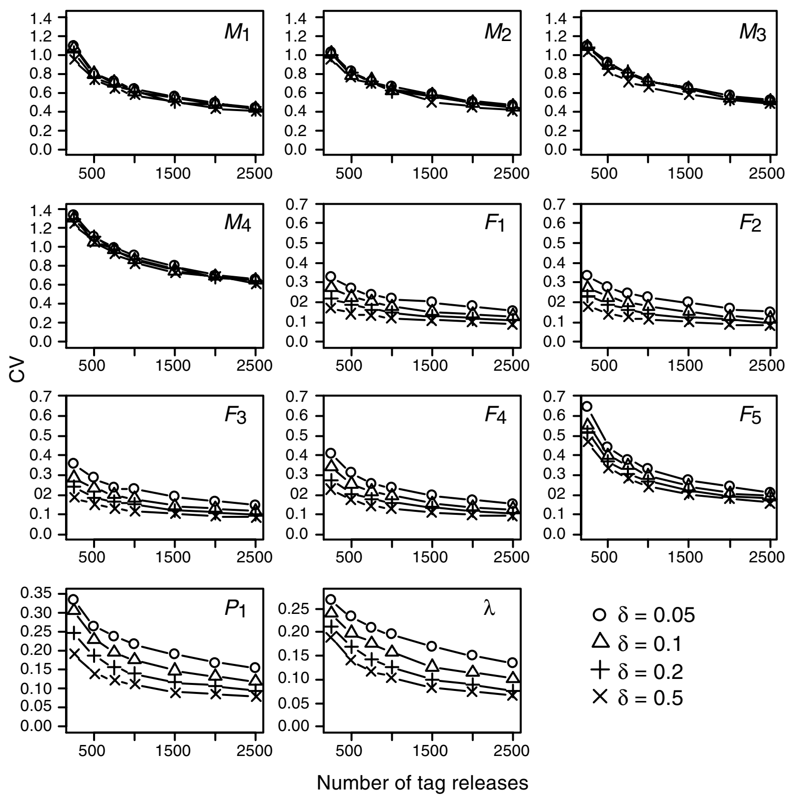

Figure 4

Effect of changing the number of tag releases at various proportions of observer coverage (δ) on the coefficient of variation (CV) of the parameter estimates. Results are shown for scenario 1 (see Table 1) and are based on 1000 simulation runs per combination of tag releases and observer coverage. M_i = natural mortality rate for age i fish; F_i = fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ = tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 1).

Discussion

The current article extends the integrated BP model for tag-recapture and catch data developed in Polacheck et al. (2006) to incorporate the estimation of reporting rates through observer data, which we refer to as the BPO model. This is an important and practical extension because nonreporting of tags is a serious problem in many commercial fisheries that needs to be accounted for in the model to obtain meaningful results, and observer data often provide the most viable means of doing so.

In the way the BPO model was formulated, increasing the level of observer coverage improves the parameter estimates not only by improving the reporting rate estimates, but also by improving the precision of the catch-at-age data. If all fish caught in the observer component were not sampled, then the improvements would not be expected to be as great. As an extreme case, the precision of the catch-at-age data could be assumed to be independent of the level of observer coverage, in which case increasing the level of observer coverage would only improve the parameter estimates through the reporting-rate estimates. However, it is difficult to envisage a situation where observers would not take age or length samples from at least a portion of the catches.

In the study by Pollock et al. (2002), where a standard Brownie model was modified to include the estimation of reporting rates when one component of a multicomponent fishery has observers (i.e., 100% reporting rates), the authors show how the overall likelihood for their model can be partitioned into two conditionally independent components. They argue that the reporting rates can be estimated by maximizing the second likelihood component, and then plugged into the first component to estimate the mortality rates, and that doing so provides the maximum likelihood estimates of the reporting rates and mortality rates for the joint likelihood. Although a similar partitioning could be done for the BPO model, the estimates obtained from

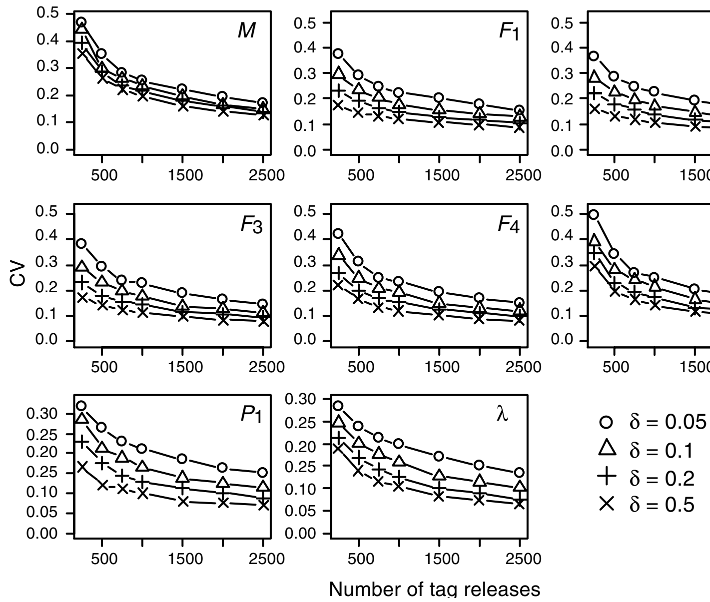

Figure 5

Effect of changing the number of tag releases at various proportions of observer coverage (δ) on the coefficient of variation (CV) of the parameter estimates. Results are shown for scenario 8 (see Table 1) and are based on 1000 simulation runs per combination of tag releases and observer coverage. M = natural mortality rate (assumed to be constant for scenario 8); F_i= fishing mortality rate for age i fish; P_1= population size of tagged cohort at age 1; λ = tag reporting rate for the unobserved component of the fishery (assumed to be constant for scenario 8).

maximizing the separate components would not be the overall maximum likelihood estimates because there is information in the catch data about the mortality rates. Furthermore, we assert that even in the model by Pollock et al. (2002), the estimates obtained from the two-step likelihood procedure are only the maximum likelihood estimates of the overall likelihood when the reporting rates are allowed to vary by year and age, and not, as the study would indicate, when there are any constraints on these parameters.

The BPO model allows for simultaneous estimation of age-specific fishing mortality rates, natural mortality rates, and reporting rates, as well as cohort size at first tagging, for a cohort tagged in consecutive years. All parameters appear to be estimated with reasonable accuracy, but the level of precision that can be achieved varies greatly, depending on the specifics of the population, the fishery, and the experimental design, and also on the parameter. Nevertheless, some general observations can be made based on our simulations. Cohort size appears to be estimated well in all situations (with a CV between 0.10 and 0.20 in the majority of scenarios considered). With the exception of the oldest age of fish at recapture, the fishing mortality rates also tend to be estimated with good precision (CVs of less than 0.20 achievable in many situations). In general, natural mortality is estimated poorly in comparison to the other parameters, with CVs above 0.60 in many cases. If, however, natural mortality can be assumed constant over enough release years (or otherwise constrained), then it too can be estimated with reasonable precision (e.g., CV on the order of 0.20 for our scenario 8 with 1000 releases per year).

Reducing the number of parameters that need to be estimated through imposing parameter constraints can greatly improve the accuracy and precision of the estimates. However, this is only true if the constraints imposed approximate reality; for example, modeling natural mortality as a constant will not lead to better parameter estimates if in fact natural mortality changes significantly with age. In practice, standard model selection techniques, such as Akaike's information criterion (AIC; Akaike, 1974) and its many variations (e.g., AIC_c for small sample sizes, QAIC for overdispersed data; see Burnham and Anderson, 1998, and references therein), can be used to determine which parameter constraints are most supported by the data.

For ease of presentation, the model was developed for, and applied to, one cohort of tagged fish. In practice, it is likely that several cohorts (i.e., age classes) would be tagged in each year of tagging. If all parameters being estimated are both year- and age-dependent, then modeling multiple cohorts simultaneously will give very similar results to modeling each cohort individually. The results would be identical if the catch-at-age data for all cohorts were modeled as independent, but there is likely to be correlation between catch estimates for different cohorts in the same year that should be accounted for. If some constraints can be put on the parameters, such as natural mortality varying only with age or fishing mortality following an age selectivity curve, then precision in the parameter estimates should improve. For example, we re-ran the simulations for scenario 1 as described in the model performance section, but with data generated for three cohorts instead of one. In fitting the model, both natural mortality and fishing mortality were allowed to vary only with age. In comparison to the CVs obtained for scenario 1 with one cohort (Table 3), the CVs obtained with three cohorts were roughly 35–40% less for all parameters. Again, to determine which parameter constraints are most appropriate in a real situation, standard model selection procedures such as AIC can be used.

As was illustrated, the BPO model can be used to evaluate the effect of releasing more tags versus increasing observer coverage on the precision and bias of the parameter estimates. Because these programs can be costly to run and resources are usually limited, it is useful to have a statistical framework for comparing how alternate allocations of resources affect the results that can be achieved. Our results confirm the general conclusion of Polacheck and Hearn (2003) that it is important to ensure both adequate numbers of tag releases and adequate observer coverage (the latter for robust estimation of reporting rates, as well as for improved estimation of catch-at-age numbers in our model). However, while Polacheck and Hearn (2003) found a relatively direct trade-off between the level of observer coverage and number of tag releases with their approximate model, we found with our more comprehensive model that the trade-off depends on the parameters of interest. In particular, greater improvements could generally be achieved in the precision of the fishing mortality and cohort size estimates by increasing the proportion of observer coverage than by increasing the number of releases. On the contrary, much larger gains were achieved in the precision of the natural mortality-rate estimates by increasing the number of tag releases than by increasing the proportion of observer coverage. Although the results will be highly case-specific, these general observations were true in all of the scenarios we considered, and we expect they will hold true in a fairly wide range of scenarios. That being said, the purpose of the simulations was not to draw any specific conclusions, but to illustrate how the model can be used to provide practical guidance about the experimental design of a tagging study.

A version of the BPO model has been used to provide advice to the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) on the levels of observer coverage and tag releases necessary to achieve their objectives for a long-term tagging program conducted on SBT. To make the model more closely resemble the situation for SBT it was necessary to extend the model to a two-fishery situation with a purse-seine fishery and a longline fishery, where tag reporting rates were estimated from planted tags in the purse-seine fishery and from observer data in the longline fishery. Simulations, similar to those presented here, were conducted with input parameter values that best simulate the situation for SBT. The results showed that the numbers of tags that were being released each year were adequate, but that an increase in the CCSBT's target level of observer coverage from 10% to about 30% was required to meet the objectives of the program regarding precision of the mortality-rate estimates.

In summary, the model presented here provides a robust statistical framework for obtaining joint estimates of mortality rates and abundance from tagging data in situations where observers are present in the fishery. The model can be used to provide insight into design issues for those starting up new, or modifying current, tagging and observer programs for the purposes of estimating mortality rates and abundance.

Acknowledgments

We thank K. Pollock, D. Peel, and three anonymous reviewers for constructive comments and suggestions on drafts of this manuscript. The Australian Fisheries Research and Development Corporation (FRDC) provided funding support for this research.

Literature cited

Akaike, H.
1974. A new look at the statistical model identification. IEEE (Inst. Electrical and Electronics Engineers) Transactions on Automatic Control 19:716–723.

Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson.
1985. Statistical inference from band recovery data: a handbook, 2nd ed., 305 p. U.S. Fish Wildl. Serv. Resour. Publ. 156.

Burnham, K. P, and D. R. Anderson.
1998. Model selection and inference: a practical information-theoretic approach, 353 p. Springer-Verlag, New York, NY.

Cormack, R. M.
1964. Estimates of survival from the sighting of marked animals. Biometrika 51:429–438.

Frusher, S. D., and J. M. Hoenig.
2001. Estimating natural and fishing mortality and tag reporting rate of southern rock lobster *Jasus edwardsii* from a multiyear tagging model. Can. J. Fish. Aquat. Sci. 58:2490–2501.

Hampton, J.
2000. Natural mortality rates in tropical tunas: size really does matter. Can. J. Fish. Aquat. Sci. 57:1002–1010.

Hearn, W. S., T. Polacheck, K. H. Pollock, and W. Whitelaw.
1999. Estimation of tag reporting rates in age-struc-

tured multicomponent fisheries where one component has observers. Can. J. Fish. Aquat. Sci.56:1255–1265.

Hoenig, J. M., N. J. Barrowman, W. S. Hearn, and K. H. Pollock.
1998a. Multiyear tagging studies incorporating fishing effort data. Can. J. Fish. Aquat. Sci. 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. Can. J. Fish. Aquat. Sci. 55:1477–1483.

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

Mosimann, J. E.
1962. On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. Biometrika 49:65–82.

Polacheck, T., J. P. Eveson, G. M. Laslett, K. H. Pollock, and W. S. Hearn.
2006. Integrating catch-at-age and multi-year tagging data: a combined Brownie and Petersen estimation approach in a fishery context. Can. J. Fish. Aquat. Sci. 63:534–538.

Polacheck, T., and W. Hearn.
2003. Designing tagging programs for pelagic longline fisheries: trade-offs between the number of releases and observer coverage. N. Am. J. Fish. Manage. 23: 810–821.

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

Pollock, K. H., W. S. Hearn, and T. Polacheck.
2002. A general model for tagging on multiple component fisheries: an integration of age-dependent reporting rates and mortality estimation. Environ. Ecol. Stat. 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. N. Am. J. Fish. Manage. 21:521–532.

Pollock, K. H., J. M. Hoenig, and C. M. Jones.
1991. Estimation of fishing and natural mortality when a tagging study is combined with a creel survey or port sampling. Am. Fish. Soc. Symp. 12:423–434.

Schwarz, C. J., and G. A. F. Seber.
1999. Estimating animal abundance: review III. Statist. Sci. 14:427–456.

Seber, G. A. F.
1965. A note on the multiple recapture census. Biometrika 52:249–259.
1982. The estimation of animal abundance and related parameters, 2nd ed., 654 p. Macmillan Publ. Co., New York, NY.