Table of Contents

List of Tables

List of Figures

List of Appendices

A1. Intellectual Property

A2. Staff

A3. An overview of different tagging experimental designs for estimating population parameters in commercial fisheries

A4. Southern bluefin tuna (*Thunnus maccoyii*) tag-recapture and catch data: details of data compilation

A5. An integrated Brownie and Petersen model for estimating mortality rates and population size in a fisheries context with known reporting rates

A6. Estimating the bias in the estimator of population size in a simple tag-recapture and catch model

A7. An integrated Brownie and Petersen model for estimating mortality rates and population size in a fisheries context with unknown reporting rates

A8. An evaluation of the consequences of aging errors in the context of a Brownie tag-recapture model

A9. Exploring trade-offs in experimental design of a 2-fishery integrated tag-recapture and catch model for estimating mortality rates and abundance

A10. Incorporating spatial structure and movement into an integrated tag-recapture and catch model using a generic movement model

A11. Incorporating spatial structure and movement into an integrated tag-recapture and catch model using movement dynamics representative of southern bluefin tuna

A12. Tagging in a spatial context: design and analysis considerations

A13. Parameter identifiability in spatial Brownie tag-recapture models

A14. Estimating tag shedding parameters from 1990s southern bluefin tuna tagging data

A15. Estimation of mortality rates and abundance for southern bluefin tuna using tag-return and catch data from 1991 to 1997

A16. Application of spatial tag-recapture and catch models to southern bluefin tuna data from the 1990s

A17. An evaluation of abundance estimates from tagging programs when tag returns are only available from one component of a multi-component fishery: an example based on the 1990s southern bluefin tuna tagging program

A18. Review of methods for estimating tag reporting rates and their applicability to southern bluefin tuna longline fisheries

A19. Updated estimates of tag reporting rates for the 1990s southern bluefin tuna tagging experiments

A20. Review of historical southern bluefin tuna tagging data for potential application of Brownie mortality estimation models

2002/015	**Estimation of mortality rates from tagging data for pelagic fisheries: analysis and experimental design**

Principle Investigator: Tom Polacheck

Address: CSIRO Marine and Atmospheric Research
P.O. Box 1538
Hobart TAS 7000
Tel: 03 6232 5312   Fax: 03 6232 5012

Objectives:

1. To provide a robust estimation framework for estimating mortality rates for tuna stocks from conventional tag-recapture data including associated confidence intervals.
2. To apply the estimation framework to tag-recapture data for southern bluefin tuna (SBT).
3. To utilize the estimation framework to provide guidance in the design of future tagging experiments – in particular to examine the trade-offs between the number of tags released, levels of observer coverage and spatial/temporal distribution of tags.
4. To explore the utility of archival tag data to improve the estimation of mortality rates from conventional tagging experiments in terms of incomplete mixing.

Non-technical Summary

Outcomes Achieved

This project has resulted in a comprehensive modelling framework for estimating mortality rates and abundance from tag-recapture and catch-at-age data. This framework was used to evaluate trade-offs in resources dedicated to the various experimental design components, and the results were considered by the CCSBT Scientific Committee in its design, review and recommendations for improving the conventional tagging program currently being conducted under the CCSBT Scientific Research Program. The modelling framework was applied to southern bluefin tuna data from tagging experiments conducted in the 1990s to provide estimates of mortality rates (natural and fishing) and abundance that are the most comprehensive to date. It is anticipated that the methods developed in this project will continue to be used in future analyses conducted for the CCSBT to aid in the evaluation of the SBT stock and in the provision of management advice.

Data from conventional tagging experiments have been and are becoming increasingly important for providing estimates of mortality rates in stock assessments of large pelagic fishery resources as they provide a powerful, and perhaps only, alternative to relying on fishery dependent catch rate (CPUE) indices of abundance. While there is an extensive scientific literature on tag-recapture experiments, a comprehensive framework for the analysis and design of tagging experiments in the

pelagic fishery context does not exist. This has meant that in the design phase of tagging experiments, it has not been possible to provide scientifically rigorous advice on various design components, such as numbers of tags to release, spatial and temporal distribution of tag releases, and necessary levels of observer coverage. In addition, extensive tagging experiments have been conducted on juvenile southern bluefin tuna (SBT) during the past several decades, particularly during the early 1990s, but a comprehensive analysis of the data taking into account all of the major potential sources of heterogeneity has not been completed because of the lack of an appropriate estimation framework. As such, development of such a framework is essential to ensure that the full value from both past and future experiments is realized.

In this project, we succeeded in developing a comprehensive modelling framework that integrates tag-recapture data and catch-at-age data that can be used to provide rigorous estimates of mortality rates and abundance. The general framework was developed in a non-spatial context and was later expanded to include a spatial dimension in order to provide unbiased estimates in situations where spatial heterogeneity (i.e. systematic incomplete mixing over large spatial scales) is a significant issue. We also demonstrated how issues of reporting rates, tag shedding and (unsystematic) incomplete mixing can be dealt with in the general framework.

The modelling framework was used to evaluate relative trade-offs in dedicating resources to various experimental design components. In particular we focussed on evaluating the relative benefits of releasing more tags versus improving the catch-at-age data and/or increasing the level of observer coverage (and thereby improving the reporting rate estimates as well as estimates of the size/age distribution of the commercial catches). Such results can be useful in providing guidance on the appropriate distribution of resources among different design components of a tagging experiment, and can potentially improve the efficiency and cost effectiveness of future experiments.

The modelling framework was applied to SBT tagging data from experiments conducted in the 1990s and catch-at-age data from the corresponding time period to provide estimates of juvenile mortality rates (natural and fishing) and abundance at age 1. This is the most comprehensive analysis of these data to date, and the results suggest that:
- natural mortality for SBT at age 1 is quite high (~0.4) and decreases to about 0.2 by age 5;
- fishing mortality rates in the 1990s were generally close to zero for ages 1 and 2, were greatest at ages 3 and 4, and declined at age 5;
- overall fishing mortality decreased in the first couple of years of the 1990s then increased fairly steadily from 1994 to 1997 to relatively high levels (~0.4 at ages 3 and 4);
- population abundance at age 1 appears to have decreased by over 50% in 1993 and 1994 compared to earlier in the 1990s.

The historic SBT tagging data from experiments conducted prior to the 1990s were considered, but we concluded that these data have limited potential to provide quantitative mortality rate and abundance estimates, and are not directly amenable to the integrated estimation framework developed within this report. This is because of

limitations in the timing and location of releases, lack of information for estimating reporting rates and systematic non-mixing issues in the 1960s data.

The results of this research have already been of direct benefit to the CCSBT Scientific Committee. Several of the appendices in this report have been submitted to meetings of the Scientific Committee and these papers have been integral in the review of the current conventional tagging program and in developing recommendations for improving the program. The methods developed here also provide a robust and improved basis for maximizing the information that can be derived from tagging experiments – in particular the integration of tag-recapture data with observer and catch data provides a comprehensive method for directly estimating mortality rates (both natural and fishing) and population sizes. These are the primary quantities required to be estimated in stock assessments, and having an approach for directly estimating them that does not require catch rate (CPUE) data provides a potentially powerful alternative for augmenting traditional stock assessment methods.

Areas where there is potential for further improvements in the modelling framework developed here have been identified. In particular, further work to fully address the issue of incomplete mixing, both systematic and unsystematic, is required. With regard to unsystematic incomplete mixing, better models and methods for estimating the level of overdispersion in the tag-recapture data are needed. With regard to systematic incomplete mixing, it would be useful to consider a wider range of spatial and movement dynamics, with the acknowledgement that the most appropriate spatial model will be situation-specific. In addition, the results in this report point to the critical need to improve the data for estimating tag reporting rates and catch-at-age data for SBT fisheries if the potential of the current large-scale CCSBT SRP tagging experiments is to be fully realized.

Keywords: tag-recapture, mortality rates, Brownie model, Petersen estimator, southern bluefin tuna, spatial modelling

Acknowledgements

Background

Reliable indices of abundance and/or fishing mortality rates are a critical component of the stock assessment process. In many fisheries, particularly pelagic ones, catch per unit effort (CPUE) is often the only, or primary, index of abundance available. The problems and dangers of relying on CPUE are well known and documented (e.g., northern cod). The development of reliable non-CPUE indices has long been recognized as a high priority both by fishery scientists, in general, and by scientists involved in the assessment of international tuna resources (e.g., CCSBT Scientific Committee). However, the development of fishery independent and non-CPUE based indices is a difficult problem for a widely dispersed stock such as SBT. Tagging experiments have been recognized as offering one (and perhaps the only viable) alternative to relying on CPUE for wide ranging pelagic species.

Conventional tagging experiments are a conceptually simple and powerful mechanism for providing estimates of critical parameters (e.g., fishing mortality rates) in fishery stock assessments. In theory, all that is required is to release a sufficient number of tagged fish and monitor the number of tags that are returned over time, which then provides a measure of fishing intensity. However, while conceptually simple, there are a number of key requirements and assumptions that must be addressed in the experimental design and analysis of the data. Failure to adequately address these can result in highly imprecise and inaccurate estimates or, in the worst case, non-interpretable results. The main issues in design and analysis of tagging experiments are concerned with heterogeneity in tag return rates as a result of factors such as tag-shedding, non-reporting and incomplete mixing (which can include emigration and/or immigration in situations involving more than a single population).

Extensive conventional SBT tagging experiments have been conducted in the past. In particular, the multi-year, multi-cohort tagging in the 1990s has provided a very informative data set that has been an important contributor in recent stock assessments. However, a comprehensive analysis of these data taking into account all of the major potential sources of heterogeneity (e.g., tag shedding, non-reporting and non-mixing) has not been undertaken. Extensive additional and on-going conventional tagging experiments are being undertaken within the CCSBT as part of its agreed Scientific Research Program.

Finally, a major source of uncertainty in any tagging program is the extent of incomplete mixing between tagged and untagged fish in the population. Archival tags have been identified as one potential source of information in this context because these tags provide information on where a fish has been between its time of recapture and release. In this regard, the CCSBT Scientific Committee has suggested that archival tagging be done to augment the currently agreed-on conventional tagging experiments under the CCSBT Scientific Research Program. However, there has been no analytical work done to determine how this information might be incorporated into the actual analysis of data from conventional tagging experiments, nor on the relative sample sizes and cost/benefit trade-offs involved.

Need

Data from conventional tagging experiments have been and are becoming increasingly important for providing estimates of mortality rates in stock assessments of large pelagic fishery resources as they provide a powerful, and perhaps only, alternative to relying on CPUE indices of abundance. There is an extensive scientific literature on tag-recapture experiments; however, a comprehensive framework for the analysis and design of tagging experiments in the pelagic fishery context does not exist. This has meant that in the design phase of tagging experiments, it has not been possible to provide scientifically rigorous advice on matters such as the relative trade-offs between the number of tags to release, the spatial/temporal distribution of tags and the levels of observer coverage (required for estimation of reporting rates and for estimating the size/age distribution of the commercial catches). In addition, a comprehensive analysis of the extensive data from past SBT tagging experiments (particularly those conducted in the 1990s) taking into account all of the major potential sources of heterogeneity has not been completed because of the lack of a comprehensive estimation framework. Thus, the full benefit of the large research investment from these experiments has not been realized. It is also anticipated that conventional tagging programs will be a major future source of assessment information in both the eastern and western tuna fisheries. As such, an appropriate framework for the design and analysis of tagging experiments is needed to ensure that the full value from both past and future experiments is realized. Finally, there is increasing data accumulating from archival tags. There is a need to understand the role, and the relative cost/benefits, that archival tags can contribute to the overall estimation of mortality rates to past and future conventional tagging experiments.

Objectives

1. To provide a robust estimation framework for estimating mortality rates for tuna stocks from conventional tag-recapture data including associated confidence intervals.
2. To apply the estimation framework to tag-recapture data for southern bluefin tuna.
3. To utilize the estimation framework to provide guidance in the design of future tagging experiments – in particular to examine the trade-offs between the number of tags released, levels of observer coverage and spatial/temporal distribution of tags.
4. To explore the utility of archival tag data to improve the estimation of mortality rates from conventional tagging experiments in terms of incomplete mixing.

Methods

Estimation Framework

Before determining the estimation framework to use in this study, we first undertook a review of different tagging experimental designs and estimators that could potentially be used for estimating population parameters in commercial fishery situations. Based on this review, we concluded that experimental designs based on multi-year tagging of the same cohort provide the most powerful experimental design in terms of the resulting ranges of parameters that can be estimated. Such an experimental design requires tagging for at least three consecutive years and can provide estimates of fishing and natural mortality rates, and potentially population size[1] when integrated with catch data. Since the experimental design requires tagging of the same cohort in multiple years, it is most applicable to juveniles where age can be inferred from size. It should be noted that even if tagging is done only on juveniles, the tagging experiment can still potentially provide useful information on older ages as fish continue to be recaptured, as long as sufficient numbers are tagged and tag retention rates are reasonably high. We also note that that if mortality rates (both fishing and natural) are not age-dependent, then tagging and tracking of the same cohorts is not essential. In such cases, a multi-year tagging experimental design could be used for the adult component of a stock where aging may be problematical.

As part of this study, we extended the traditional Brownie estimator for multi-year tagging experiments (Brownie et al. 1985) to incorporate estimates of the catch-at-age from the commercial catches. The integration of a multi-year tagging experiment with catch-at-age data from the commercial fishery can provide robust estimates of stock size as well as fishing mortality rates (Appendix 5). This integrated approach was used as the basic framework in most of the analyses conducted within this project. The estimates of stock size from this integrated approach are not reliant on having relative abundance indices. This is particularly important in the context of assessment for large pelagic tuna stocks since it provides an approach for estimating stock sizes which is independent of CPUE indices. Such experimental designs are particularly

---

[1] Population size in this context and generally throughout this report refers to the size of a cohort or group of cohorts (those which are tagged).

applicable to tagging experiments for juvenile SBT given their availability to surface fisheries over three age classes (i.e., ages 1-3). This means that it is viable to tag large numbers of individuals from the same cohort over multiple years.

There are a number of factors that need to be accounted for in order for tagging experiments to provide robust and reliable estimates. These include:

1. reporting rates of recaptured tags;
2. spatial heterogeneity/incomplete mixing;
3. tag shedding rates;
4. accuracy of catch-at-age data.

In the initial development of estimators from tagging experiments, these factors are frequently ignored (e.g., reporting rates are assumed to be 100%). While ignoring these factors in the initial development provides an indication of what can ideally be achieved with a particular experimental design, it is important in any actual implementation to ensure that appropriate data are collected to address these issues and, where necessary, to have an integrated estimation framework for incorporating these data. In particular, in the design phase of a tagging experiment having such a framework is critical for being able to assess alternatives with respect to the effort devoted to numbers of tag releases, their distribution in space and time, and to the estimation of reporting rates. In addition, when catch-at-age data are integrated with multi-year tagging data to estimate mortality rates and abundance, it is important to consider the uncertainty associated with the estimates of the catch-at-age. Catch-at-age data are also required for estimating reporting rates (particularly in a multi-fishery context). Therefore, in considering alternative experimental designs, the effort devoted to the estimation of catch-at-age is a further dimension that needs to be considered. In Appendices 7, 9, 10, 11, 14 and 15, we have developed estimators that deal with these issues and that allow potential experimental designs to be evaluated.

Reporting Rates

The estimation of reporting rates is a particularly critical component in the design of tagging experiments and the precision with which these are estimated can limit the overall precision that can be obtained from a tagging experiment. Attempting to estimate reporting rates retrospectively is problematical. As such, it is critical that a strategy for estimating reporting rates is included within the initial experimental design. There are four basic approaches to the estimation of reporting rates. The applicability of these different approaches in a fishery context is dependent upon the actual nature of the fishery. These approaches are reviewed in Appendix 18. It should be noted that reporting rates are likely to differ among fishery components. A different method can be used to estimate reporting rates in each component, and all of the results incorporated into the final analysis. In terms of tuna longline fisheries, Appendix 18 concludes that the use of observers to estimate reporting rates is currently the only viable approach, while tag seeding is the most viable approach for surface fisheries involving farming or canning. There are two fundamental approaches for incorporating reporting rates into the integrated tagging and catch-at-age model. First, the reporting rates can be estimated from an independent analysis, and then the estimates and their standard errors can be inputted into the model (e.g., Appendices 9 and 15). Second, the estimation of reporting rates can be integrated

directly into the model (e.g., Appendices 7, 10 and 11). In the case of tag seeding, a likelihood estimator for the reporting rates can be developed from the tag seeding data independent of the tagging and catch-at-age data. In this case, the tag seeding likelihood can simply be multiplied with the tagging and catch likelihoods to provide an overall estimator (Appendix 11). However, in situations where observers are used to estimate reporting rates, the estimation is more complex since the reporting rate estimates use some of the tagging and/or related catch data. In this case, the likelihoods for the tagging and catch data need to be modified. Appendix 7 expands the non-spatial integrated tagging and catch model developed in Appendix 5 to incorporate the estimation of reporting rates for the situation in which observers monitor a portion of the catches; Appendices 10 and 11 do so in a spatial context. We use the expanded model in Appendix 7 to address experimental design issues such as the trade-off in the number of releases relative to the level of observer coverage. In Appendix 9, we consider a two-fishery situation in which reporting rates are derived from observer data in one fishery and from tag seeding data in the other. The estimator in Appendix 9 can be used similar to that in Appendix 7 to look at design trade-offs.

Spatial Heterogeneity – Systematic Incomplete Mixing

The issue of spatial heterogeneity and incomplete mixing is another critical component that needs to be considered in the design of tagging experiments. In Appendix 10, we expand the integrated multi-year tagging and catch-at-age model to allow for abundance and fishing mortality to differ among defined regions, and for fish to move among these regions. Within this spatial estimation framework, we have also developed the appropriate likelihoods for estimating reporting rates from observer data. This allows us to evaluate alternative experimental designs and to examine the trade-offs between the number of tags released, the level of observer coverage and the spatial/temporal distribution of tags.

The estimation framework developed in Appendix 10 is generic, without any restrictions on the movement patterns (i.e., transition probabilities) between areas. As long as the underlying spatial structure and time periods in the model are appropriate and transition probabilities are allowed to vary with age and year/time-period, this framework should be able to reasonably represent most movement dynamics. However, the cost is high in terms of the number of parameters to be estimated and the demand for adequate tagging in many areas, particularly as the number of areas and time periods increases. This can easily lead to intractable estimation and design problems. In addition, such a generic structure provides little basis for prediction as it assumes no consistency in the movement dynamics with time. Thus, imposing some form of spatial and/or temporal structure on the transition probabilities is highly desirable. The structure imposed will depend on the population being modelled, however it is not unreasonable or unrealistic to expect that there are some consistent underlying biological processes that determine how and when animals move. In Appendix 11, we develop two examples of such models. These examples were constructed to represent an idealized representation of 'SBT-like' spatial dynamics and take advantage of known (or assumed) prior information on the movement dynamics of juvenile SBT to reduce the parameterization of the transition probabilities. In both models, we define four areas and two seasons per year with movement only allowed between certain areas at the end of certain seasons.

However, in one of the models movement is represented as a seasonal Markov (i.e., memoryless) process, whereas in the second model there is site-fidelity in the movement dynamics. As these two models were constructed to represent SBT-like fishery situations, the estimation framework incorporates both observer and tagging seeding data for estimating reporting rates in different fishery components within the model.

Overdispersion – Unsystematic Incomplete Mixing

The tag-recapture component in the basic estimation framework assumes a multinomial distribution for the tag returns. This assumption is a common feature in most tag-recapture estimation models (e.g., Seber 1982). However, this is only valid if all fish of a particular age have the same probability of being caught. If there is unsystematic incomplete mixing of tagged and untagged fish[2] (e.g., if fish tagged in the same school and/or in close proximity on the same day have positively correlated recapture probabilities), then the numbers of returns at age will have more variability than a multinomial distribution would predict, referred to as overdispersion. Differential age/size selectivities among fishing vessels will also contribute to overdispersion if tagged fish are not homogeneously mixed within the untagged population. We extended the basic non-spatial estimation framework to incorporate overdispersion in the tag-return data using a Dirichlet-multinomial distribution (Appendix 9). We also expanded the spatially-explicit estimation framework to allow for overdispersion within each spatial stratum using the Dirichlet-multinomial distribution (Appendix 10). Essentially, the probabilities of return corresponding to a particular set of releases are modelled as Dirichlet random variables with a variance parameter that governs the level of overdispersion. Then, the numbers of returns conditional on the probabilities of return follow a multinomial distribution, and the unconditional numbers of returns follow the compound distribution referred to as the Dirichlet-multinomial (see Annex A of Appendix 9). Through simulations, we explored the effect of incorporating overdispersion in the tag return data on the estimation of the mortality rate and abundance parameters (Appendix 9).

Tag Shedding

Double-tagging experiments are the primary approach used for estimating shedding rates and extensive work has been done on the problem of modelling and estimating shedding rates from such experiments (Kirkwood and Walker 1984; Barrowman and Myers 1996; Cadigan and Brattey 2003). Methods for estimating tag shedding rates based on double tagging were reviewed (Appendix 14). Based on that review, models that had parameters which allowed for an initial immediate shedding and continuous shedding rate were explored for use with the 1990s SBT tagging experiments. We also developed an approach that allowed taggers with statistically similar shedding

---

[2] Unsystematic incomplete mixing is meant to refer to situations where there is still large amounts of mixing among tagged and untagged fish and the pattern of mixing has a large random component such that on average the probability of recapture of tagged and untagged fish are the same (often referred to as 'patchiness'). This should be distinguished from the situation discussed in the previous section on spatial heterogeneity where there is a systematic and repeatable pattern of incomplete-mixing between tagged and untagged fish – for example, if all tagging was done late in the season in one location and fish in that location and time period tend to remain in one part of the stock's overall range. Such systematic incomplete mixing will induce biases into the population and mortality estimates if it is not accounted for within the estimation model, whereas unsystematic incomplete mixing should not bias the parameter estimates but will increase the variance of the estimates.

rates to be combined (Annex A of Appendix 14) and applied this approach to the return data from the 1990s tagging experiment. We reviewed approaches for integrating tag shedding rates into the basic estimation framework from multi-year tagging programs and developed a maximum likelihood approach for combining estimates of tag shedding rates derived from independent analyses of the double tagging data into the basic integrated estimation framework (Appendices 14 and 15).

Archival Tagging

Archival tags are tags with electronic sensors and memory for storing the data from these sensors. These tags provide substantial information on the biology, habitat and location of a fish. In terms of the estimation framework developed here, the relevant data are those that allow the position of the fish to be determined. Some archival tags have the ability to detach themselves after a pre-set period, float to the surface and communicate their data via satellite. This has the advantage of providing movement and potentially natural mortality rate information that is completely fishery independent. However, their size currently limits their use to large fish. In addition, attachments for periods exceeding one year are problematical. The more commonly used archival tags need to be recaptured for the data stored on them to be retrieved. Within the context of the spatially-explicit models, archival tags provide additional data for estimating movements to the data provided by conventional tags since they provide information on the location of the fish during the intervening period between release and recapture. We developed a maximum likelihood approach for combining archival tag-recapture data (with its positional information) and tag-recapture data from conventional tags from multi-year experiments with spatially-explicit components.

We note that because archival tags exploit modern micro-electronics, they are very expensive relative to conventional tags. The number of archival tags that can be deployed will be small in practice. If archival tags were released separately from conventional tags, we would have to estimate fishing mortality and transition parameters specifically for the archival tags. This would render them practically useless in this regard. Thus, we assumed that the archival tags are released in parallel with conventional tags. In developing this approach, we also assumed that the positions of the archival tagged fish are known without error. While there can be large inaccuracies in the positional information obtained from archival tags, the spatial stratum in the spatially-explicit models considered here are large enough (i.e. 1000s of km) that the effect of location errors should be small.

Assuming that the probability of recapturing a tagged fish is independent of whether it has an archival or conventional tag means that the likelihoods for the two types of tags can be multiplied together. Note that initial tagging mortality may be higher for archival than conventional tags. If data were available to estimate this, it would be straightforward to incorporate such a differential into the estimation procedure. It should be further noted that in developing the likelihood for the archival data we assumed that reporting rates were similar for archival and conventional tags. In many situations this may not be the case, because higher rewards are generally provided for the return of archival tags. Again, modifying the likelihood component to allow for differential reporting rates would be straightforward if data were available for the differential to be estimated. The assumption of equal initial tagging mortality and

reporting rates for the two tag types was made as a pragmatic approach for addressing the general question of the potential for archival tags to enhance conventional tagging experiments. Incorporation of differential tagging mortality and reporting rates would only have complicated the analyses and presentation without changing the general results and conclusions.

Given the above two assumptions, the likelihood that an archival tag is not returned is exactly the same as that of a conventional tag released from the same region at the same time. However, for tags that are returned, the likelihoods differ for archival tags than conventional tags because of the additional information on location of the fish provided by archival tags. Essentially, the data from archival tags are like having multiple recaptures of the fish occurring at every intermediate time period between tagging and final capture. Explicit expressions for the likelihoods that take into account the multiple recapture nature of the archival tags are provided in Appendix 10.

Catch-at-age Estimation

The catch-at-age data has two main sources of errors: one that results from fishing being a random process (i.e., if we think of catching fish as a random experiment then we would not get the same outcome if we were to repeat the experiment); and a second that results from the age distribution of the catch being estimated from a sample (and generally a sample of lengths from which age must be inferred). We refer to the first source as process error and the second as sampling error. If sampling error did not exist, then the process error of the catch-at-age data could be modelled as multinomial if we assume each fish of a particular age has an equal probability of being caught, or, alternatively, as overdispersed multinomial if we assume there is overdispersion in the data (see previous discussion of overdispersion for the tag-return data). However, in most commercial fishery situations (including SBT), the sampling error will be large and we expect the process error to be negligible in comparison. As such, we chose to model all of the error in the catch-at-age as sampling error and ignore any process error. Both process and sampling error could be included but this would require a substantially more complex estimation approach with considerable computational overheads (e.g., a Kalman filter). As long as sampling error dominates the error associated with the catch-at-age data, there would be no gain from undertaking such an approach (see Appendices 5 and 7 for more detail).

We model the sampling error in the catch-at-age data as Gaussian with a coefficient of variation (CV) that depends on the level of sampling. The CV is intended to capture variability in the catch-at-age data due to non-homogeneous spatial and temporal distribution of fish as well as different size/age selectivities among vessels (i.e., if these factors are significant, then the CV of the catch-at-age data would be large because the age distributions derived from different samples could vary greatly). We note that the actual error structure in any particular application will depend upon the details of the sampling and age estimation procedures. The choice of a Gaussian distribution provides a general framework for exploring the effects of sampling error in the catch data on the mortality rate and abundance estimates. However, other error structures that accommodate a particular sampling regime can easily be incorporated as long as an appropriate likelihood function can be developed. We also note that we have assumed that the errors in the catch-at-age estimates between years (i.e., ages for

a single cohort) are independent. This is a reasonable assumption for many fisheries in which sampling and aging data are collected each year. However, in some situations (particularly where age at length is being estimated from a growth curve), covariance in the estimates between years may exist and should be accounted for.

In situations where observers are used to estimate reporting rates, the components of the fishery with observers will generally have additional catch sampling for estimating the catch-at-age. In such situations, the CV of the catch data will in part be determined by the level of observer coverage since this affects the level of sampling. The CV of the catch-at-age would be expected to decrease as the level of observer coverage increases. Although the exact relationship will depend on the situation, we developed a generic relationship to examine how changes in observer coverage would affect the estimation of mortality rates and abundance (see Appendix 7).

Returns from Only One Fishery Component

In addition to the basic multi-year tagging framework, we also considered what estimation framework could be used for tagging experiments in which tag-return data with reliable reporting rates and catch-at-age data are available from only one component of a multi-component fishery (Appendix 17). The approach developed was to consider the tagging and catch-at-age data in the context of a Petersen-type mark and recapture estimator of abundance (Seber 1982). A Petersen estimator is based on the ratio of the observed number of tags returned within samples taken from the population given the known number of tags released into the population. In a fishery context, the catch-at-age data constitute a sample from the population. However, unlike most situations in which a Petersen-type estimator is used, the size of the sample examined for tags is estimated rather than being known exactly.

Data

Extensive data exist from a large number of tagging experiments on SBT conducted over the last 40 years. Most of the tagging has been conducted on juvenile fish captured by surface gear in coastal waters off Australia. In this project, the primary data used in the analyses were from the SBT tagging experiments conducted in the 1990s and the associated catch-at-age data. We focused on these tagging experiments because they were designed with the intention of ensuring that multiple cohorts were each tagged at three different ages. In addition, information, while limited, exists for developing estimates of tag reporting rates from these experiments (Appendix 19). The lack of information about reporting rates for the pre-1990s tagging experiments is a major limitation which hinders the direct application of the estimation framework developed here. In the 1990s, tagging was done in every year between 1992 and 1997. In most years, tagging was done in both Western Australia (WA) and South Australia (SA), with most of the fish tagged being ages 1 to 3. Appendix 4 provides a description of the tag-recapture data from the 1990s experiments and the corresponding catch-at-age data, as well as details of how the data were compiled.

In addition to the 1990s tag-recapture data, we examined the potential of the data from the pre-1990s tagging experiments to provide quantitative mortality rate estimates within the context of the generalized estimation framework developed in this report (Appendix 16). Extensive tagging was conducted in most years in the 1960s. Tagging in the 1970s was limited both in the number of years of tagging and

the actual number of releases. In the 1980s, there were extensive releases but almost all of these occurred in 1983 and 1984. Pre-1980, tagging took place in WA, SA and New South Wales (NSW). In the 1980s tagging only took place in WA and SA since surface schools of juvenile fish had disappeared from NSW. For the analyses involving the pre-1990s tagging data, the compilation of data was the same as that for the 1990s data as described in Appendix 4 except that tag releases by fishermen and from troll caught releases were also considered[3].

Results

Estimation Framework

Extensive exploration was conducted on the performance and statistical properties of the integrated multi-year and catch-at-age estimation framework developed in this project (Appendices 5, 6, 7, 8, 9, 10, 11 and 12). This was done primarily through simulations mincing a range of potential experimental designs and parameter values for the underlying population dynamics. These simulations also provided indications of the consequences of different potential trade-offs in effort dedicated to the various experimental design components (e.g., numbers of releases, distribution of releases, levels of observer coverage, etc.).

Reporting Rates Known

In Appendix 5, simulations were conducted using the basic estimation framework under ideal conditions, namely known reporting rates and complete mixing, to demonstrate the value of directly incorporating catch-at-age data into the estimation framework and also to illustrate the trade-off between number of releases and accuracy of catch data in terms of parameter estimation. The results of the simulations demonstrate that there are substantial benefits in using the integrated catch-at-age and tagging model developed here compared to the more common tagging-only (Brownie) model for multi-year tagging of a cohort. The integrated model can provide maximum likelihood estimates not only of mortality rates (natural and fishing) but also of initial population size (at the time of first tagging) with reasonable levels of precision. Population size estimates are not obtainable if the catch-at-age data are not included. In addition, the simulation results suggest that including catch data can improve the precision of fishing mortality rate estimates by up to ~40% and natural mortality rate estimates by up to ~10%. They also show that reducing uncertainty in the catch data can lead to significant improvements in the precision of the population size estimate and the fishing mortality rate estimates (especially at young ages and with low numbers of releases), whereas increasing the number of releases can lead to large improvements in the natural mortality rate estimates as well as the fishing mortality rate estimates (especially when the uncertainty in the catch data is high).

---

[3] These additional tags were included because of the large number of fishermen/troll releases. However, the examination in Appendix 20 is largely concerned with differential in the spatial distribution of releases and returns and also includes detailed consideration of differences in returns from scientest-tagged and fishermen-tagged fish.

Reporting Rates Estimated

We explored the properties of the basic estimation framework for situations where reporting rates are unknown (see Appendix 7). We focused on the situation in which observer data are used to estimate the unknown reporting rates because observers are the only currently viable approach for estimating reporting rates from tuna longliners (Appendix 18). Tagging and observer programs can be costly to run and resources available for doing so are usually limited. Thus, in the simulations, we explored the effect of releasing more tags versus increasing observer coverage on the precision and bias of the parameter estimates. In particular, we were interested in the estimates of natural mortality, fishing mortality and abundance – reporting rate estimates are also provided by the model but these are usually of secondary interest. We found that increasing the number of releases improved the precision of all parameters, especially the fishing mortality estimates at older ages and the natural mortality estimates. The improvement tended to be greatest at low release numbers such that increasing the number of releases beyond 1000 resulted in only marginal gains for most parameters (natural mortality rates being an exception). Increasing the level of observer coverage also improved the precision of all parameters, especially the fishing mortality estimates and the population size estimate. Although the improvement per unit increase in observer coverage was greatest at low levels (i.e., in going from proportion coverage of 0.05 to 0.1), substantial gains in precision were still achieved in all parameters (except reporting rates) by increasing observer coverage from 0.2 to 0.5. The results depended on the reporting rate value assumed. The precision of the parameter estimates was greater when the reporting rate was high; however, the above general observations held true regardless.

In some situations considered, biases were evident in some of the parameter estimates, however these diminished as the number of releases increased and as the level of observer coverage increased. In any case, biases were not generally an issue of concern except when both the number of releases was low (500 or less) and the levels of observer coverage was low (0.1 or less). Even then, when the reporting rate was high, the biases were never more than 12% for any of the parameter estimates. An unexpected observation regarding biases was that some of the parameter estimates (the age 1 and 2 fishing mortality estimates and the population size estimate) showed almost no signs of biases in the simulations with a low reporting rate (0.25) but did show small biases in the simulations with a high reporting rate (0.75); the reason for this is unclear[4].

We also explored the properties of the basic estimation framework in a two-fishery situation where reporting rates are unknown and reporting rates are estimated by two different methods (see Appendix 9). The simulations in this case were conducted within the general context of the current SBT fishery. Thus, one fishery was considered to be a longline fishery in which reporting rates are estimated from observer data and the second fishery was a surface fishery in which reporting rates are estimated from tag seeding experiments. In this investigation, we also allowed for overdispersion in the tag-return data. The results from the model with overdispersion

---

[4] Note that with high reporting rates and low observer coverage some bias may be introduced because by chance the proportion of tags returned from the observed catches may be less than those returned from the unobserved catches. In such a case, the reporting rate estimate would be truncated to equal 1.0, and this may be a source of estimation bias.

suggest that in order to achieve coefficients of variation of 20% or less for the longline fishing mortality rates at ages 1 to 3, observer coverage must be at least 30% (and at least 20% for the model without overdispersion). Estimates of fishing mortality in the surface fishery are chiefly unaffected by the level of observer coverage in the longline fishery, provided fairly accurate estimates of surface fishery reporting rates and catch-at-age by fishery exist. It is important to note, however, that these results depend on the assumption of no systematic incomplete mixing. If this assumption is violated, then the level of observer coverage in the longline fishery would become more influential on the accuracy and precision of parameter estimates. Without good observer data, and thus good information on differential tag reporting and return rates between fishery components, there would be little power to test the assumption of non-mixing and, if necessary, develop spatially-explicit tag recovery models to account for heterogeneity in recapture probabilities. The results also demonstrate the importance of having reliable and precise estimates of the catch-at-age for each fishery when applying the estimation model presented here. This emphasizes the need to develop appropriate sampling and error models for the catch data; having representative and adequate observer coverage can help to accomplish this in longline fisheries.

Overdispersion

We explored the properties of the basic estimation framework when the recapture data were modelled as Dirichlet-multinomial to account for overdispersion in the recapture probabilities (Appendix 9). Qualitatively, the results from varying any of the factors were very similar in the model with Dirichlet-multinomial returns as the model with multinomial tag returns. As expected, the parameters were almost always estimated with less precision (i.e., their CVs were larger) with Dirichlet-multinomial returns, but the relative changes in precision and general observations about the trade-offs in design components did not change significantly between the models. However, our simulations suggest that the extent of overdispersion is consistently underestimated. This means that if overdispersion exists the estimated precision of the parameter estimates using a Dirichlet-multinomial model is still likely to be overestimated (i.e., variances underestimated).

Spatially-Explicit Models

We explored the properties of the basic estimation framework when more than a single spatial region is included in the model (Appendices 10, 11, 12 and 13). With more than a single region, identifiability problems (i.e., some parameters are not estimable) can exist if tagging does not occur in every region and time period. The extent of this will depend upon the spatial structure in the model and any constraints placed upon the parameters, particularly those that determine the movement among areas (see Appendices 12 and 13). Due to the large number of parameters as the spatial complexity of the model increases, it is difficult to make general conclusions about the estimation properties with complex spatial structures because any conclusions will depend on the range and details of the simulations conducted. However, the results of the simulations that we conducted suggest that unbiased parameter estimates can be obtained with reasonable levels of precision as long as 'reasonable' numbers of tags are released in most, if not all areas, and if 'reasonable' data are available for estimating reporting rates and the catch-at-age. As the number of tags released in some areas becomes small and/or the quality of data for estimating

the catch-at-age parameters diminishes, biases can be introduced (Appendix 12). Overdispersion, particularly when high, can degrade performance substantially compared to simulations with simple multinomial recaptures and can introduce biases into some parameter estimates (Appendix 10). In general, estimates of natural mortality (particularly for ages 2 and above combined) can be problematical with small numbers of releases or poor data for estimating reporting rates and catch-at-age (Appendix 12). Simulations using low fishing mortality rates (and therefore low numbers of returns) were found to yield biases in some of the parameter estimates. The biases appear to decrease or disappear with increased numbers of releases (and therefore increased numbers of returns), which indicates that they are primarily small sample size effects.

Population wide estimates were generally estimated better than the corresponding area-specific ones (Appendices 10, 11 and 12). In fact, unbiased and relatively precise population wide estimates were often obtained even when the data were insufficient to yield unbiased or reasonably precise area-specific estimates. In contrast, estimates based on pooling data across areas and ignoring the spatial structure showed some substantial biases when large differences existed in recapture probabilities among areas. The degree of bias is roughly related to the extent of non-mixing. Overall, the simulation results indicate that if spatial heterogeneity is likely to exist it is important to include a spatial dimension within the analysis to achieve unbiased estimates of population wide parameters. In terms of the precision of the estimates, there appears to be little or no cost for including a spatial dimension even when complete mixing does in fact exist (Appendix 12).

The simulations also indicate that the distribution of tag releases is not critical in terms of population wide estimates. In fact, for some spatial structures and movement dynamics, it is not essential to have releases in all areas to get meaningful estimates at the population wide scale. In Appendix 12, unbiased and reasonably efficient estimates of population wide parameters were obtained when tagging only took place in one region in an 'SBT-like' situation. Moreover, tagging in only one area yielded no real loss in precision as long as the same total number of tags was released. However, we note that with more complex spatial structures and movement dynamics, which might be expected in many situations, releasing tags in only one area can result in an intractable estimation problem with some parameters becoming unidentifiable (Appendix 13).

Estimates for SBT from Conventional Tagging Experiments

1990s Experiments – Reporting Rate Estimation

An essential component for applying the estimation framework developed in this report to the data from historical SBT tagging experiments is the estimation of reporting rates. Estimating reporting rates for SBT is complicated because of the nature of the SBT fishery, which comprises multiple components with varying reporting rates. For the 1990s tagging experiments, there are only limited data for estimating reporting rates and these data do not permit a rigorous statistical analysis. Nevertheless, if assumptions are made about the relationship between reporting rates in fishery components with and without data, then the data do provide sufficient information to derive estimates of the reporting rates. In previous analyses of the

1990s tagging experiments (Polacheck et al. 1996, 1998; Preece et al. 2001), a range of reporting rate estimates has been used that correspond to different options for the relationship between the reporting rates in components of the Japanese fishery with and without observers and for the reporting rates in the Australian surface fishery. The most recent reporting rate estimates were produced in 2001; since then there have been revisions to the catch-at-age data used in estimating the reporting rates as well as new data from tag seeding experiments. As such, we produced updated reporting rates for use in this report that reflect these revised catch data and also incorporate the tag seeding data.

Appendix 19 documents how the updated reporting rates were calculated and discusses the differences among the various (eight) options presented. As discussed in Appendix 19, only option 8 among the eight options is actually information based for all major non-observed fisheries with non-zero reporting rates. For the other seven options, the reporting rate for at least one of the fishery components is based on what can be considered the most optimistic assumption for that fishery. In this sense, option 8 could be considered the most 'realistic' or plausible. As such, the reporting rate estimates corresponding to option 8 (given in Table 1) were used as the primary reporting rate estimates in the analyses conducted for this report.

Table 1: Year- and age-specific reporting rate estimates used in the primary analyses conducted for this report (taken from Table 5a, option 8, of Appendix 19).

Age	1991	1992	1993	1994	1995	1996	1997
1	0.933	0.887	0.926	0.522	0.748	0.32	0.805
2	0.654	0.577	0.75	0.498	0.632	0.387	0.775
3	0.597	0.543	0.625	0.6	0.632	0.475	0.727
4	0.196	0.327	0.471	0.559	0.397	0.413	0.639
5	0.198	0.276	0.402	0.457	0.258	0.25	0.597
6	0.171	0.278	0.381	0.39	0.192	0.267	0.537
7	0.18	0.277	0.411	0.405	0.179	0.275	0.568
8	0.186	0.292	0.423	0.418	0.187	0.254	0.596

1990s Experiments – Tag Shedding Estimation

Tag shedding is another important component that needs to be accounted for in the analysis of tag-return data. Estimates of tag shedding were not included in previous analyses of mortality rates from the 1990s tagging experiments (e.g., Polacheck et al. 1996, 1998); this reflected both a lack of resources for undertaking the analyses combined with a perception that tag shedding rates were relatively low. Incorrectly assuming that there is no tag shedding will result in biased estimates of mortality rates and abundance, although the biases will be relatively small if shedding rates are very low. Nevertheless, it is important that tag shedding be included within comprehensive analyses of tag-return data both to confirm that the rates are in fact low and to remove any potential biases.

Estimating shedding rates and incorporating them into the analysis to estimate mortality rates and/or abundance can be statistically and numerically complex because of the need to account for differences in tag shedding rates among individual taggers (Appendix 14). Estimates of tag shedding rates for the 1990s tagging experiments were calculated for individual taggers and six groups of taggers identified as having statistically similar shedding rates (see Appendix 14 for more detail). The estimates of the shedding rates were generally small, particularly in light of the fact that all tagged fish were double tagged. Thus, the probability of a fish losing both tags over the primary period of recovery (i.e., over four years of liberty) was generally less than 10% (Figure 1). However, the shedding rate for tagger group 6 was substantially greater than for the other taggers (e.g., over 20% of the fish tagged by group 6 are estimated to have lost both tags after four years at liberty). The differences among taggers emphasises the importance of allowing for differences among taggers in the estimation of mortality rates and population sizes rather than using an average shedding rate calculated from pooling the data across all taggers. Although the shedding rates were generally small, excluding them from the model for estimating mortality rates and abundance (i.e., assuming 100% tag retention) did have an effect on the parameter estimates. For example, estimates of initial population sizes were 2-6% greater if 100% tag retention was assumed and the estimates of natural mortality were also sensitive.
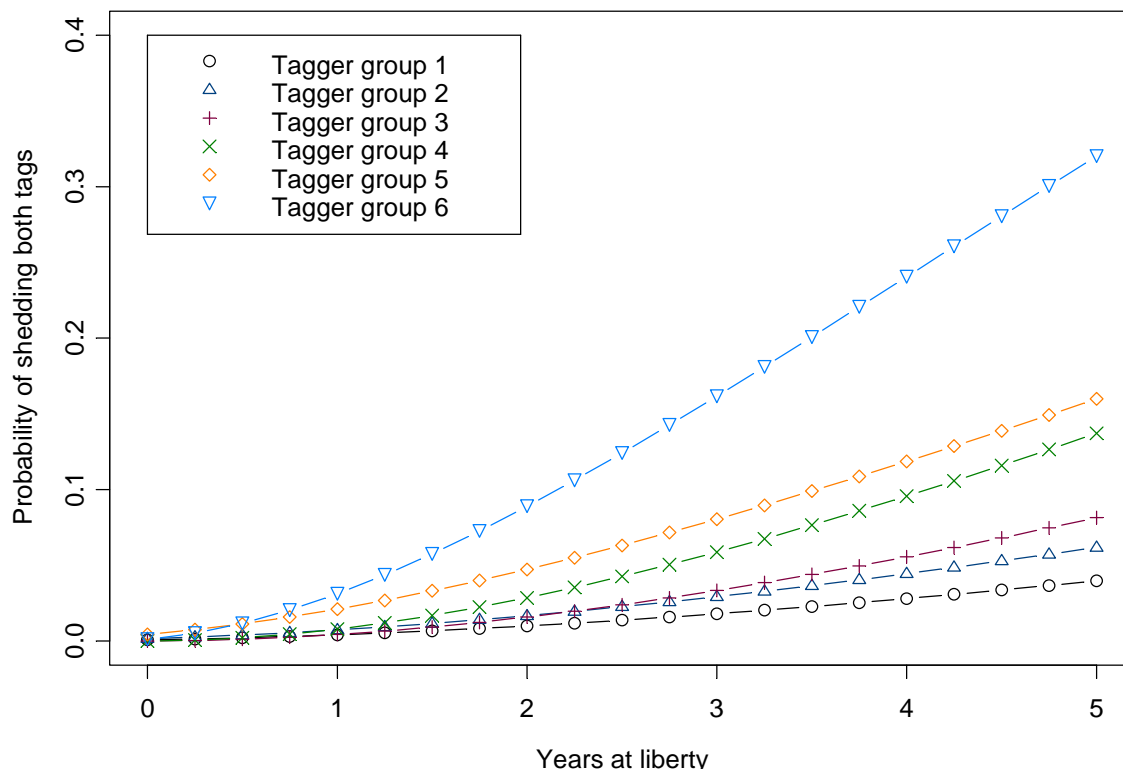
Figure 1: Estimates of the probability of a tagged fish losing both its tags as a function of its time at liberty for the 6 tagger groups defined in Appendix 14 (see Appendix 14 for more details).

1990s Experiments – Integrated Model with No Spatial Structure

A comprehensive analysis for estimating mortality rates and abundance for southern bluefin tuna using tag-return and catch-at-age data from the 1990s was conducted with the assumption of no systematic non-mixing after the first year of release (i.e., no need for spatial structure) (Appendix 15). The analysis included two parameterizations for fishing mortality rates – one in which the age- and year- specific fishing mortality rates were unconstrained (model 1) and one in which they were constrained to have separable age and year effects (model 2). Both models led to similar parameter estimates and the same general conclusions. The results suggest that natural mortality at age 1 is quite high (~0.4) and decreases to about 0.2 by age 5; however, the uncertainty in the age 5 estimate is very high and the estimate is sensitive to changes in either the model or the data inputs. Tagging cohorts at age 4 (in sufficient numbers) in addition to ages 1 to 3 could provide valuable information for better estimating natural mortality at older ages. Fishing mortality estimates were generally close to zero for ages 1 and 2, were greatest at ages 3 and 4, and declined at age 5 (Figure 2). The results also suggest that fishing mortality decreased in the first couple of years of the 1990s then increased fairly steadily from 1994 to 1997.

Population abundance appears to have decreased from about 2.5-3 million age 1 fish in 1989 to just over 1 million age 1 fish in 1993 and 1994 (Figure 3).

In the analysis, we have attempted to account for uncertainty in all of the data components in the model. However, there is very limited information for estimating the uncertainty associated with both the reporting rates and the catch-at-age data. Accounting for the potential uncertainty associated with these required relatively arbitrary assumptions about effective samples sizes and expected levels of precision. If the assumed values are too low, then the variance of the mortality rate and abundance estimates would be underestimated. In addition, we assumed a multinomial distribution for the tag return data, but the variance in the number of returns is likely to be greater than predicted by a multinomial distribution due to incomplete mixing and heterogeneity in the capture probabilities of fish. The Dirichlet-multinomial, as described in Appendix 9, would be one approach for dealing with this; however, this approach requires an assumption about the level of overdispersion, either assuming it is known or keeping it constant since it cannot be estimated within the model otherwise. Moreover, the rather arbitrary assumptions about the uncertainty associated with the reporting rates and catch-at-age data renders questionable the value to be gained from more sophisticated variance estimation. The necessity for and potential gain from incorporating overdispersion in the tag-return model for southern bluefin tuna is an area for further investigation, particularly for future tagging experiments.
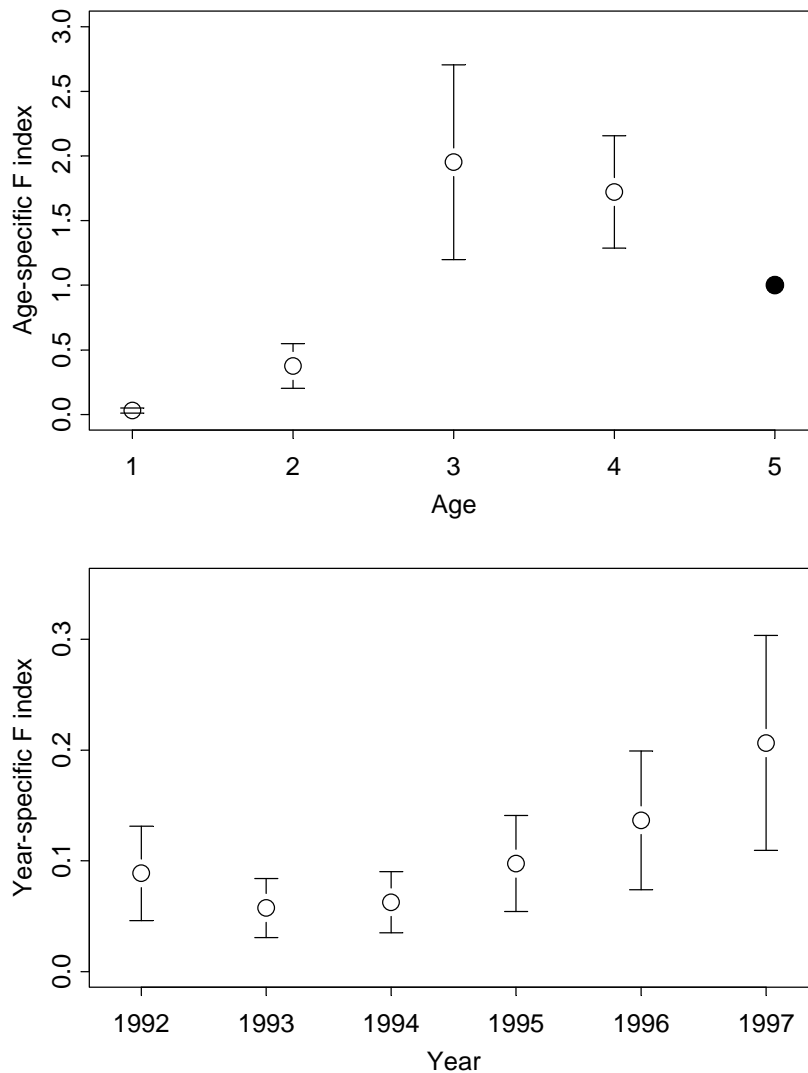
Figure 2: Estimates (± 1 standard deviation) of the age-specific fishing mortality rate effect (upper panel) and the year-specific fishing mortality effect (lower panel) for the model with separable fishing mortality rates (model 2). Note that the latter estimates should be interpreted as relative indices, where the age effect at age 5 has been fixed at 1.0 (see Appendix 15). The estimate of the year-specific component for 1991 has been omitted because it is based on very little data and has large uncertainty associated with it.

Figure 3: Comparison of population size (P) at age 1 estimates and their standard deviations (SD) by cohort for the model with unconstrained fishing mortality rates (model 1) and the model with fishing mortality rates constrained to have separable age and year effects (model 2). Black circle = model 1 results; blue triangle = model 2 results. For the 1989 cohort, only a direct estimate of P at age 2 is obtained from the models, so the age 1 estimates shown are post-calculated (see Appendix 15 for more details).

1990s Experiments – Integrated Model with Spatial Structure

The integrated spatially-structured model developed in Appendix 11 was applied to the multi-year tag-recapture and catch-at-age data from the 1990 to 1994 cohorts (see Appendix 16). It should be noted that the 1990s experiments had not been designed in the context of such a spatially-explicit model and almost all tag releases occurred off WA or in the GAB. As demonstrated in Appendix 16, the parameters are still identifiable, however the power to estimate the spatially-explicit parameters would be expected to be low, particularly given the low reporting rates in the longline fishery. This was in fact the case and the spatially-explicit parameters were not well constrained by the data. Improved performance was obtained by placing some 'sensible' restrictions on the transition probabilities between regions and the distribution of age 1 fish by region.

The results provided from fitting the spatially-structured integrated model were generally consistent with our basic understanding of the spatial dynamics of the SBT fishery and juvenile stock. Thus, selectivity in the GAB/WA surface fishery was estimated to be dome-shaped, peaking at age 3 and declining to close to zero at ages 1 and 5. Similarly, selectivity in the South-East Indian Ocean (SEIO) longline fishery

(which includes most of the Taiwanese longline catches) was also estimated to be dome-shaped with a peak at age 3; however, it was lower at age 2 and higher at ages 4 and 5 than in the GAB. In the South African and Tasman Sea regions, fishing mortality rates were estimated to be greatest at ages 4 and 5. The estimates of age-specific components of the regional and seasonal transition probabilities suggest that the majority (75-90%) of fish move from the GAB to the SEIO at the end of season 1 at all ages, with the exception of age 4 where a large fraction (about 40%) also move to the Tasman region. The transition probability estimates back to the GAB at the end of season 2 suggest that almost no fish return from South Africa, whereas large fractions return from the SEIO and Tasman regions. The age component of these transition estimates suggests that a greater proportion of fish return to the GAB as age increases. It is difficult to know if the latter finding is real or simply an artefact of the limited data (and confounded by the estimates of the initial distribution at age 1 for which there is little information). It is also probably due in part to the lack of return data from the longline regions and at older ages. Estimation and interpretation of the results is confounded by the estimates of reporting rates as well as potential problems with the age distribution of the catch data. Temporally-changing sampling biases and discarding of small fish have been identified as issues in the SBT fisheries during the time covered by this study, and could be distorting the spatially-explicit parameter estimates.

Comparison of the results from fitting the spatial and non-spatial models to the 1990s tagging and catch-at-age data indicate similar and consistent estimates at the population wide level, both for annual fishing mortality rates and for initial (age 1) cohort size estimates when similar hypotheses/constraints are used for the natural mortality rates (e.g., that M is constant past age 2). In both cases, estimates of fishing mortality rates were highest for ages 3 and 4 and increased during the 1990s (compare Table 2 and Figure 2). Also, the overall rates were similar in magnitude. The estimates and trends in the abundance at age 1 by cohort were also similar (Figure 4), as were the actual estimates of natural mortality (i.e., 0.42 compared to 0.46 for age 1 and 0.40 compared to 0.37 for ages 2+ for the spatial structured and non-structured models respectively). The similarity and consistency between the two sets of results suggest a high degree of robustness in the results and that unsystematic non-mixing issues may not be a substantial concern in these experiments.

Table 2: Estimates of the average yearly fishing mortalities by year and age for SBT when the tag release/recapture and catch at data were fit to the integrated spatially-structured estimation model.

Year	Age 1	2	3	4	5
1991	0.041				
1992	0.005	0.037			
1993	0.002	0.018	0.100		
1994	0.003	0.017	0.096	0.100	
1995	0.006	0.039	0.194	0.185	0.107
1996		0.071	0.344	0.270	0.129
1997			0.454	0.445	0.195
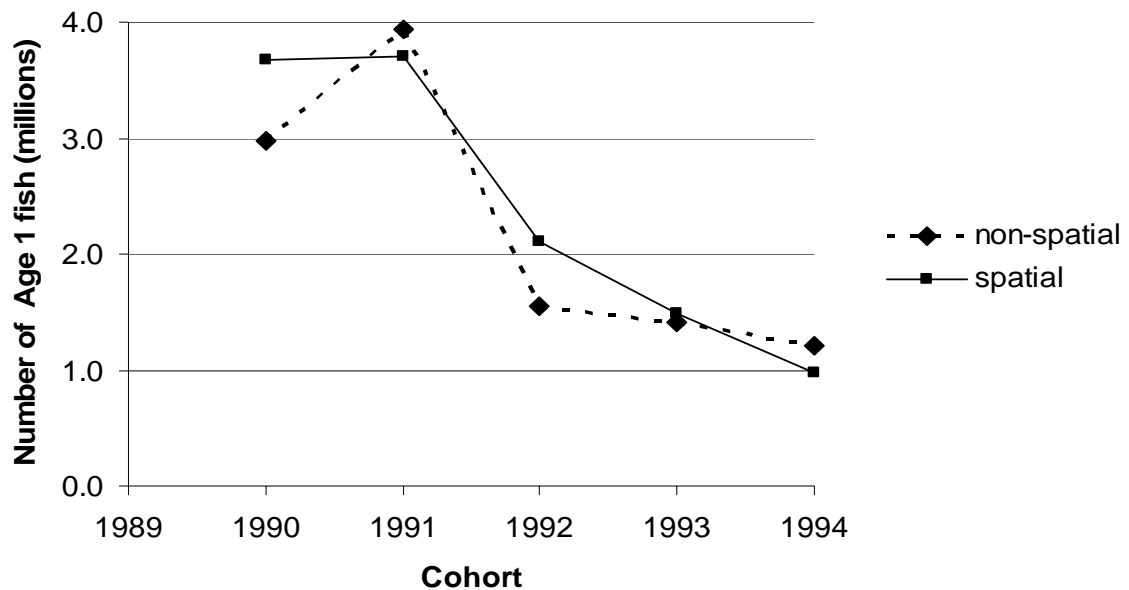
Figure 4: Estimate of the population size (*P*) at age 1 (millions) for the 1990 to 1994 cohorts from the integrated model with spatial structure compared to the estimates for the non-spatially structured model when similar assumptions are made for changes in natural mortality with age (see Appendix 16 for details).

1990s Experiments – Using Only Returns from the Great Australian Bight

A Petersen-type mark and recapture estimator of abundance was used to examine the potential information that can be derived from the 1990s tagging experiments if only returns from the Great Australian Bight are considered (Appendix 17). The purpose was to determine the type of performance that might be expected from the current CCSBT tagging program, given that reliable reporting rates are unlikely to be available for the longline component of the fishery. The results suggest that using only the SBT tag returns from the surface component of the entire juvenile fishery may still allow for information on juvenile abundances and/or trends, but this requires relatively consistent mixing patterns of tagged fish with the complete population of juvenile fish. It also requires that reliable estimates of reporting rates and of the age distribution of the surface catches are available, which emphasizes the need for developing appropriate statistical estimators for these quantities.

The results from applying the estimator to the 1990s SBT tagging data lend reasonable support to the assumption of consistent and high levels of mixing. They also indicate no increase (and possibly a decrease) in the strength of cohorts at age 1 during the first half of the 1990s, and suggest a declining trend in abundance by ages 2 and 3 for the surviving members of these cohorts (Figure 5). One advantage of using this approach when there is little or no information on the reporting rates from the longline fisheries is that it can provide an indicator of trends in juvenile abundances over the period of the tagging experiments independent of any assumptions about tag

24

returns and catches in the longline fishery. Such indicators can provide a useful independent check on overall complex stock assessment results.
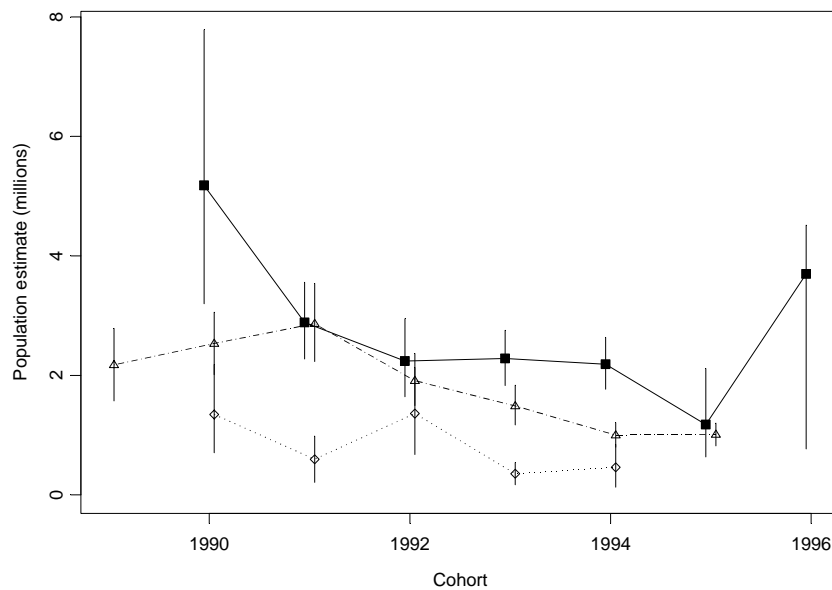
Figure 5: Comparison of Petersen-type estimates for the number of 1-, 2- and 3-year-old SBT ($P^*_{A,a,c,r}$ as defined in Appendix 17) based on age 1 releases (solid squares), age 2 releases (triangles) and age 3 releases (diamonds), respectively, from Western Australia and the Great Australian Bight combined (note there were no age 3 releases in Western Australia). The estimates shown are based on the pooled returns and catches for each age of release. Error bars are estimated approximate 90% confidence intervals (see Appendix 17 for more detail).

Pre-1990s Experiments

The historic (pre-1990s) SBT tagging data were examined with respect to their potential for providing quantitative mortality rate and abundance estimates (Appendix 20). The examination conducted indicated that there are substantial problems with using the pre-1990s data in this regard, and we concluded that the data are not directly amenable to the integrated estimation framework developed within this report. For the 1970s and 1980s releases, there was simply not enough multiple tagging of the same cohorts at consecutive ages to permit the application of a Brownie-type estimator. For the 1960s releases, there was sufficient multiple tagging of some cohorts to possibly allow for a Brownie-type estimator to be used; however, the differential return rates for tags released in different areas provides strong evidence for substantial incomplete mixing among releases from NSW with those from WA and SA (e.g., Table 3). This indicates that the non-spatial models of Appendices 7 and 9 would not be appropriate and that a spatially-explicit model would be required to obtain reliable estimates. The spatially-explicit models developed for use with the 1990s tagging experiments (Appendices 10 and 11) cannot be used for the 1960s data because these models make no allowance for a NSW juvenile component of the stock since this component disappeared in the early 1980s (Caton 1991). The generic spatial model of Appendix

10 is also not directly applicable because of the lack of releases and fishing in all areas (particularly in the longline areas) and the differing seasonality among the fisheries. A spatially-explicit model that incorporates hypotheses with respect to the possible movement dynamics along the lines of those in Appendix 11 but tailored for the stock and fisheries of the 1960s could be developed. The development of such a model was beyond the scope of the current study, particularly given the additional complexities of the 1960s data including lack of data for estimating reporting rates, high shedding rates and concerns associated with the large number of fishermen releases (e.g., the results would be highly dependent upon assumptions about reporting rates and whether or not fishermen releases were included) – see Appendix 20 for more detail. Nevertheless, further analyses of the 1960s tagging experiments could be informative, especially with respect to the NSW surface component of the stock and its disappearance in the early 1980s. They would also provide additional estimates of juvenile natural mortality that would be useful for comparison with those from the 1990s experiments.

Even without a quantitative estimation model, the examination of the tag-return and catch data from the 1960s in Appendix 20 suggests that historically a large degree of spatial structure and spatial heterogeneity existed among juvenile SBT found within Australian waters. The combination of the high proportion of NSW returns from releases in NSW combined with the high return rates from releases in NSW (but not from other areas) associated with increased catches in NSW in the late 1960s (see Appendix 20) suggests the possibility of a substantial degree of separation among juvenile SBT between NSW and other areas (WA and SA). This further suggests a high degree of vulnerability to over-exploitation and localized depletion for the NSW component, which would have implications for the rebuilding of the SBT stock.

Table 3: The number of surface tag returns by recapture location (SA, NSW, TAS) and the number of longline tag returns for tags released in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania.)

Release Location	Recapture Location			
	SA	NSW	TAS	Longline
WA	278	166	1	128
SA	179	128	0	144
NSW	47	1622	1	77
TAS	3	1	0	3

Designing Tagging Experiments

The evaluation of experimental design for a tagging experiment can be considered from two perspectives:

1. Will the experimental design provide robust and sufficiently reliable estimates (or what level of resources is required to achieve acceptable performance)?

2. What is the optimal allocation of resources among the competing data needs (e.g., number of releases, data for reporting rate estimates, etc.)?

There is clearly strong interaction between these two considerations as the level of resources required will depend in part upon their allocation among competing data needs. Determining what level of resources is needed and the 'optimal' experimental design for a given amount of resources is complex because of the large number of design parameters and the difficulty in determining realistic costs for achieving specific values for them, particularly in international multi-fishery situations (e.g., observer costs). Within the estimation framework developed here, there are five basic components that potentially compete for the resources and that will determine whether the experiment is likely to achieve acceptable performance. These are:

1. the number of tag releases distributed appropriately over time and space;
2. the number of tags returns (i.e., resources devoted to rewards, promotion and recovery activities);
3. the collection of data for estimating reporting rates;
4. the collection of data for estimating the catch-at-age data;
5. the collection of data for estimating tag shedding rates.

In addition, it is essential that well-developed procedures are available for catching, tagging and releasing fish that have minimal impact on the survival of tagged fish. Where possible effort should be taken to estimate any tagging related mortality and incorporate such estimates if not essentially zero into the overall estimates of mortality rates. In any case, the condition of all fish that are released should be assessed and recorded (e.g. whether bleeding or not). Only fish deemed to be in 'good' condition at the time of release (where 'good' is clearly defined in observable terms as part of the tagging protocol) should be included in the analyses.

It should be emphasized that if data from any of these components are not available then valid estimation is not possible. Thus, the first step in the design evaluation stage is to determine whether it is feasible to collect the data for each of these components. If it is not possible, serious consideration should be given to whether to proceed because any results from the experiments will be dependent upon assumptions for the missing components[5]. For example, there are numerous instances in which tagging experiments have been conducted without having built the collection of data for estimating reporting rates into the design (e.g., the SBT experiments conducted in the 1960-1989 period), and the potential value of these experiments with respect to estimating population parameters (e.g., mortality rates and abundance) have been compromised or limited. In this regard, it is important that data are not only collected for estimating all necessary components but also that they are collected using appropriate and documented sampling designs so that the variances associated with the estimates can be estimated. Otherwise, there is no statistical basis for evaluating the overall accuracy and precision of the resulting mortality rate and abundance

---

[5] In some situations the results can still provide useful information and it may be worth proceeding. For example, some parameter estimates may be relatively insensitive over the range of reasonable assumptions for the missing data components. In particular, if actual fishing mortality and tag return rates are high, then the returns themselves provide the basis for estimating a minimal reporting rate while the estimates for fishing mortality over the remaining range of possible reporting rates could be informative.

estimates. In this regard, estimates of catch-at-age data are often problematical in that, although they often exist, the way they have been compiled often does not allow for proper evaluation of their accuracy and precision.

Among these five components, design considerations for collecting tag shedding data are the most straightforward to deal with. This is because the data for estimating shedding rates comes from fish that are double-tagged. Thus, the maximum resources that can be devoted to improving estimates of tag shedding will be determined by the number of releases. In many cases, the marginal resources required to double tag all releases (i.e., the cost of the tags and the extra rewards[6]) will be small relative to the overall resources required for implementing a tagging experiment. As such, a robust approach is simply to double tag all fish and thus simplify this aspect of the experimental design[7]. In cases where the marginal cost of double tagging all releases is substantial, then it is important to ensure that a sufficient number of double-tagged fish are released by each tagger to get sufficient returns with which to reliable estimates of their shedding rates (see Appendix 14). The number of releases that is sufficient will depend in part on the recapture rates and the reporting rates since these determine how many of the releases will be returned. In the results for SBT presented in Appendix 14, the shedding estimate for the tagger with the least number of returns (39) had a CV of 30%; however, the results also suggest that the correlation between number of returns and precision of shedding estimates can be quite weak. Even if shedding rates can be estimated accurately, high levels of shedding can compromise the entire experimental design (e.g., if shedding rates are sufficiently high such that the probability of still retaining a tag after two years is low, the information from the multi-year tagging of the same cohort over three years would lost). As such, more important than determining resources to devote to double tagging in the design phase is to ensure that sufficient resources are devoted to developing robust tagging techniques and adequately training taggers so that tag shedding rates are consistently low. This development and training should be a prerequisite in the experimental design.

Among the other four design components, resource requirements are more difficult to specify because the expenses involved with each of them can be high and there can be substantial interaction/trade-offs among them[8]. In addition, the resources available for some components may be determined by needs outside of the tagging experiment and not amenable to being traded-off against other components. We would emphasize, however, that all four of these components are essential and if the data for any component is insufficient then the overall estimates from the tagging program will be

---

[6] Note that if tag shedding is relatively low, most tag returns will involve the return of two tags and thus it would be the combined reward for both tags that would be the effective incentive. Thus, double tagging all fish should not necessarily be seen as doubling the cost of rewards.

[7] It is important to ensure that in fact both tags are reported when a fish is recaptured with two tags. Thus, the rewards should be based on the number of tags returned (and not the number of fish from which tags are recaptured). This needs to be clear in promotional material and rewards should be such that there is sufficient incentive to return more than one tag. Also, it is important to have direct communication, if possible, with tag returners to ensure that both tags are being returned.

[8] It should be noted that in the case where observers can observe each individual fish caught and the resulting data are used to estimate reporting rates, increasing resources devoted to observers would have the synergetic effect of improving both the actual reporting rates and the catch-at-age estimates, and 100% observer coverage could eliminate the need for other forms of tag recovery activities and catch sampling.

compromised. As such, it is important to ensure that there is a commitment for sufficient resources for each component to achieve what would be considered minimal levels of reliability. A commitment in this case is not only simply for financial resources but must also involve access/availability, such as access to a representative sample of vessels in the case of observers or availability of representative landings for tag seeding experiments.

It should be noted that there are interactions between the expected results from a specific design for a tagging experiment and the underlying dynamics of the population and fishery involved. Thus, fundamental design parameters will be affected by the underlying dynamics. For example, the number of releases that may be adequate to achieve what is considered to be an acceptable level of precision in a situation with high fishing mortality rates can be inadequate if fishing mortality rates are low (e.g., see Figure 4 in Appendix 5). In addition, as the basic (ideal) estimation framework is expanded to incorporate all factors that need to be accounted for in any specific implementation, the number of design parameters increases and the evaluation of alternative experimental designs becomes increasingly complex (the number of tags to release at each age, location and timing of releases, the level of observer coverage in different fisheries, the number of tags seeded in each year, the effort devoted to estimating the catch-at-age data, etc.).

The extensive simulations that we conducted under this project (see Appendices 5, 7, 9, 10, 11 and 12) demonstrate that the trade-off among these various design components can be quite complex with substantial interactions among the various parameters (e.g., the trade-off between number of releases and observer coverage can depend upon the precision of the catch-at-age estimates). As such, it is not meaningful or fruitful to attempt to provide specific guidance on values to be used for the design parameters or on the trade-off among them without knowing specific conditions of the tagging experiment to be implemented. As such, for any given tagging experiment, simulation testing of potential experimental designs conditioned on the prevailing situation is the only feasible approach for determining the appropriate design parameters and evaluating the trade-offs among them. Simulation testing is important to ensure that robust estimates can be achieved. This is particularly the case if complex spatial/temporal structure needs to be taken into account in subsequent analyses (see Appendix 12). The estimation framework and simulation approaches developed in this project provide a general basis for conducting such simulations. While it is not possible to provide detailed generic design criteria for tagging experiments, the simulation testing that we conducted in this project provides some insights into the basic requirements and trade-offs, both in general and in 'SBT-like' fishery situations.

Number of Releases

To a rough first approximation, the variance for many of the parameters of interest will be proportional to the number of tags returned. As such, ideally it is important to ensure that sufficient tags are released so that that there are a 'reasonable' number of expected returns from each of the release/return strata in the model. While in general, the mortality rates and population size estimates should be asymptotically unbiased[9],

---

[9] Because of the way that we have modelled the variance of the catches in the simulations so that they have a constant CV, the estimates of population sizes in the simulations have a small statistical bias.

this is not necessarily the case with small sample sizes. Some substantial biases were evident in both the non-spatial (Appendices 5 and 7) and spatial (Appendix 12) simulation results when the number of releases was small. It is not possible to provide generic guidelines on what constitutes sufficient numbers of releases both to avoid substantial biases and to achieve reasonable levels of precision in parameter estimates because what is sufficient will be highly dependent upon the population/fishery dynamics and the reporting and shedding rates. Nevertheless, simulations conducted in Appendix 12 for an SBT-like situation with spatial structure (i.e., one in which there is seasonal migration between discrete areas) to investigate the effect of differing sample sizes provides some indication of what level of releases may be required at least within an SBT-like context. The results suggest that with high quality data for estimating the catch-at-age and reporting rates, releases on the order of 1000 tags per year would be sufficient to obtain estimates of the initial population size and fishing mortality rates with CVs of ~20% or less. In contrast, for with low data quality, not even 2400 releases per year was sufficient to achieve similar levels of precision nor were the biases always small. Estimates of natural mortality rates are more poorly estimated and to achieve CVs of less than 40% would entail releases on the order of 2400 per year for the high data quality situations and substantially greater for the low data quality situations.

Overdispersion

Overdispersion in tag-return data (relative to a basic multinomial model) is likely to be present for most tagging experiments as a result of incomplete, but unsystematic, non-mixing at relatively small to medium scales in relationship to the timing and distribution of tags (e.g., schooling). Thus, in trying to assess whether an experimental design will provide robust and sufficiently reliable estimates, it is critical to allow for overdispersion. Results in Appendices 9 and 10 provide results on how overdispersion in the tag-return data can degrade the precision of and, in more extreme cases, induce biases into parameter estimates. In Appendix 9, we assumed a factor of three for the increase in the variance of the returns at age relative to a multinominal distribution to provide an indication of what level of precision might be required for SBT tagging experiments. However, the value of three was rather arbitrary. In most cases there would be little basis for anticipating in the design phase what level of overdispersion would be realistic to expect. The general issue of how best to quantify overdispersion in tagging experiments needs further research. Nevertheless, it is important to allow for overdispersion in the design phase and to allow for increased sampling effort above that determined to achieve acceptable performance under a multinominal tagging model.

Spatial Considerations

The introduction of a spatial dimension into the experimental design will clearly complicate the analysis and, in most instances, the implementation of the tagging experiment. As such a critical question in the design of a tagging experiment is whether it is necessary to include an explicit spatial component. Clearly, if the purpose of the experiment is to get measures of movement rates or area-specific

---

However, in actual applications in which direct estimates of the sampling variances in the catch-at-age data are available it would not be necessary to make this assumption and thus the estimation bias would not be of concern (see Appendix 6).

population sizes and fishing mortality rates, then the spatial component cannot be ignored. However, if the primary purpose of the experiment is to provide overall estimates of parameters of primary interest for stock assessment purposes (i.e., natural mortality, fishing mortality and population size), then whether a non-spatial design and analysis will be sufficient is worth considering. In terms of experimental design, this is an overarching question. As such, a focus of the simulations we conducted with explicit spatial structure was a comparison of the precision and accuracy of population wide estimates of abundance and mortality rates when these estimates were derived from the spatial estimation model and when they were derived from the non-spatial estimation model applied to data pooled spatially (see Appendices 10 and 11 for details of the two estimation approaches).

If the spatial component is ignored, the expectation is that the resulting estimates will be biased if there is systematic incomplete mixing of tags across the entire spatial range. The extent of bias would be expected to depend on the extent of non-mixing and the relative differential in abundances and fishing mortality rates by area/seasons. Results in Appendix 12 provide comparisons of the estimates derived using the spatial estimation model with those obtained from pooling the data spatially for an SBT-like situation. Overall, the simulation results indicate that if systematic spatial heterogeneity exists, it is important to include a spatial dimension within the analysis phase to achieve unbiased estimates of population sizes. In terms of the precision of the estimates, there appears to be little or no cost for including a spatial dimension when in fact complete mixing exists (see Appendix 12 for more detail).

Estimation of Reporting Rates and Catch-at-age

We would expect that having better data with which to estimate reporting rates and catch-at-age will ultimately lead to better estimates of the mortality rate and abundance parameters. To explore this, all of the simulations conducted with explicit spatial structure (Appendix 12) were performed using two levels of data quality for estimating reporting rates and catch-at-age. The high quality data scenarios, as expected, always provided more robust parameter estimates than the low data quality scenarios. Somewhat surprising was that the differences in performance were not larger than observed. With poor information on reporting rates we had anticipated that there might have been much greater confounding between reporting rates and mortality rates than was apparent in the results. In general, the parameters in the low quality data scenarios remained unbiased. The CVs, while being substantially greater than in the high data quality scenarios, were often within what might be considered acceptable limits in a stock assessment context (i.e., less than 30%). However, we would reiterate that the CVs in the simulations are likely to be over-optimistic to what would be realised in practice because of overdispersion in the tag-return data.

Competition may exist between resources devoted to collecting sufficient data for reporting rate and catch-at-age estimation and resources directly involved in tag releases. Appendices 7 and 9 examine the trade-off between the number of releases and the level of observer coverage in a non-spatial context. Note that the level of observer coverage affects the estimation of both the catch-at-age and the reporting rates. Not surprisingly, we found that the marginal gain will be greatest for the weakest element in the experimental design. Thus, if observer coverage is low (e.g., less than 10%), the potential improvement from increased observer coverage will be

substantially greater than if it is 50%. Overall, it is important in the design phase to strive for a reasonable balance among the resources devoted to each of the data collection components, as it is the component with the poorest quality data that will contribute most to the uncertainty in the results.

In design considerations regarding catch-at-age estimation, it is important to ensure that the degree of sampling error in the catch-at-age is accurately reflected. In particular, differences in catchability and selectivity among vessels and areas will result in more sampling variability in the catch-at-age data than if a simple random sample of the catch were available. In cases where observers perform most or all of the catch sampling, the increased variability is likely to be substantive at low levels of observer coverage (e.g., observers coverage of 10% of the total catch may result in sampling from only a small number of vessels/fishing trips operating in high seas longline fleets where a single trip may be several months or more). Thus, when considering the level of effort in the design phase to put into the estimation of the age distribution of the catch, it is important to represent the actual level of sampling variability. This will be fishery specific and further work is needed for almost all fisheries on how best to characterize this. In terms of the estimation framework developed here, this might best be done by developing a functional relationship between the actual observer/sampling coverage and the effective coverage in terms of simple random sampling.

Tagging of Multiple Cohorts

Most of the simulation results presented in this report are for a tagging experiment involving a single cohort. In practice, it is likely that two or more cohorts would be tagged in a given year. If any parameters are thought to be the same between cohorts, then tagging multiple cohorts can improve the information available for estimation and inference about not only these parameters, but the other model parameters as well.

We conducted simulations in a non-spatial context to explore the effects on the mortality rate and abundance estimates if a tagging experiment includes multiple cohorts (see Discussions of Appendices 5, 7 and 9). In these simulations, fishing mortality rates and reporting rates were allowed to differ by year and age but natural mortality rates were constrained to differ only by age, not year. Having additional cohorts resulted in large improvements in the precision of the natural mortality rate estimates and also in the fishing mortality rate estimates at older ages. For example, in the specific simulations conducted in Appendix 7 (which directly incorporate the estimation of reporting rates through observer data), having data from two cohorts reduced the CV of the natural mortality rate estimates by up to 20%, and by almost 30% with three cohorts. There was very little difference in the precision of the estimates of the fishing mortality rates for younger ages or for the initial population sizes regardless of the number of cohorts. However, if we had considered a situation where a separable model is realistic for the fishing mortality rates, then considerable improvement in the precision of the estimates of fishing mortality rates could be achieved from tagging multiple cohorts. If separable assumptions were found to be applicable, this may also improve the overall estimation performance of the remaining parameters, particularly given the high correlation among them (see Discussion of Appendix 5 with regard to parameter correlations).

Multi-cohort tagging would also improve the estimation of area-specific parameters in a spatially-explicit model. Simulation results presented in Appendix 12 found that with spatially-explicit models the population wide estimates were generally well estimated even through the area-specific estimates of population size, fishing mortality rates and transition probabilities often were not. The spatial model for a single cohort has a large number of parameters to estimate relative to the number of data points (i.e., the number of tag returns by age in each area and time period). In a situation where transition probabilities are assumed to be age-dependent but not year-dependent, having return information from multiple cohort for estimating the transition probabilities would reduce the parameterization and may provide a powerful approach for gaining insights into the actual spatial dynamic processes through the testing of alternative model structures for the movement dynamics.

Tag Recovery Activities – Promotion and Rewards

Tag recovery promotion and rewards are an essential component of the experimental design for a tagging program because achieving high reporting rates can substantially improve mortality rate and population size estimates. Resources for this are essential unless observers or automatic detection devices are to be the only source of data for recaptured tags. However, this is perhaps the most nebulous component of the experimental design to evaluate what would constitute appropriate resources. This is because it is not possible to specify the relationship between the resources expended upon promotion and rewards and the level of reporting that will result. Nevertheless, experience gained from past tagging experiments indicates that promotional activities and rewards can have a large effect on the actual return rate (e.g., Polacheck et al. 1995).

We conducted simulations in the non-spatial context to explore the effect of different reporting rates on parameter estimation in the case where reporting rates are estimated from observer data (Appendix 7). As would be expected, all parameters were estimated more precisely when the reporting rate was high (0.75) than when it was low (0.25), regardless of the number of releases or the level of observer coverage. However, the magnitude of the improvement differed among parameters, being greatest for the reporting rate estimates but also substantial for the fishing mortality rates at older ages and the initial population size, and it also depended on the number of releases and level of observer coverage, generally being greatest at low levels of releases and/or low levels of observer coverage. Refer to Appendix 7 for specific results.

Roughly speaking, increasing the reporting rate has a similar effect to increasing the number of releases because both result in more tag returns. However, the increase in tag returns due to increasing the reporting rate depends on the level of observer coverage – if the observer coverage is close to zero, then tripling the reporting rate in the unobserved component will result in almost triple the total tag returns (i.e., effectively a tripling of the number of releases), whereas if the observer coverage is close to one, then tripling the reporting rate in the unobserved component will have almost no effect on the total tag returns (i.e., effectively no change in the number of releases). Of course, it is not quite that simple because increasing the number of releases leads to proportional increases in tag returns in both the observed and

unobserved components, whereas increasing the reporting rate only leads to increased tag returns in the unobserved component.

In terms of experimental design, the critical question is what level and types of promotional activities and what type and value of rewards will result in high return rates from fishermen. Experience from past tagging experiments indicates that this is likely to be fishery specific (e.g., cash rewards can be unacceptable in some cultures, while non-cash rewards can quickly lose their incentive for fishermen that recapture numerous tags). Experience also suggests that fishermen respond positively to personalized contact promoting tag returns (e.g., pre-season discussion explaining the tagging program and in-season meeting of vessels returning from fishing trips) and that such personalized contact can be highly effective in improving return rates (T. Polacheck, personal communications). For example, the tag-return rates from Japanese longline vessels in South Africa appear to have been substantially greater from vessels that were visited by tag collect liaison officers than from vessels without such visits. Thus, personal and direct interaction with fishermen should be an important component of any tag recovery strategy.

Age Estimation Issues

The multi-year tagging framework investigated in this project assumes that tagged fish can be accurately aged and, thus, correctly assigned to their respective cohorts. Therefore, an important consideration in the experimental design will be the method used to estimate the age of tagged fish. The simplest and most easily implemented approach, and the approach used for SBT, is to estimate age based on length using an estimated growth curve and cohort slicing. In fact if data are collected on the size at release and recapture, the tag-return data from the experiment can be used as the basis for estimating or updating the growth curve (although this would entail some delay before ages could be assigned to tagged fish). However, cohort slicing from a length curve will result in aging errors when ages are assigned to individual fish or groups of fish. Thus, when cohort slicing is used to estimate the age of tagged fish, it is important to know whether aging errors, if ignored, would induce substantial errors into the population dynamics parameter estimates derived from tagging models. Simulations were conducted in Appendix 8 to address this issue.

The results in Appendix 8 indicate that in many situations the effects of aging errors from cohort slicing on the mortality rate estimates obtained are minimal. In particular, Appendix 8 suggests that using cohort slicing for juvenile SBT tagging experiments is unlikely to induce substantial biases. However, this is not always the case – in situations where the number of fish tagged in a given year from the cohort being modelled is much smaller than the number tagged from adjacent cohorts and where the degree of overlap in lengths among ages is large[10], some of the estimates (namely, fishing mortality at older ages and natural mortality) have large biases. Furthermore, the true mortality rates affect the results. We found larger biases in the parameter estimates when fishing mortality decreased with age than when it increased. In

---

[10] Note that the degree of overlap is a function of both the variance around the mean length at age for any given age and the actual rate of growth. Thus, a high degree of overlap can be achieved even with relatively low variance in the length at age if growth rates are correspondingly high (see discussion in Appendix 8).

situations where mortality rates (both natural and fishing) are the same at all ages, aging errors have no consequences.

In cases where there is potential for aging errors from cohort slicing to introduce substantial biases, an alternative to relying on cohort slicing for aging of tagged fish should be considered in the experimental design phase. These would include:

1. Collection of a scale from each tagged fish if scales can be used to reliably estimate age for the species being tagged;
2. Collection of otoliths from a sample of fish during tagging operations in order to be able to produce an age-length key that could be directly applied to the tagged fish;
3. Incorporation of aging error directly into the estimation framework.

It should be noted that use of an age-length key would still necessitate the incorporation of aging errors directly into the estimation framework. This is because an age-length key does not provide a unique estimate of the age for each tagged fish but only a probability distribution for its age given its length. Incorporation of aging errors directly into the estimation framework is conceptually straightforward. It would involve specifying a likelihood function for the age of each released fish given its length. Then, conditional on its age of release, the other likelihood functions developed in this report can be used to estimate the conditional likelihood of recapturing a tagged fish. These conditional likelihoods would then need to be integrated over the possible ages of releases to provide an overall unconditional likelihood. While conceptually straightforward, it would be computationally complex and time consuming. Thus, in any specific application it is important to determine whether aging errors are likely to be important early in the design phase.

CCSBT SRP Tagging Program

As noted above, specific results about experimental designs are dependent upon the specific dynamics of the stock and fishery particularly as the conditions under which the experiments are conducted departed from the 'ideal' situation (i.e., known reporting rates and complete mixing). Appendix 9 provides an example of how the estimation framework combined with simulation results can be used to provide experimental design advice in a specific application intended to represent SBT. The CCSBT as part of its international Scientific Research Program (SRP) is currently undertaking a large-scale, multi-year conventional tagging program. The analyses in Appendix 9 were presented at a technical review of the tagging program conducted by the CCSBT Scientific Committee in 2004 and were the basis for a number of the conclusions agreed upon at that review (Anon. 2004)

The focus of the simulations conducted in Appendix 9 was to investigate levels of observer coverage and tag releases necessary to achieve reasonable precision in mortality rate and abundance estimates in a tagging program for juvenile SBT given the current fisheries and population abundances. The results suggest that the number of tags that have been released or plan to be released as part of the CCSBT SRP tagging program are adequate, but that increasing observer coverage in longline

fisheries that capture juvenile SBT[11] from current levels could potentially lead to significant improvements in the precision of the fishing mortality rate estimates for the longline fishery, as well as smaller improvements in the estimate of population abundance. The results from the model which allowed for overdispersion in the tag returns suggest that in order to achieve CVs of 20% or less for fishing mortality rates estimates at ages 1 to 3 in the longline fishery, observer coverage must be at least 30% (and at least 20% for the model without overdispersion). The agreed current target observer coverage by CCSBT is 10% and this target is not being achieved in all of the major longline fisheries. The results indicate that estimates of fishing mortality in the surface fishery are chiefly unaffected by the level of observer coverage in the longline fishery, provided fairly accurate estimates of surface fishery reporting rates and catch-at-age by fishery exist. It is important to note, however, that the results depend on the assumption of no systematic non-mixing. If this assumption is violated, then the level of observer coverage in the longline fishery would become more important because without good observer data, and thus good information on differential tag reporting and return rates between fishery components, there is little power to test the assumption of non-mixing and, if necessary, develop spatially-explicit tag recovery models to account for heterogeneity in recapture probabilities. The simulation results also demonstrate the importance of having reliable and precise estimates of the catch-at-age for each fishery when applying the estimation model presented here. This emphasizes the need to develop appropriate sampling and error models for the catch data; having representative and adequate observer coverage can help to accomplish this in longline fisheries.

Potential Utility of Archival Tags

Archival tags have the potential to improve the estimation in models with spatially-explicit structure and can be combined with data from conventional tags to improve estimates of movement rates (Appendix 10). However, if adequate conventional tagging can be conducted in all areas and reporting rates are not low, the additional information will be of limited direct value in improving the precisions of the estimates within the context of the spatial estimation models developed in this report (see Appendix 10 for more details). In essence, data from conventional tags are sufficient for providing the information required for estimating the parameters of the model, while the additional movement information from each archival tag can be compensated for by releasing additional conventional tags. Given the current relative differences in the cost of archival and conventional tags (i.e., ~$1500 versus $1-2), archival tags would not be a cost effective approach for improving tagging experiments in which there were no spatial/temporal constraints or problems with releases or returns[12].

In a wide range of simulations that we conducted, we were unable to find situations in which having large percentages of archival tags compared to having similar numbers of conventional tags results in markedly improved performance in terms of the precision of the estimates except for the estimates of the transition probabilities in

---

[11] Some longline fisheries (most notably those on the SBT spawning grounds) do not catch juvenile SBT. Also, observers do not provide a viable approach for estimating reporting rates from the surface fishery and other approaches (e.g. tag seeding) are used for this component of the fishery.

[12] Possible exceptions would be where the marginal cost of tagging additional individual fish was high (e.g., adult bluefin tuna, which can have a market value of thousands of dollars, or where the cost of daily tagging operations are high and catch rates are low in some areas).

some cases (Appendix 10). This may seem surprising given the seemingly increased information retrieved from archival tags. However, the actual additional information relative to the parameters being estimated is small. This is because the main information for estimating total mortality rates is whether or not a tag has been recaptured and so archival tags provide little or no additional information in this respect. While the archival tags do provide some additional information on movements, even this is relatively small within the structure of the large-scale area models developed here because the fraction of archival tags that are likely to be captured with information on movements beyond a single time step will usually be small. As such, even for the estimation of transition parameters, archival tags will only provide small amounts of additional data in most cases.

Nevertheless, a relatively small number of archival tags may be highly informative for developing the appropriate spatial/temporal structures to use in the estimation model as they provide direct information on where fish have spent their time and the type of movements that need to be allowed for (e.g., which transition probabilities can reasonably be assumed to be equal to zero). Archival tags can also be informative with respect to the underlying movement dynamics. For example, with seasonal migrations, archival tags provide direct data for estimating the degree of site fidelity (i.e., the extent to which the probability of a fish returning to an area depends upon whether it has been there previously). In this regard, archival tags can contribute to appropriate experimental design as well as estimation model development and selection. This is particularly true in fisheries where there is little or no prior information on the movement and spatial dynamics of a population.[13] As such, we think that archival tagging should be considered an important component in the overall context of experimental design for a tagging program. However, the appropriate balance among conventional and archival tags will be more a matter of judgment because of the primary role of archival tags in model development and selection in contrast to parameter estimation.

Discussion

Analyses of the 1990s SBT Tagging Experiments

The analyses of the 1990s tagging experiments presented in this report (Appendices 15, 16 and 17) are the most comprehensive that have been conducted to date. They not only include integration of the tag-return and catch-at-age data to produce estimates of fishing mortality, natural mortality and population abundance estimates, but also include consideration of large-scale spatial heterogeneity. In addition, they are the first analyses that have examined and incorporated estimates of tag shedding rates. They also use updated estimates of reporting rates that were generated as part of this project. The three different analyses are all relatively consistent with respect to recruitment/juvenile abundance trends in the 1990s. They suggest that population abundance at age 1 (i.e. essentially recruitment) decreased by over 50% for the 1993 and 1994 cohorts compared to the three preceding cohorts. The analyses which provide direct estimates of mortality rates indicate that natural mortality for SBT at age 1 is quite high (~0.4) and decreases with age. In addition, fishing mortality rates

---

[13] Although in such situations, archival tags still do not contribute substantial additional information for estimating mortality rates and abundance.

in the 1990s were generally close to zero for ages 1 and 2, were greatest at ages 3 and 4, and declined at age 5. Further, fishing mortality increased from 1994 to 1997 to relatively high levels (~0.4 at ages 3 and 4).

Although we endeavoured to make our analyses as comprehensive as possible, there is still scope to further improve the parameter estimates, primarily with respect to variances. In particular, exploration and application of a model for the tag-return data, such as the Dirichlet-multinomial, that allows for more variability in the returns than a multinomial model would provide better representation of the variability. However, the marginal gain of using such a model in terms of providing improved/realistic variance estimates for the mortality rate and abundance estimates is likely to be small relative to improving other components of the model. Specifically, there are clear limitations in the data for estimating reporting rates and the catch-at-age. In particular, data for estimating reporting rates are missing for large components of the fishery, and the available information and sampling protocols for the catch data are insufficient to derive variances for the catch-at-age estimates. As such, arbitrary assumptions were required about the level of uncertainty in these two data sets, and the variance estimates for the mortality rate and abundance estimates are highly dependent upon the assumptions made.

In addition to variance estimation issues, the design of and data collection from the 1990s SBT tagging experiments were inadequate to fully address issues related to non-mixing. These inadequacies stem from (1) lack of tagging in regions outside of Western and South Australia and (2) lack of data for estimating reporting rates from the high seas fisheries that catch juvenile SBT (particularly the Japanese and Taiwanese fisheries in the Indian Ocean). This combination of inadequacies meant that reporting rates and movement/spatial dynamics are highly confounded in any analysis (i.e., to what extent were the reported number of tags in the longline fisheries a function of availability of tagged fish versus reporting rates). While there appears to be some robustness in the estimates (particularly with respect to temporal trends), the reliability of the results depends to a large extent on the perception that a large fraction of SBT juveniles are found in the southern Australian waters and the assumption that mixing is reasonably consistent and complete.

The limitations inherent in the design and data available from the 1990s tagging experiments have implications for the design and implementation of future experiments (including the current SRP tagging program being conducted by the CCSBT). It is imperative to ensure that adequate data are collected for estimating reporting rates from all major fisheries which catch juvenile SBT. In addition, it is important to ensure that the estimation of the catch-at-age is based on a statistically valid sampling approach that would ensure the estimates are unbiased and that would provide the bias for direct variance estimation. Also, releasing tags in areas other than the Great Australian Bight would greatly increase the extent to which issues of large-scale spatial heterogeneity could be addressed, although these issues can be at least partially addressed with releases only in southern Australia if reporting rate estimation problems were resolved (see Appendix 12 for further discussion).

Viable solutions for addressing the inadequacies in the 1990s tagging experiments in future tagging experiments exist. These have been discussed in the appendices and within meetings of the CCSBT. Issues of reporting rate estimation and catch-at-age

estimation for longline fleets could be resolved using observer programs that are adequately designed and implemented to ensure representative and sufficient coverage. The problem of wider geographic distribution of tag releases could also be resolved through having observers tag and release juvenile fish[14]. For the current surface fishery, in which 100% of the catch goes into tuna farms, the problem of reporting rates is resolvable by tag seeding (Stanley and Polacheck 2003; Polacheck and Stanley 2004) and catch-at-age estimation by improved length frequency sampling. However, neither adequate longline observer programs (Anon. 2004) nor improved surface fishery length frequency sampling have been implemented in conjunction with the current CCSBT SRP tagging program. While recognition of the critical importance of observers for the SRP tagging program exists, there has been insufficient cooperation, support and resources for adequate observer programs to have been implemented. As such, it appears that the data from the current CCSBT SRP tagging program for addressing the critical reporting rate, catch-at-age and spatial heterogeneity issues may be no better, and perhaps worse, than that available from the 1990s. For SBT, tagging experiments have the potential to provide estimates of critical parameters in the stock assessment, to decrease uncertainty in these parameter estimates, and to reduce reliance on CPUE indices. However, without the necessary resources and access to the fishery to address these data collection issues, the potential is unlikely to be realized.

Experimental Design and Estimation Framework

The estimation framework developed in this report along with the simulation results demonstrate the power of tagging experiments to provide robust estimates of fishing mortality rates, natural mortality rates and population size, which are the main parameters needed for the assessment of fish stocks. As such, tagging experiments have a large potential to reduce uncertainty in many stock assessments, particularly those reliant upon CPUE indices of abundance. The estimation framework developed here results in a synergy from combining traditional multi-year (i.e., Brownie) tagging experiments and catch-at-age data. The catch-at-age data will generally be required in any case in order to be able to estimate reporting rates and incorporating it into the estimation framework allows for joint maximum likelihood estimates of mortality rates and population size. However, both the design and implementation of tagging experiments are demanding and complex. Poorly designed and/or executed experiments can yield either uninterpretable data or parameter estimates with substantial biases and uncertainty. In particular, it is essential that the issues of reporting rates, catch-at-age estimation, tag shedding and incomplete mixing are addressed in the design and analysis stage. It is also essential that there is a commitment to ensure necessary access and availability to the fishery for the collection of data for reporting rates and catch-at-age estimation (e.g., observers, tag seeding, etc.). Simply releasing large number of tags without ensuring that these prerequisites are met is unlikely to yield informative results.

While resolving issues of reporting rates, catch-at-age estimation, tag shedding and incomplete mixing can be challenging, the estimation framework and simulation

---

[14] Pilot tagging projects by observers on longline vessels have demonstrated that this is a feasible approach for tagging juvenile SBT in areas where they are caught by commercial longliners. Return rates from such tagging suggest that tag-induced mortality is not substantially different from fish tagged in the surface fishery.

results demonstrate that there are tractable solutions. In particular, the combination of observers and tag seeding provides a viable approach for estimating reporting rates for most components in large-scale tuna fisheries. Observers can also simultaneously provide data for catch-at-age estimation, and for fishery components in which observers would not be effective (e.g., in tuna farming operations) well-designed port sampling programs can provide this information. Observer and port sampling programs are clearly expensive but need to be evaluated in the overall data collection strategy for a fishery; that is, the data provided by such programs will often be an essential component of the monitoring strategy for a fishery irrespective of whether tagging experiments are being conducted. It should be noted that automatic tag detection systems exists but are not currently a feasible alternative for many fisheries (see Appendix 18). However, these technologies are developing and may make automatic detection systems cost effective and feasible alternatives for reporting rate estimation in the future. If tagging is to be used as a long term monitoring and assessment strategy in a fishery, the development and implementation of automatic detection systems should be undertaken and evaluated.

The issues related to spatial heterogeneity and incomplete mixing are the most challenging for the design and analysis of tagging experiments. They may also limit the applicability of tagging experiments if there are systematic spatial structure issues that cannot be taken into account in the analysis because of design or implementation limitations (e.g., lack of access to areas for tagging). In such cases, the mortality rate and abundance parameters may become inestimable or else be estimated with large biases and uncertainty. Spatial heterogeneity and incomplete mixing would not be an issue if tags were released in proportion to abundance (i.e., random sampling). However, the problem of incomplete mixing can occur across a continuum of spatial scales.

At the smaller geographic scale, ensuring that tagging effort is spread locally and allowing for some period of non-mixing should mediate smaller scale spatial heterogeneity/incomplete mixing effects (e.g., those related to schooling) and avoid effects that would introduce systematic biases (i.e., a preponderance to tag fish that would be more or less vulnerable to recapture in the commercial fisheries). Nevertheless, overdispersion in the returns relative to a multinomial distribution would be likely (see Discussion in Appendix 5). If overdispersion exists, it needs to be accounted for in order to prevent unrealistically high expectations and levels of confidence about the precision of the parameter estimates. The use of the Dirichlet-multinomial distribution as developed in Appendices 9 and 10 is one approach that allows for such overdispersion to be addressed. In terms of experimental design, in most cases there would be little basis for anticipating in the design phase what level of overdispersion would be realistic to expect. As such, knowing what would constitute sufficient resources (e.g., number of releases, observer coverage) to achieve an acceptable level of performance is problematical. Calculations of precision for experimental designs based on assumptions of multinomial sampling for the tag returns provide an indication of the maximum level of achievable precision. The level of additional resources needed to allow for overdispersion would largely be a matter of judgment without information on realistic levels of overdispersion to expect; however, calculations based on a Dirichlet-multinomial model would provide some guidance. We also note that the primary effects of overdispersion appear to be on the levels of precision rather than on the relative trade-off in effort devoted to the

different design components of the experiment (Appendix 9); therefore, results based on a multinomial tag-return model or a Dirichlet-multinomial model with an assumed level of overdispersion should provide some guidance on the relative resources to devote to the various components.

At the larger scales for widely distributed populations, it is difficult to conceive of practical experimental designs that would ensure that releases were distributed in proportion to abundance and thus be assured to avoid bias problems associated with spatial heterogeneity and incomplete mixing. In such cases, it is important that the experimental design anticipates and allows for spatially-explicit modelling of the release and recapture data (e.g., Appendices 10 and 11). The best way to ensure that spatial heterogeneity issues can be fully addressed in the analysis phases is by spreading tagging effort and releases across the geographic range of the population being tagged[15]. In the analysis stage, the range of spatial dynamic models that can be tested will depend upon the spatial/temporal pattern of releases, with the range being greater as the distribution of tag releases is broader. As such, when designing tagging experiments, it is critical to consider what is the appropriate range of spatial hypotheses/models and ensure that the release strategy will allow these to be explored and tested. In this context, simulation testing of the design is critical to ensure that robust estimates can be achieved, particularly for situations with complex spatial structure (see Appendix 12 for further discussion).

The simulation results generally suggest that natural mortality rates are particularly difficult to estimate accurately and precisely without relatively large numbers of releases and high quality supporting data for estimating reporting rates and age distribution of the catch. However, this needs to be considered in relationship to what other alternatives are available for estimating natural mortality, because for most fish stocks there are no direct alternatives. Most commonly, stock assessments use an assumed value for natural mortality (or an assumed range of values to allow for uncertainty)[16]. In addition, the simulation results were generally based on tagging one cohort of fish. Most tagging experiments would involve tagging multiple cohorts over a number of years, and such multi-cohort experiments would provide a more robust basis for the estimation of natural mortality rates (as shown in Appendices 5, 7 and 9).

In summary, tagging experiments can be extremely informative and provide perhaps the only alternative to relying on CPUE indices for assessment of many tuna and other fish stocks. However, to be successful it is critical that the experiments are well designed with sufficient resources for implementation, and that appropriate mechanisms and access are available for the collection of the required data.

---

[15] In many cases there will be practical limits to covering the boundary areas or "fringe" portion of the stock. Such practical limits exist no matter what method is used to monitor a population (e.g. surveys, CPUE indices). Pragmatic approaches are required that attempt to ensure all major components are sampled so that any bias would be small. Note one advantage of tagging is that if mixing of tagged fish with those in marginal areas occurs it mitigates the effect of not actually having tagged in those areas.

[16] In some cases, natural mortality rate is an estimable parameter within the stock assessment model but the information for its estimation is dependent upon input priors and upon structural assumptions within the model (e.g., selectivity constraints and stock recruitment relationships).

Benefits

The results of this research have already been of direct benefit to the CCSBT Scientific Committee in its design, review and recommendations for improving its juvenile conventional tagging program being conducted under the CCSBT Scientific Research Program (SRP). Thus, several of the appendices in this report were submitted to meetings of the Scientific Committee. These papers have been integral to the review of the current program and for developing recommendations for improving the program, including those for improving reporting rates and their estimation. Hopefully, this will provide a catalyst for actual improvements. The methods developed here also provide a robust and improved basis for maximizing the information that can be derived from tagging experiments – in particular the integration of tag-recapture data with observer and catch data provide a comprehensive method for directly estimating mortality rates (both natural and fishing) and population sizes, while accounting for reporting rates. These are the primary quantities required to be estimated in stock assessments, and having an approach for directly estimating them that does not require catch rate (CPUE) data provides a potentially powerful alternative for augmenting traditional stock assessment methods.

The research has demonstrated both the importance and feasibility of accounting for large-scale spatial dynamics within tagging experiments aimed at estimating mortality rates. This should lead to both more robust designs, implementation and analyses in future experiments. The results also provide the basis for evaluating trade-offs in resources that should be devoted to different data collection components of a tagging experiment. This should provide the basis for improved efficiency and cost effectiveness in future experiments.

The results of this research have also provided comprehensive estimates of mortality rates and abundance from the SBT tagging experiments conducted in the 1990s. These results are of direct benefit to the SBT stock assessments by (1) suppling direct estimates of natural mortality rates for juvenile SBT and (2) providing a basis for evaluating the appropriateness of the stock assessment results by directly comparing the estimates from the two different analyses. In addition, the methods developed here provide the basis for a more robust and statistically appropriate approach for directly incorporating the data from these and future tagging experiments (including the current CCSBT SRP program) into future SBT stock assessments. This should reduce the reliance on CPUE within the SBT assessments.

The results from this research provide the first analytical evaluation of the value of archival tags within the context of conventional tagging experiments. The results indicate that under most situations archival tags are likely to be of limited direct benefit at the estimation stage of a conventional tagging experiment. However, archival tagging data are likely to be of critical importance in generating hypotheses about the spatial and temporal dynamics of the species and in model selection. The results provide the basis for evaluating the relative roles of the two types of tagging and their appropriate use within an overall research framework.

Finally, while the results of this research demonstrate the value and potential of tagging experiments to improve the information basis for stock assessments and

reduce the reliance on CPUE, this is only achievable if these experiments are appropriately designed and implemented. Hopefully, the results of this study will provide impetus for careful consideration and appropriate implementation of such experiments across a number of fisheries and, consequently, result in an improved basis for assessing and managing them.

Further Development

A number of areas where there is scope for improvement in the analytical and statistical methods, or for further analysis of the historical SBT tagging data, are identified in the appendices of this report. Particular problems that would warrant further investigation include:

1. Methods for modelling and estimating overdispersion in tag-recapture experiments taking into account the problems of estimating overdispersion parameters in a likelihood estimation context;
2. Methods for estimating variance in complex tagging experiments - in particular bootstrap and Monte-Carlo approaches;
3. Methods for estimating the uncertainty associated with catch-at-age data;
4. Applicable spatial models for the 1960s tagging data taking into account the NSW surface fishery component;
5. Development of statistical models for intermediate movement dynamics between Markovian (random) transition probabilities and transition probabilities with complete memory.
6. Consideration of alternative spatial models for the SBT tagging data in conjunction with analyses of the accumulating archival tag data as a source of information for the development/specification of alternative hypotheses and model structures.

In addition, the results in this report point to the critical need to improve the data on reporting rates and the catch-at-age data from SBT fisheries if the potential for the current large-scale CCSBT SRP tagging experiments is to be realized. They further suggest the importance of developing and implementing feasible approaches for automatic detection of recaptured tagged fish for improving future tagging experiments, both for SBT and other species.

Planned Outcomes

The primary planned outcome from this project was the development and exploration of a robust estimation framework for estimating mortality rates for tuna stocks from conventional tag-recapture data. Such a framework was developed in the course of this project and used to evaluate trade-offs in the design components of tagging experiments. This basic framework and several extensions of it were completed in the course of this project and have formed the basis of working papers to the CCSBT Scientific Committee (Polacheck et al. 2003; Eveson et al. 2004, 2005) and have played an important role in the design and review of the CCSBT SRP tagging program. In addition, a manuscript based on Appendix 5 has been submitted and accepted for publication in the primary scientific literature. We also anticipate preparing several additional manuscripts based on results from this research.

The other main planned output from this project was the application of the estimation framework to tag-recapture data for southern bluefin tuna. This was accomplished in Appendices 15 and 16. The results from Appendix 15 were provided to this year's CCSBT Scientific Committee and it is anticipated that the remaining results will be provided in subsequent papers to the CCSBT Scientific Committee, as well as written up for publication in the primary literature. It should also be noted that the methods developed for incorporating tag shedding estimates from double-tagging data (Appendix 14) are being incorporated into the chapter of a book near completion on the use of multi-year tagging experiments (e.g. Brownie-type models).

In addition, to the above initially planned outcomes, this project was extended without any additional FRDC funding to allow work directly related to the project's objectives to be completed. This additional work was specifically requested by the CCSBT Scientific Committee as input to a review of the CCSBT SRP conventional tagging program conducted in 2004. Three specific outcomes were identified in this extension. These were completed (Appendices 9, 17 and 18) and presented as working papers to the CCSBT Scientific Committee Meeting in 2004 (Eveson et al. 2004; Polacheck et al. 2004a,b).

Conclusions

All of the primary objectives of this project have been completed. An integrated estimation model for the analysis of tag-recapture and catch data was developed for estimating natural mortality rates, fishing mortality rates and population size. Trade-offs among the various design components of a tagging experiment were explored using this estimation framework and guidelines were provided that should be considered when designing such experiments.

In summary, the results from this project suggest the following general conclusions:

1. The integrated tagging and catch estimation model developed here provides substantial benefit for the analysis of multi-year tagging programs, both in terms of the parameters that can be estimated and improved precision of the estimates, particularly fishing mortality rates.
2. Based on comprehensive application of the integrated tagging and catch model(s) to the 1990s tagging data for SBT:
   - natural mortality for SBT at age 1 is quite high (~0.4) and decreases to about 0.2 by age 5;
   - fishing mortality rates in the 1990s were generally close to zero for ages 1 and 2, were greatest at ages 3 and 4, and declined at age 5;
   - fishing mortality decreased in the first couple of years of the 1990s then increased fairly steadily from 1994 to 1997 to relatively high levels (~0.4 at ages 3 and 4);
   - population abundance at age 1 appears to have decreased by over 50% in 1993 and 1994 compared to earlier in the 1990s.
3. The historic (pre-1990s) SBT tagging data have limited potential to provide quantitative mortality rate and abundance and are not directly amenable to the integrated estimation framework developed within this report.

4. The tag-return and catch data from the 1960s suggest that, historically, a large degree of spatial structure and spatial heterogeneity existed among juvenile SBT found within Australian waters.

5. In the design of tagging experiments, five essential components compete for resources:
   - the number of tag releases (distributed appropriately over time and space);
   - the number of tags returns (i.e., resources devoted to rewards, promotion and recovery activities);
   - the collection of data for estimating reporting rates;
   - the collection of data for estimating catch-at-age data;
   - the collection of data for estimating tag shedding rates.

   If data from any of these components are not available then valid estimation is not possible and serious consideration should be given to whether to proceed.

6. The trade-offs among these various design components in a tagging experiment can be quite complex. For a specific proposed experiment, simulation testing provides the best method of evaluating possible experimental designs and should be undertaken to ensure that robust estimates can be achieved and to evaluate the trade-offs among design parameters.

7. Issues related to spatial heterogeneity and incomplete mixing are the most challenging for the design and analysis of tagging experiments. They may also limit the applicability of tagging experiments if there are systematic spatial structure issues that cannot be taken into account in the analysis because of design or implementation limitations (e.g., lack of access to areas for tagging). However, given appropriate design and implementation, the spatial models developed in this report provide a feasible approach for dealing with spatial heterogeneity.

8. The estimation of reporting rates is essential for the successful implementation of tagging experiments, and the combination of observers and tag seeding provides a viable approach for estimating reporting rates for most fishery components in large-scale tuna fisheries.

9. Archival tags do not provide a cost effective approach for improving the direct estimates of mortality rates and population sizes compared to conventional tags. Nevertheless, relatively small numbers of archival tags can be highly informative for developing the appropriate spatial/temporal structures to use in both the design and analyses of conventional tagging experiments.

10. Tagging experiments can be a powerful method to provide robust estimates of fishing mortality rates, natural mortality rates and population size, which are the main parameters needed for the assessment of fish stocks. As such, tagging experiments have a large potential to reduce uncertainty in many stock assessments, particularly those reliant upon CPUE indices of abundance.

References

Anon. 2004. CCSBT. Report of the Extended Scientific Committee for the Ninth Meeting of the Scientific Committee. 13-16 September 2004. Seogwipo City, Jeju, Republic of Korea.

Barrowman, N.J., and Myers, R.A. 1996. Estimating tag-shedding rates for experiments with multiple tag types. Biometrics, 52: 1410-1416.

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

Cadigan, N.G., and Brattey, J. 2003. Semiparametric estimation of tag loss and reporting rates for tag-recovery experiments using exact time-at-liberty data. *Biometrics* **59**, 869-876.

Caton, A.E. 1991. Review of aspects of southern bluefin tuna: biology, population and fisheries. Inter-Amer. Trop. Tuna Comm., Spec. Rep. 7: 181-357.

Eveson, J.P., T. Polacheck and G. Laslett. 2004. Exploring trade-offs in experimental design of a 2-fishery integrated tag-recapture and catch model for estimating mortality rates and abundance. CCSBT-ESC/0409/16.

Eveson, J.P., T. Polacheck and G. Laslett. 2005. Estimation of mortality rates and abundance for southern bluefin tuna (*Thunnus maccoyii*) using tag-return and catch data from 1991 to 1997. CCSBT-ESC/0509/Info-3.

Kirkwood, G.P., and M.H. Walker. 1984. A new method for estimating tag shedding rates, with application to data for Australian Salmon, *Arripes trutta esper* Whitley. Aust. J. Mar. Freshw. Res., 35, 601-606.

Polacheck, T., W. Hearn, and W. Whitelaw. 1995. Interactions between surface and longline fisheries for southern bluefin tuna based on recent tagging results: The importance of reporting rates. Interactions of Pacific Tuna Fisheries, FAO, Tech.Paper.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1996. Estimation of natural and fishing mortality for juvenile southern bluefin tuna based on multi-year tagging of cohorts. CCSBT/SC/96/11.

Polacheck, T., W. Hearn, C. Millar, and C. Stanley. 1998. Updated estimates of mortality rates for juvenile SBT from multi-year tagging cohorts. 1998. CCSBT-SC/9808/20.

Preece, A., T. Polacheck, D. Kolody, P. Eveson, D. Ricard, P. Jumppanen, J. Farley, and T. Davis. 2001. Summary of the primary data inputs to CSIRO's 2001 stock assessment models. CCSBT-SC/0108/21.

Polacheck, T., P. Eveson and G. Laslett. 2003. Exploring the trade-off between tag releases and observer coverage in the estimation of mortality rates through an integrated Brownie and Petersen mark-recapture estimation approach. CCSBT-ESC/0309/22.

Polacheck, T. and C. Stanley. 2004. Update on tag seeding activities and preliminary estimates of reporting rates from the Australian surface fishery based on tag seeding experiments. CCSBT-ESC/0409/15.

Polacheck, T., P. Eveson and G. Laslett. 2004a. Review of methods for estimating tag reporting rates and their applicability to SBT longline fisheries. CCSBT-ESC/0409/14.

Polacheck, T., P. Eveson and G. Laslett. 2004b. An evaluation of abundance estimates from tagging programs when tag returns are only available from one component of a multi-component fishery: an example based on the 1990s southern bluefin tuna tagging program. CCSBT-ESC/0409/17.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. Charles Griffin, London. 654 pp.

Stanley, C.A. and T. Polacheck. 2003. Report from a Pilot Tag Seeding Program for Estimating Tag Reporting Rates from the Australian Surface fishery. CCSBT-ESC/0309/25.

Appendix 1: Intellectual Property

No commercial intellectual property arose from this work.

Appendix 2: Staff

Tom Polacheck
Geoff M. Laslett
J. Paige Eveson

Appendix 3:

An overview of different tagging experimental designs for estimating population parameters in commercial fisheries

Tom Polacheck, J. Paige Eveson and Geoff M. Laslett

Introduction

The literature on mark-recapture experiments is enormous, with a variety of different experimental designs and estimation models. This reflects the fact that mark-recapture experiments are one of the primary tools for studying the dynamics of wild animal populations and that differing objectives exist for conducting the experiments. In addition, different applications often require developing case-specific variants of a general approach to cope with the complexity of individual situations. The parameters that can be estimated with the differing approaches can also depend upon the actual experimental design and on what assumptions are made about the population dynamic processes.

While there are a number of summaries of the mark-recapture literature that provide technical details of one or more general approaches (e.g., Ricker 1975; Seber 1982; Quinn and Deriso 1999; Morgan et al. 2002), it can be difficult to grasp an overview of the primary features of different approaches and the inter-relationship among them. Such an overview is important in the initial consideration and design of tagging experiments. Having a guide to what parameters can be estimated from different experimental designs and the data requirements associated with them is necessary for evaluating whether a tagging experiment is able to fulfil important information needs and is logistically feasible in any specific situation. It is not infrequent for tagging experiments to have been conducted without careful consideration of the estimation methods that will be applied. After the fact the analyst is faced with trying to find an applicable estimation model. Forethought in the design stage may substantially increase the information that can be extracted from such (often expensive) tagging experiments.

The current appendix presents an overview of potential basic tagging experimental designs for use in commercial fishery situations. We suggest that most mark-recapture experiments in a large-scale fishery context can be classified into one of three basic types based on where the information for estimating the parameters of interest stems from. In this context, we assume that the parameters of primary interest are those related to the population dynamics (i.e., either absolute abundance or mortality rates). We consider possible integration across these different approaches and the role auxiliary data (e.g., fishing effort; abundance surveys) can have in increasing the information derived from tagging experiments.

It should be noted that in this appendix, mark-recapture experiments refer only to experiments in which a single recapture of a tagged animal is possible (i.e. re-releasing of a recaptured tagged animal does not occur).   In experiments where animals are re-released when recaptured, the history of recaptures for an individual provides substantial additional information not available in mark-recapture experiments with only single recaptures. There is extensive literature on tagging experiments involving multiple recaptures (e.g., Ricker 1975; Seber 1982). However, such experiments are not generally applicable to tagging experiments involving commercial fisheries since recaptured fish are generally killed or injured in the commercial harvesting process. Moreover, re-release is generally unfeasible because of the commercial value of harvested fish. Further, ensuring the reliability of the re-released data can be logistically challenging and problematical. As such, we have not

attempted to include such experimental designs in the general overview presented here.

In addition, we only consider the situation of homogenously mixed populations to avoid the complexity of adding a spatial dimension. However, for all approaches, extensions exist, or can be developed, that take into account spatial dynamics (e.g., incomplete mixing, movements and migrations). These extensions are conceptually rather straightforward, using the information on the spatial location of release and recaptures to estimate the spatial dynamics parameters, but can be quite complex mathematically, statistically and computationally (e.g., see Sibert 1984; Deriso et al. 1991; Appendices 10 and 11 of the current report).

Finally, we generally limit our discussion to closed populations in terms of recruitment. While at first this may seem artificial and overly restrictive, a particular cohort or group of cohorts within a population constitutes a "closed" population in this context. At least in a fisheries context, animals from individual cohorts are frequently the main focus of interest. Moreover, if substantial heterogeneity exists in recapture probabilities with age or size, this will likely require stratification of the releases and captures into appropriate groupings. The basic approaches discussed here can be extended to "open" population situations, but require assumptions about the nature of the recruitment process (such as constant recruitment) or that the analysis of the tagging experiment be embedded within a more general population model or stock assessment framework. Nevertheless, the same basic classification of experimental designs would apply in terms of the parameters for which the tagging data would be informative and where the information for estimating them is derived. Thus, we consider that the basic classification developed here provides a useful guide when considering whether tagging experiments should be conducted in particular situations.

A Basic Classification of Mark-Recapture Experiments

While there are a number of alternative ways in which mark-recapture experiments can be classified, we consider the fundamental distinction to be whether data are collected and available only on the number of tagged animals that are recaptured or whether, in addition, data are available on the number of animals examined for recaptured tags. This distinction divides mark-recapture experiments into those in which the primary object is to estimate mortality rates (either natural, exploitation or both) and those that aim to estimate absolute abundance. In the latter case, the information for estimation comes from the proportion of recaptured animals that have tags. The first published use of this approach dates back to 1917 (see Le Cren 1965) and is generally called the Petersen method (e.g. Seber 1982)[1]. As such, we will refer to this general class of experiments that derive their information from the proportion of recaptured animals that have tags as the "Petersen" approach. The first class of experiments, which only uses information on the number of recaptured tags over time, we will refer to as the "return rate" approach. This approach requires information from multiple recapture events (at least three for statistical estimation of variance).

---

[1] Although this method for estimating abundance is commonly called the Petersen method, it appears that Petersen never actually estimated population size using this method and that K. Dahl was the first one to develop and apply it in 1917 to trout populations in Norwegian tarns. Dahl, however, does acknowledge the role of Petersen's tagging experiment in 1889 as providing the starting point for his work (see Le Cren 1965).

Moreover, this class of experiments can be subdivided into two fundamental types: (1) those that utilize only a single release event and (2) those that utilize multiple release events. In the first case (1), the information for estimation comes principally from the rate at which tags are returned over time. We will refer to this class of experiments as the "tag-attrition" approach. In the second case (2), the information for estimation comes primarily from comparing the return rates over time from the multiple release events. Much of the initial development of the theory and estimation methods underlying this type of mark-recapture experiment was done by Brownie et al. (1986), so we will refer to this approach as the "Brownie" approach.

A somewhat obvious extension to these basic approaches is to combine information on both the rate of tag returns and the proportion of recaptured animals with tags within a single experimental design and estimation framework. There has been work along this line within the context of tag-attrition models (Parker 1955; Ricker 1975), although this work has generally not been discussed as an integration of the two approaches. There appears to have been little work and no general framework for integrating the Petersen and Brownie approaches. Part of the reason for this is probably the very different "experimental" situations for which the Petersen approach and the return rate approaches were originally conceived. The Petersen approach has primarily been designed for situations in which the researcher not only tags and releases the animals, but also conducts the recapture component of the experiment. The return rate approaches have primarily been designed for situations in which the researcher is only directly responsible for the release component of the experiment and is dependent on others for the recaptures (e.g., fishers in commercial fishing situations; the general public in bird banding experiments).

Finally, an alternative and related extension is to combine information on the rate of tag returns with auxiliary data, such as measures of fishing intensity (e.g., fishing effort) or abundance indices (e.g., survey results). Such extensions start to expand the analytical problem beyond a mark-recapture context and into a more complex integrated modelling framework involving multiple data sets with common parameters.

Common Assumptions

There are several assumptions common to all of the different mark-recapture approaches, at least in their conceptual development. These include:
1. The animals tagged are a representative sample of the target population[2,3];
2. There is no tag loss;
3. Survival and recapture rates are not affected by the tagging operation or by an animal having a tag attached;
4. The fate of tagged animals is independent of the fate of other animals;
5. Times of release and recovery are accurately recorded within the timescales required by the estimation model;
6. Tagged animals are correctly identified with respect population attributes (e.g., stock, age-group, sex);

---

[2] The target population may be a sub-component of a larger population; e.g. a particular cohort, sex, etc.

[3] Often the fish tagged will not be a random sample but it is assumed that the rapid and complete mixing occurs subsequently.

7. 100% of the recaptured tags are returned or else tag reporting rates are known (note this assumption is not strictly required for the Brownie approach but without it the information obtained is substantially limited – see below).

In most applications, one or more of these assumptions is likely to be violated. This does not necessarily invalidate the approach as long as appropriate data can be collected to estimate the extent of the violation and to provide the basis for accounting for such violations in the estimation framework. In fact, a large fraction of the tagging literature is devoted to accounting for violations of the above assumptions (e.g., estimation and incorporation of shedding rates, reporting rates, incomplete mixing, etc.).

Petersen Approach

The Petersen approach is designed to provide estimates of absolute abundance at the time of tagging. A Petersen estimator is based on the ratio of the observed number of tags returned within samples taken from the population given the known number of tags released into the population. If the animals tagged are a representative sample of the population and thoroughly mixed, then the expected proportion of tags in a random sample from the population will be equal to the proportion of the population that was tagged. Thus,

$$E(P_0) = \frac{C_i N_0}{R_i} \tag{1}$$

where
P_0 = the population size at the time of tagging
N_0 = the number of tags released
R_i = the number of tags recovered at time i
C_i = the number of animals examined for tags at time i (i.e., the catch in a commercial fishery situation)

The Petersen approach has been primarily applied in controlled experimental situations in which the experimenter both conducts the tagging and the subsequent sampling. In such cases, there is no uncertainty about the number of animals sampled for returns. In a fisheries context, the catch data can be considered to constitute a sample from the population. However, in this case, the size of the sample examined for tags is usually not known exactly but is an estimated quantity, and the uncertainty associated with it needs to be accounted for (see Appendix 17).

Note that for the Petersen estimator, tag recoveries can be pooled from recaptures and samples taken over an extended period of time and no assumptions are required about natural or fishing mortality rates (e.g., see Seber 1982; Appendix 17 of the current report).

Population Dynamics Model

In order to be able to partition estimates of total mortality rates into natural and fishing mortality, it is necessary to have a population model specifying how these two interact. The most common practice in fishery models is to assume that they represent competing risks and that they are constant over the time step used to model the

population (commonly a year). For a group of animals of a given age, this yields the following deterministic equations for the number of animals that survive to the next age and the number of animals caught (e.g., see Hilborn and Walters 1992):

$$P_{i+1,t+1} = P_{i,t} \exp\left(-F_{i,t} - M_{i,t}\right)$$

$$C_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} P_{i,t} \left(1 - \exp\left(-F_{i,t} - M_{i,t}\right)\right)$$

where
$P_{i,t}$ = the number of individuals of age i at time t
$C_{i,t}$ = the catch of individuals of age i at time t
$F_{i,t}$ = the instantaneous fishing mortality rate for individuals of age i at time t
$M_{i,t}$ = the instantaneous natural mortality rate for individuals of age i at time t.

Note that $P_{i,t}$ can be replaced by $N_{i,t}$ (the number of tagged animals of age i in the population at time t) and $C_{i,t}$ can be replaced by $R_{i,t}$ (the number of tagged animals recaptured at age i and time t) in the above equations to give predictions of the expected number of recaptures over time from a tagging experiment. In particular:

$$N_{i+1,t+1} = N_{i,t} \exp\left(-F_{i,t} - M_{i,t}\right) \tag{2}$$

$$E\left(R_{i,t}\right) = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} N_{i,t} \left(1 - \exp\left(-F_{i,t} - M_{i,t}\right)\right) \tag{3}$$

Equations such as 2 and 3 that can be used to predict numbers of recaptures form the basis of the estimators used in both tag-attrition and Brownie approaches. Other population dynamics equations can and have been used with either of these approaches, in particular those that represent situations where fishing is highly seasonal or occurs essentially in a single pulse during a year. It should be noted that the Brownie approach can provide estimates of total mortality for the time between tagging events ($Z_{i,t} = F_{i,t} + M_{i,t}$) independent of any population dynamics model.

Tag-Attrition Approach

Equations 2 and 3 can be used to predict the number of tag returns over time from a single release of tags conditional on natural and fishing mortality rates. This is the basis of the tag-attrition approach. However, for every age/time-step, two parameters are required (one for each fishing and natural mortality), while there is only one data point (the number of tag returns at that age or time). Even if natural mortality is assumed to be constant over age/time, which is common in fishery models, there is still one more parameter to be estimated in total (over all ages/times) than data points. In order to obtain a model that is not over-parameterized without including additional data to the tagging data, either natural mortality needs to be assumed known or both fishing and naturally mortality rates need to be assumed constant. Both assumptions have been used (Ricker 1975).

If natural mortality is assumed known, then estimates of age/time-specific fishing mortality rates can be estimated by iteratively solving equations 2 and 3 for the predicted number of tags recaptured over time and the number of tags still remaining at liberty. Even though a value for natural mortality needs to be assumed, the results can still be informative. For example, the range of values of fishing mortality obtained over a "reasonable" (prior) distribution for natural mortality can provide estimates of whether overfishing relative to some reference fishing mortality rate has occurred. In addition, results from experiments conducted at different times can be used to provide estimates of relative changes in fishing mortality rates (assuming natural mortality has not effectively changed over time).

If both fishing and natural mortality rates are assumed constant over time and age, equations 2 and 3 can be used to estimate both quantities. This can be seen by noting that for a deterministic model with constant mortality rates, the ratio of returns in two successive years equals the net survival rate; i.e., from equations 2 and 3, the logarithm of the ratio of returns is $\log (R_{i+1}/ R_i) = -Z = -(F + M)$. Given the net survival rate and the number of releases, equation 3 can be solved for F, and M can then be obtained from $Z–F$. Although the assumption of time and age independent mortality rates would not be very realistic in many situations, the above estimator can still provide useful indications about the average levels of mortality, particularly in situations where there are little other data and over age/size ranges where size and age effects would not be expected to be large.

Finally, there are other potential ways to constrain the parameters so that they are estimable. For example, the tag-return data are considered in a VPA (virtual population analysis) context by selecting or externally estimating a terminal fishing mortality rate (Deriso et al. 2001; Bayliff 1971). Alternatively, it is only necessary to assume constant fishing mortality rates for a limited number of years/ages; time varying fishing mortality rates could then be estimated for the remaining periods. Similarly, instead of assuming constant fishing mortality rates, fishing mortality could be assumed to be a separable process[4] and an age-specific selectivity function could be specified when analysing the releases for a given age. Such an approach would be primarily of use in a multi-year tagging experiment where fish of a particular age (say age 2) are tagged in each year.

Brownie Approach

The Brownie approach uses equations 2 and 3 in an analogous manner to the tag-attrition approach to predict the expected number of tag returns from each release event (Table 1). However, by having multiple release events over time (e.g., releasing fish from the same cohort in consecutive years), most of the over-parameterization problems inherent in the tag-attrition approach are resolved. For example, for three release events and three recapture periods, there are six mortality rate parameters (three F's and three M's) and there are six data points; however, not all data points provide independent information so only five parameters are actually estimable. Generally speaking, an experiment with n release years and m recapture years has $2*m$ parameters but only $m+(n-1)$ of these are estimable. Normally, constraints on the

---

[4] A separable process in this context means that age and year specific fishing mortality rates can be separated into separate, multiplicative year and age components.

mortality rates are imposed to reduce the number of parameters; for example, natural mortality rates are often assumed to be constant over recapture years $n-1$ to m, or a parametric form can be assumed for natural mortality, such as a linear decrease with age (see Appendices 5 and 15).

A potentially useful feature of the Brownie approach is that it can provide estimates of total mortality ($Z=M+F$) for each year of releases even when there is no information on reporting rates. Further, if reporting rates are assumed constant over time, then the approach, in theory, can provide estimates of F, M and reporting rates without any additional data. However, the precision of the estimates is usually unsatisfactory (Hoenig et al. 1998). It is possible in some cases to modify the basic Brownie design to improve the reporting rate estimates, still under the assumption that these are constant over time (e.g., by having multiple releases within a year – Hearn et al. 1998). However, in many fishery situations, particularly ones involving multiple fisheries, reporting rates are likely to vary with time and age. In such cases, additional data are required to account for non-reporting (see Appendices 7 and 18). Alternative parameterisations of fishing or natural mortality rates (such as assuming fishing mortality is separable into year and age effects, and that natural mortality is constant or a function of age) are straightforward to implement and can potentially result in improved performance with tagging of multiple cohorts or if the number of release years is extended.

Integrated Approaches

While much of the mark-recapture literature has been focused on experimental designs in which the primary data collected is either the proportion of tags recovered in a sample from the population (i.e., the Petersen approach) or the rate of return over time (i.e., the tag-attrition or Brownie approaches), in many situations both types of data will be available and there are synergies to be gained by combining them in a single estimator. In particular, both types of data will be available in many fishery situations – tag-return data will be collected in the commercial fisheries over extended time periods (i.e., as long as tags remain in the fish and there is a process for reporting the data), and estimates of the catch-at-age will often be available to provide direct estimates of the relevant numbers of animals examined for tags. In many cases, estimates of the catch-at-age will also be critical in the estimation of reporting rates, and this needs to be considered in the experimental design (e.g., when reporting rates are based on observers, or in multi-fleet fisheries with different reporting rates and selectivities – see Hearn et al. 1999 and Appendix 7).

Combining Petersen and Tag-Attrition Approaches

Combining Petersen and tag-attrition approaches can provide simultaneous estimates of the population size and mortality rates, subject to similar assumptions about natural and fishing mortality rates as discussed above with respect to tag-attrition models. Parker (1963) developed such an integrated method within a tag-attrition context (i.e., for an experiment involving a single release) in which mortality rates are assumed constant. In this case, the model developed was for an open population with respect to recruitment and can also provide estimates of the recruitment rate. The model is equally applicable to a cohort or closed population, in which case the recruitment rate is zero. Extension of Parker's model have been undertaken to accommodate variable recruitment rates and/or variable fishing mortality, but the natural mortality must be

assumed known to obtain estimates of fishing mortality rates that vary over time (see Seber 1982). In an alternative variant, Kleiber et al. (1987) developed a combined tag-attrition and catch model that assumed constant total mortality (Z) but time varying fishing mortality rates. While noting that this is a "paradoxical" assumption, they suggest that it may not be unreasonable if $Z>>F$. This model has been applied to provide mortality rate and abundance estimates for skipjack in the western tropical Pacific (Kleiber et al. 1987).

Within the applications involving a combined Petersen and tag-attrition approach, we are not aware of any analyses that have considered the uncertainty in the number of animals examined for tags (i.e., the uncertainty in the catch and catch-at-age data). Similarly, we are unaware of any analyses that have investigated whether the estimates of the mortality rates are improved (i.e., more precise) by including the catch data within a tag-attrition context. However, if the catch data are known relatively precisely, substantial improvement would be expected, at least for a closed population, because under the assumption of constant mortality rates the decline in the catch over time provides an estimate of total mortality, Z (see Ricker 1975; Hilborn and Walters 1992).

Combining Petersen and Brownie Approaches

Petersen and Brownie approaches can be combined to provide estimates of initial population size along with estimates of natural and fishing mortality rates, with no constraints or additional assumptions needed beyond those required by the basic Brownie approach. Nevertheless, we are not aware of any previous study that has combined these two approaches within a single estimator. The general statistical development of this combined approach is presented in Appendix 5 of this report, and a large portion of this report is devoted to evaluating the performance of this approach in different situations and to extending the approach to include such factors as estimation of tag reporting rates, spatial heterogeneity, initial non-mixing, and more realistic variance structures for the tag return data. The results demonstrate that there are substantial benefits from incorporating catch data directly into a Brownie approach; in particular, including catch data not only allows for abundance to be estimated but it can also improve the precision of both fishing mortality and natural mortality rate estimates (particularly the former – see Appendix 5 for more detail).

Incorporating Auxiliary Data

In addition to combining a Petersen approach with either a tag-attrition or Brownie approach, it is possible to incorporate auxiliary data which in themselves contain information that can be used to directly estimate abundance, mortality rates or trends in either of these. If the auxiliary data are independent of the tagging data and can be put into a likelihood framework, developing an integrated estimator for shared parameters is straightforward. For example, data for estimating abundance from an independent survey may be available that can be combined with a Petersen tagging abundance estimate to provide a single joint estimate; the benefit of such a combined estimate will depend upon the relative precision of the two estimates when calculated separately. If one estimate has much greater precision, then the benefit of incorporating additional data into the estimator will be small.

Effort Data

The more important situations where auxiliary data can contribute to a tagging experiment is when the auxiliary data combined with the tagging data can expand the scope of relevant parameters that can be estimated and/or eliminate assumptions or constraints on the parameters being estimated. In this context, the potential benefits that can be provided by auxiliary data are most relevant to tag-attrition models. The incorporation of fishing effort as a direct measure of trends in fishing mortality rates has been one approach used to resolve the over-parameterization problem inherent in the tag-attrition approach (Lucas 1975; Kleiber et al. 1987). In this case, fishing effort is assumed to be directly proportional to abundance: $F_i = qE_i$, where E_i is the fishing effort in time period i and q is the catchability coefficient. With this assumption, it is possible from combining effort data with tag-attrition data to estimate time varying fishing mortality rates and a constant natural mortality rate (Lucas 1975). The integration of effort data and tag-attrition data has been used to estimate mortality rates for a number of tuna stocks (e.g., Kleiber et al. 1987; Sibert 1984; Bertignac et al. 1998). The further addition of catch data would also allow for the initial abundance at the time of tagging to be estimated.

The precision and accuracy of the parameter estimates obtained from an integrated tag-attrition and effort approach will depend upon the variance and trends in q. Catchability in most commercial fisheries is highly variable due to variability in gear, vessels, fish behaviour and availability. If realistic estimates of uncertainty are to be provided it is essential that the variability in catchability is addressed; however, this has not been the case in most applications that have used effort data. In addition, catchability has generally been found to increase over time (e.g., as a result of new technologies); such temporal trends in q will induce biases in the parameter estimates unless accounted for in the model.

It should be noted that the incorporation of effort data into tagging experiments has essentially the same inherent problems as the use of CPUE indices as measures of relative abundance. In this context, tagging experiments that depend upon the use of effort data do not overcome the problem with fishery-dependent CPUE indices. In fact, concerns about the reliability and interpretation of CPUE data have been a primary motivating factor for considering undertaking large scale tagging programs in commercial fisheries situations.

Effort data can also be brought into the analysis of tagging data from Brownie-type tagging experiments (Hoenig et al. 1998). As the effort data provide a direct estimate of trends in fishing mortality, they can improve the precision of the mortality rate estimates, the extent to which will depend largely on the degree of variability in q. Inclusion of effort data would also eliminate the need for the constraints on natural mortality that are generally needed in a Brownie model – i.e., a separate M_i could be estimated for each release age[5].

Independent Abundance Data

In theory, independent estimates of absolute or relative abundance (e.g., from acoustic, line transect or trawl surveys) corresponding to the whole population or a

---

[5] This feature was not recognized by Hoenig et al. (1998) in their consideration of integrating effort data into a Brownie model.

portion of the population being tagged could be combined with either Brownie or tag-attrition data to provide a complete set of abundance and mortality rate estimates. However, a single independent abundance estimate would not directly contribute to the mortality rate parameters being estimated from the tagging data, nor would the tagging data contribute to the abundance parameter being estimated from the independent survey data. In contrast, if a time series of abundance estimates (either absolute or relative) were available, there would be definite benefits from combining the two sets of data in a single analysis. The difference between successive abundance estimates provides a direct estimate of total mortality (Z) during the time period. Thus, a time series of such independent estimates would contribute to the estimation of mortality rates in a similar way as the incorporation of effort data (but without the problems associated with the interpretation of effort data). Besbeas et al. (2002) have examined this issue within a Brownie context for bird banding data. We are unaware of any work along this line within a tag-attrition context. Further, we note that in many fishery situations, it is unlikely that resources would be available to undertake simultaneous large scale tagging and abundance surveys. In fact, for many fisheries, feasible and practical methods for undertaking such large scale abundance surveys do not exist (e.g., for fisheries on widely dispersed pelagic stocks), and the lack of such methods has been one of the underlying reasons for undertaking tagging experiments.

Discussion

The overview of different tagging experimental designs presented here should not be considered to be exhaustive. In particular, there are a large number of variants for each of the three basic approaches discussed. However, the overview provides what we consider to be fundamental distinctions among different approaches that are important to consider in the design and analysis of a tagging experiment. The distinctions are both in the types of data that need to be collected, the temporal span of the experiment and, perhaps most importantly, the type and range of population parameters that can be estimated if the experiment is successfully implemented. Tables 2-4 provide a summary of the major distinctions among the different approaches. Note that our tabling of the parameters that can be estimated is not exhaustive in that alternative assumptions and/or parameterizations could have been chosen. For example, in Table 2 for the tag-attrition model we state that age-specific Fs can be estimated if M is assumed known; however, we could equally have stated that age-specific Ms can be estimated if F is assumed known. As another example, in Table 3 the integrated models could have been parameterized in terms of the population size at each age (P_i) and only one F parameter (as opposed to F at each age and only the initial population size P_0). The possible variations are numerous, so we chose to present the tables using the more conventional model assumptions and parameterizations.

There are a number of tag experimental design approaches based on having exact times of releases and recaptures over extended periods (e.g., Leigh et al. in press). The additional complexity of these approaches may make it difficult to see their relationship to the basic overview presented here. However, these exact time approaches can be considered as continuous analogues to the grouped time approaches described here. Essentially, if releases have been spread out over sufficient lengths of time such that there will be sufficient contrast in the expected number of returns from early and late releases, the experimental design will be

capable of separating natural and fishing mortality rates (i.e., analogous to a Brownie model). If the release period is relatively restricted, assumptions similar to those involved with a tag-attrition model will be required.

Based on the overview presented here, the integrated Brownie and Petersen approach is the most powerful of the approaches in terms of the range of parameters that can be estimated without underlying assumptions or constraints on the mortality rate parameters. Such an approach can provide direct estimates for the main population dynamics parameters that traditional stock assessment attempt to estimate. The extra resources required to undertake multiple releases for a Brownie experiment relative to a single release tag-attrition experiment would appear to be worth it in terms of removing constraints on year/age-specific fishing mortality rates (which can be highly variable) or having to rely on effort data to provide independent age-specific estimates. Furthermore, the integrated Petersen/Brownie approach (as well as the Petersen/tag-attrition approach) puts extra demands on the data requirements because estimates of the catch-at-age and their statistical properties (e.g., variances and distributions) are needed. Note that these estimates are also fundamental to many fishery stock assessments, and their reliability and precision are important if assessments are to provide reliable estimates of uncertainty and risk associated with different management options. However, in too many fisheries, little focus has been given to rigorous estimation of the statistical properties of the catch estimates. Thus, if implementing a combined Petersen/Brownie tagging experiment increases focus on improving catch estimation, then this would be considered an added benefit rather than a cost – particularly when considered in the broader context of stock assessment and management purposes.

Literature Cited

Bertignac, M., Hampton, J., and Coan Jr, A.L. 1998. Estimates of exploitation rates for North Pacific albacore, *Thunnus alaunga*, from tagging data. Fish. Bull. 97: 421-433.

Bayliff, W.H. 1971. Estimates of the rates of mortality of yellowfin in the eastern Pacific Ocean from tagging experiments. IATTC Bull. 15: 381-436.

Besbeas, P., Freeman S.N., Morgan, B.J.T, and Catchpole, E.A. 2002. Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58: 540-570.

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

Deriso, R.B., Pumsly R.G., and Bayliff, W.H. 1991. A Markov movement model of yellowfin tuna in the Eastern Pacific Ocean and analyses for international management. Fisheries Research 11: 375-395.

Hearn,W.S., Pollock, K.H., and Brook E.N. 1998. Pre and post season tagging models: estimation of reporting rates and fishing and natural mortality rates. Can. J. Fish. Aquat. Sci. 55: 199-205.

Hearn, W.S., Polacheck, T., Pollock, K.H., and Whitelaw, W. 1999. Estimation of tag reporting rates in age-structured multicomponent fisheries where one component has observers. Can. J. Fish. Aquat. Sci. 56: 1255-1265.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London. 570pp.

Hoenig, J.M., Barrowman, N.J., Hearn, W.S., and Pollock, K.H. 1998a. Multiyear tagging studies incorporating fishing effort data. Can. J. Fish. Aquat. Sci. 55: 1466-1476.

Kleiber, P, Argue, A.W., and Kearney, R.E. 1987. Assessment of Pacific skipjack tuna (*Katsuwonus pelamis*) resources by estimating standing stock and components of population turnover from tagging data. Can. J. Fish. Aquat. Sci. 44: 1122-1134.

Le Cren, E.D. 1965. A note on the history of mark-recapture population estimates. J. Animal Ecology 34: 453-454.

Leigh, G.M., Hearn, W.S., and Pollock, K.H. In press. Time-dependent instantaneous mortality rates from multiple tagging experiments with exact times of release and recovery. Environmental and Ecological Statistics.

Lucas, C. 1975. A method for estimating fishing mortality rates from tag recoveries when fishing is not constant. Aust. J. Mar. Freshwater Research 26: 75-79.

Morgan, B.J.T., and Thomson, D.L. (Editors). 2002. Statistical analysis of data from marked bird populations. J. Applied Statistics 29(1-4), 669pp.

Parker, R.A. 1955. A method for removing the effect of recruitment on Petersen-type population estimates. J. Fish. Res. Board Can 12: 447-450.

Pollock, K.H., Hoenig, J.M., and Jones, C.M. 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.

Quinn, T.J., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York. 542pp.

Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bulletin of the Fisheries Research Board of Canada 191. 382pp.

Ricker, W.E. 1948. Methods of estimating vital statistics of fish populations. Indiana Univ. Publ. Sci. Ser. 15. 101pp.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. Charles Griffin, London. 654 pp.

Sibert, J.R. 1984. A two-fishery tag-attrition model for the analysis of mortality, recruitment and fishery interaction. South Pacific Commission. Noumea, New Caledonia. Tuna and Billfish Assessment Programme Tech. Rep. 13.

Table 1. Expected number of returns at age from each release event in a multi-year tagging study involving a single cohort of fish. Note in this table $S_i = \exp(-F_i - M_i)$ and $u_i = \dfrac{F_i}{F_i + M_i}(1 - S_i)$; the subscript for time has not been included for clarity.

Release age	Number releases	Expected number of returns from age class i				
		1	2	3	4	5
1	N_1	$N_1 u_1$	$N_1 S_1 u_2$	$N_1 S_1 S_2 u_3$	$N_1 S_1 S_2 S_3 u_4$	$N_1 S_1 S_2 S_3 S_4 u_5$
2	N_2		$N_2 u_2$	$N_2 S_2 u_3$	$N_2 S_2 S_3 u_4$	$N_2 S_2 S_3 S_4 u_5$
3	N_3			$N_3 u_3$	$N_3 S_3 u_4$	$N_3 S_3 S_4 u_5$

Table 2: Summary of three basic approaches for tagging experiments in which only a single recapture of any tagged animal is possible[1].

Approach	Petersen	Return Rate	
		Tag-Attrition	Brownie
Parameters estimated	P_0	F and M if both constant; **or** F_i if M assumed known	Z_i, F_i, M_i [2]
Data requirements	N_0, R_i, C_i	N_0, R_i	N_i, R_i
Required release events	Single	Single	Multiple (at least 3 consecutive time periods)
Required recapture periods	Single	Multiple (at least 3)	Multiple (at least 3)
Required external parameters	None	None if F and M constant; M for time varying F	None
Requirements regarding tag reporting rates	Known or data for estimating	Known or data for estimating	Not needed for Z_i, but external data required[3] for separating M_i and F_i

1. Variable definition in this and subsequent tables

P_0 = population size at the time of tagging (first tagging event for multiple release events);

Z_i = total mortality during time period i;

M_i = natural mortality during time period i;

F_i = fishing mortality during time period i;

N_0 = number of tags released for single release event approaches;

N_i = number of tags released at start of time period i for Brownie approach;

R_i = number of tags recovered during time period i;

C_i = number of animals examined for tags (e.g. catch) during time period i;

E_i = fishing effort during time period i;

q = catchability coefficient (assumed constant over time)

2. M_i can only be estimated for one less time period than release events (i.e., if there are 3 release events, then 2 M parameters can be estimated).

3. If reporting rates are assumed constant then no external data are needed but generally performance is poor.

Table 3: Summary of the integration of the Petersen approach with either of the two tag return approaches (i.e. tag-attrition or Brownie).

Integrated Approach	Petersen/Tag-Attrition	Petersen/Brownie
Parameters estimated	P_0, and: F and M if both constant; **or** F_i if M assumed known	P_0, Z_i, F_i, M_i [1]
Data requirements	N_0, R_i, C_i	N_i, R_i, C_i
Required release events	Single	Multiple (at least 3 consecutive time periods)
Required recapture periods	Multiple (at least 3)	Multiple (at least 3)
Required external parameters	None if F and M constant; M for time varying F	None
Requirements regarding tag reporting rates	Known or data for estimating	Not needed for P_0 and Z_i but external data needed[2] for separating M_i and F_i

1. See note 2 of Table 2.
2. See note 3 of Table 2.

Table 4: Summary of the two integrated approaches (Petersen/tag-attrition and Petersen/Brownie) when auxiliary information is incorporated.

Auxiliary information	Abundance index (A_i)		Effort data (E_i)	
Integrated approach	Petersen/Tag-Attrition	Petersen/Brownie	Petersen/Tag-Attrition	Petersen/Brownie
Parameters estimated	P_0, Z_i, F_i, constant M	P_0, Z_i, F_i, M_i (for all i, in contrast to Tables 2 & 3)	P_0, Z_i, F_i, constant M, q	P_0, Z_i, F_i, M_i (for all i, in contrast to Tables 2 & 3), q
Data requirements	N_0, R_i, C_i, A_i	N_i, R_i, C_i, A_i	N_0, R_i, C_i, E_i	N_i, R_i, C_i, E_i
Required release events	Single	Multiple (at least 3 consecutive time periods)	Single	Multiple (at least 3 consecutive time periods)
Required recapture periods	Multiple (at least 3)	Multiple (at least 3)	Multiple (at least 3)	
Required external parameters	None if M assumed constant	None	None if M assumed constant	None
Requirements regarding tag reporting rates	Known or data for estimating	Not needed for P_0 and Z_i but external data needed[1] for separating M_i and F_i	Known or data for estimating	Not needed for P_0 and Z_i but external data needed[1] for separating M_i and F_i
Additional assumptions (associated with the auxiliary data)	None	None	Relationship between F_i and E_i (e.g. constant catchability q)	Relationship between F_i and E_i (e.g. constant catchability q)

1. See note 2 for Table 2.

Appendix 4:

Southern bluefin tuna (*Thunnus maccoyii*) tag-recapture and catch data: details of data compilation

J. Paige Eveson, Tom Polacheck and Geoff M. Laslett

FRDC Project 2002/015

Introduction

This appendix provides a description of the southern bluefin tuna (SBT) (*Thunnus maccoyii*) tag-recapture and catch data used in the various analyses throughout this project, and gives details of how the data were compiled.

SBT are spawned in the northeast Indian Ocean between Indonesia and the northwest coast of Australia, generally between the months of September and April. The newly spawned fish migrate southward along the west coast of Australia, and then a large (but unknown) percentage of them travel eastward along the south coast of Australia to the Great Australian Bight (GAB). These fish spend the Austral summer in the GAB before leaving in the spring, either traveling east towards the Tasman Sea or west into the southeast Indian Ocean, with some traveling as far as South Africa. A large number of juveniles return to the GAB every summer, starting to arrive around November, with the majority having left again by April. It is not known what percent of juveniles of each age class migrate into the GAB each year or how much the percent varies between years, although it is known that the percentage reduces with age until almost no fish return after age 5. Moreover, it is not known whether fish that go into the GAB in one summer are more likely than those that did not to go into the GAB the next summer, or whether they have any fidelity to a particular winter location. The answers to these questions can be important in analyses of tag-recapture data for estimating mortality rates and/or abundance that require assumptions about mixing between tagged and untagged fish. They are also important if one wishes to incorporate a spatial component in any analysis of SBT data.

For most of the analyses presented in this report, we are only interested in tag-recapture and corresponding catch data for years 1991 to 1997. These are the years for which tag releases occurred during the 1990s tagging program, and 1997 is the last year for which we have reliable reporting rate estimates (which are needed in our analyses). Therefore, only the data for these years will be described.

Tag-recapture data

Extensive tagging experiments were conducted by CSIRO Marine Research from 1991 to 1997 in which juvenile SBT were caught, tagged, and released in the coastal waters off southern Western Australia, South Australia and southeastern Tasmania. Most of the tagged fish were initially caught using pole and line gear with a barbless hook, although a small number were caught with troll lines. After a fish had been hooked, it was hauled aboard the vessel and its nose to caudal fork length was measured. The fish was then tagged with two 12 cm plastic spaghetti dart tags, generally referred to as "conventional" tags. Tags were inserted into a fish about 4 cm to the rear of the second dorsal fin, one on each side of the fin. Tagging operations were designed to minimize handling time, and fish were re-released to the water within about 30 seconds of being brought on board. The tag numbers and length of each fish were recorded, together with the location and date of release. Additional information about the release was also recorded, such as the quality of tagging, the health of the fish, and the name of the tagger and the vessel. This information was later transferred to a computer database.

The age of a fish at the time of tagging was estimated based on its length using the growth curve currently adopted by the Commission for the Conservation of Southern

Bluefin Tuna (CCSBT) (Anon. 2001). Fortunately, SBT grow rapidly as juveniles so there is good separation between length distributions for the ages being tagged, and the number of aging errors should be small. All tagging was done between November and April, so the ages were adjusted in order that fish tagged in November or December from a given year-class/cohort were placed in the same age grouping as those tagged after December.

Recaptures occurred throughout the geographical distribution of SBT, ranging in longitude from 0 to 180°E and in latitude from 30 to 50°S. Not all tagged fish that were recaptured will have been reported, and for some fisheries non-reporting rates have been estimated to be substantial. Data used to estimate reporting rates come from observers placed in longline fisheries and from tag seeding experiments conducted in farms off of South Australia. These data are not sufficient in themselves, and various assumptions are also required. The method used to estimate reporting rates for SBT, and references where further details can be found, are given in Appendix 19.

Upon recapture, the finder measured the caudal fork length of the fish and recorded this length along with the tag number, the date and location of recapture, and sometimes the weight of the fish. This information along with the tags (either one or two) was sent to CSIRO and entered into the computer database, along with the name of the finder and the vessel and a judgment about the quality of the recapture information.

The number of tag returns and the return rate of tags over time provide information for estimating SBT mortality rates and abundance. However, some of the tag releases and recaptures are either unreliable or unsuitable for these purposes. Using the subsidiary information recorded upon release and recapture, we applied a rigorous screening process to the data.

To the release data we applied the following screening criteria:

- Only fish released into the wild were included. A few hundred fish were tagged and released into the South Australian tuna farms in 1997 and 1998 as part of a tag seeding program to estimate reporting rates, and these fish were excluded from analyses.
- Only releases where the fish was caught by pole and line were included. This method of catching fish is least likely to cause lasting injury to the fish. A small percentage (~1.5%) of tagged fish were caught by vessels using other gear types, mostly longline, and these releases were excluded.
- Only releases for which both tags were recorded as being inserted correctly were included to reduce the chance of tag shedding biasing our analyses. Only a small percentage (~1.8%) of releases had to be excluded on this basis.
- Only fish for which the injury due to tagging was regarded as slight were included to reduce the chance of fish mortality due to tagging biasing our analyses. Again only a small percentage (~1%) of releases had to be excluded on this basis
- Only fish whose length was recorded at the time of tagging were included. Most analyses conducted for this project required releases to be divided into age classes, and since release age is estimated based on length, those releases for which release length was not recorded had to be excluded. There were only 18 such releases.

To the recapture data we applied the following screening criteria:

- Only recaptures corresponding to releases that met the above release criteria were included.
- Only recapture records from fish caught in the wild were included. In the 1990s, tuna farming commenced and many tagged fish ended up being caught and put into farms. Tags were then returned on subsequent harvesting from the farm. The data base has two records for these fish: one corresponding to the original capture from the wild and one corresponding to the harvest from the farm. For the purposes of estimating fishing mortality, we are only interested in the information (date and location) of the capture from the wild.
- Only records corresponding to terminal recaptures (i.e. fish that were killed) or fish caught and released into the farms were included. A small percent of tagged fish were recaptured soon after release by the tagging vessel and re-released, and on the rare occasion, a tagged fish was recaptured by a fishermen and re-released; these records are omitted (in doing so, we are assuming that the fish's survival was not affected by being caught and re-released).
- Only tags recaptured by vessels with flags of Japan or Australia (i.e. Japanese or Australian quota) were included. This includes New Zealand and Australian joint venture catches (flag=Japan). For vessels with flags of other countries (i.e. catching quota for other countries), the reporting rates are assumed to be zero so we do not want to include any tags actually returned by these vessels (only 85 such tag returns).

The above release and recapture criteria were used in all non-spatial analyses of the 1990s SBT tagging data for estimating mortality rates and/or abundance (Appendices 5, 15 and 17) and were also used in the reporting rate analysis (Appendix 19). After screening, a total of 65047 releases from 1991 to 1997 were available for inclusion in these analyses. Corresponding to these releases, the total number of returns during the same time period that were suitable for our analysis was 6232. A summary of the release and recapture data after screening by year and age of release and year of recapture is provided in Table 1.

Catch data

There are several commercial fisheries that catch SBT, and these themselves can be comprised of several components. We identify the components as:

1. Australian domestic (includes surface, farm and longline catches)
2. Japanese longline
3. Japanese-Australian joint venture
4. Japanese-New Zealand joint venture
5. New Zealand domestic
6. Taiwanese longline and gillnet
7. Indonesian spawning ground fishery
8. Other (Korea, Philippines, South Africa, Singapore, and miscellaneous catches)

The level of available catch information differs considerably between components (e.g., the spatial and temporal grid at which data are provided; whether catch data are provided in weight or numbers; how much, if any, length sampling was done for age

estimation). As such, the method used to determine the number of fish caught and the age distribution of the catch differs between components and can be a complicated process. Post-processing of the raw catch data has been carried out by CSIRO to provide catch at age numbers by year for each of the above components for use in their 2001 and 2004 SBT stock assessments (Preece et al. 2001; Preece et al. 2004). The catch data from the 2004 assessments are used in the analyses presented throughout this report. The only differences are:

- Significant numbers of small SBT were caught and released by Japanese longline vessels in 1995 and 1996, and we have chosen to include the estimated non-surviving portion of the discarded catches in our catch data (Preece et al. 2001), whereas the data used in the 2004 assessments did not.

- The catch at age data for the 2004 assessments were compiled by calendar year (starting 1 January), whereas for our analysis we compiled the data by 'fishing' year, defined as starting 1 November[1], to be more consistent with the major fishing seasons for SBT. For all of the major fishery components (1 to 4), the catch data are available by month, so breaking them down by fishing year opposed to calendar year was straightforward. For the other components, which compromise a relatively small fraction of the total juvenile catch, the catch data were not available by month so we needed to use some approximations and assumptions. For the New Zealand domestic fishery, very few catches occur in November and December so using the data by calendar year was satisfactory. Fish caught in the Indonesian spawning ground fishery are essentially all greater than 7 years old, which is older than we are interested in for our analysis, so we do not need to worry about adjusting these catches. Lastly, catches from all remaining fisheries (components 6 and 8) are dominated by the Taiwanese catches, which generally occur in the austral winter, so we assume that no adjustments to year and age are necessary.

Total catch numbers by year and age are given in Table 2. Only catches from years 1991 to 1997 and ages 1 through 7 are provided since these are the only data used in the analyses in this report.

Compiling data for spatial analysis

For the spatial analysis of the 1990s SBT data (Appendix 16), we required that the tag-return data and the catch data be broken down not only by age and year but also by season and region (as defined below). In order to divide the data into regions, the location of release and recapture were required, so an additional condition was added to both the release and recapture criteria that the latitude and longitude fields could not be blank or recorded as unknown. Furthermore, the release condition that only releases from pole and line vessels should be included was relaxed to include releases from all vessels. All pole and line releases occurred in coastal waters south of Western Australia or in the GAB. For the spatial analysis we wanted releases from as diverse locations as possible; therefore, we included the small percent of longline releases that occurred mainly in the Tasman Sea but also in the southeast Indian Ocean.

---

[1] Explicitly, fishing year t is defined as the period from 1 November of calendar year $t-1$ to 31 October of calendar year t.

We defined 2 seasons as:
- Season 1: November – April (corresponding roughly to Australian surface fishery)
- Season 2: May – October (corresponding roughly to Japanese longline fishery)

We defined 4 regions as:
- Region 1: GAB (roughly corresponding to ≥120°E, <140°E, ≥25°S, <35°S)
- Region 2: Tasman (roughly corresponding to ≥120°E, <190°E, ≥25°S, <60°S minus the GAB)
- Region 3: SE Indian Ocean (SEIO) (roughly corresponding to ≥60°E, <120°E, ≥25°S, <60°S)
- Region 4: South Africa (SAfrica) (roughly corresponding to ≥−20°E, <60°E, ≥25°S, <60°S)

The spatial model that we propose and develop for SBT in Appendix 11 and apply in Appendix 16 is a simplified version of the truth in which we assume all fishing in the GAB occurs in season 1 (corresponding to the Australian surface fishery) and all fishing in the other 3 regions occurs in season 2 (corresponding to the major longline fisheries).

Note that the coordinates defining the regions are approximate because the regions are meant to correspond to fisheries, which do not adhere to strict geographical divisions (for example, some longline catches occur within the coordinates designated for the GAB but for our purposes should be classified as either Tasman or SEIO accordingly).

Releases were initially divided into regions using the above latitude and longitude definitions. These divisions were then modified as follows:
- all non-longline releases defined as SEIO were changed to GAB (these are Western Australia releases of very young, predominantly age 1, SBT)
- all non-longline releases defined as Tasman that were <140°E and <40°S were changed to GAB (there were some releases south of our 35°S boundary for the GAB that should clearly be included in the GAB)

Figure 1 shows the distribution of releases by region.

Similarly, recaptures were initially divided into regions using the above latitude and longitude. These divisions were then modified as follows:
- any non-longline recaptures defined as SEIO were changed to GAB
- any non-longline recaptures defined as Tasman that were <140°E and <40°S were changed to GAB
- any longline recaptures defined as GAB were changed to Tasman

Figure 2 shows the distribution of recaptures by region.

To divide the releases and recaptures into seasons, we initially used the season definitions given above. We then modified the season for both the releases and recaptures to adhere to our simplified model as follows:
- any recaptures that occurred in the GAB in May through October were defined as season 1 (there were no such releases and <3% such recaptures)

- any releases or recaptures that occurred in SAfrica, SEIO or Tasman in November through April were defined as season 2 (there were only 7 such releases and <9% such recaptures)

Note that changing the season definition sometimes required a change to the fishing year as well. For example, if a fish was caught in November or December of *fishing year* 1993 (calendar year 1992) in the Tasman, then we changed it from being caught in season 1 of fishing year 1993 to being caught in season 2 of fishing year 1992. Similarly, if a fish was caught in July through October of 1993 in the GAB, then we changed it from being caught in season 2 of fishing year 1993 to being caught in season 1 of fishing year 1994.

Dividing the catch data into seasons and regions was slightly more complicated because the information available for doing so varied between fishery components. Furthermore, the catch data do not correspond as nicely as the tag-recapture data to our simplified model.

To divide the catch data into seasons and regions, we used the following procedure for each fishery component:

- For the Australian domestic fishery, we do not have latitude and longitude information available to us, but instead we have general areas designations, the relevant ones being New South Wales, Tasman, South Australia, Western Australia, Albany and Esperance. Catches from New South Wales and Tasman were assigned to region Tasman, and catches from the remaining areas were assigned to region GAB. Month information is available so we initially divided the catches into seasons according to our season definitions, then modified the season (and fishing year when necessary) to adhere to our simplified model in the same way we did for the releases and recaptures. This required a fairly large percentage of changes (~25%) since, in actuality, a fair number of fish are caught in the GAB outside of season 1 and in the Tasman outside of season 2.

- The Japanese longline catches could be divided into regions based on latitude and longitude information (none were in the GAB). There were a few catches outside of our region definitions, and they were classified as SAfrica, SEIO or Tasman accordingly (based on proximity and judgment). Similarly, the catches could be divided into seasons based on month information, then we modified the season (and fishing year if necessary) to adhere to our simplified model (this required ~13% changes from season 1 to season 2; no changes were required in the other direction since there were no GAB catches).

- The Australian joint venture catches were divided into regions and seasons in the same manner as the Japanese longline catches (again with a few catches outside of our region definitions that were classified as SEIO or Tasman accordingly). No catches occurred in region SAfrica. The modification of the season definitions required almost no changes to the GAB catches, but ~19% changes from season 1 to season 2 for the other regions.

- The New Zealand joint venture catches were also divided into regions using latitude and longitude information. All catches belonged to the Tasman (no modifications were necessary). As such, all catches were defined as belonging to season 2 (~12% of the catches would have been classified as season 1 based on month).

- The New Zealand domestic catches all occurred in the Tasman, therefore we assigned all catches to season 2. Only a very small percent of catches would have been assigned to season 1 based on month.
- For the Taiwanese fishery, we attributed 75% of the catches to SEIO and the remaining 25% to SAfrica, all in season 2. This is based on very limited information available from country reports provided to the CCSBT but should be reasonable for our purposes.
- The Indonesian catches are outside the four regions we defined, but this is not of concern because fish caught on the spawning ground are older than those we are interested in for our study and can simply be omitted.
- For all other catches, we have limited information available. We divided the catches by age in a particular year into region – either SEIO or SAfrica – in proportion to the Japanese longline catch at age in these two regions in the same year. All catches were assumed to belong to season 2.

References

Anon. 2001. Report of the fifth meeting of the scientific committee. Commission for the Conservation of Southern Bluefin Tuna. 19-24 March 2001. Tokyo, Japan.

Preece, A., Polacheck, T., Kolody, D., Eveson, P., Ricard, D., Jumppanen, P., Farley, J., and Davis, T. 2001. Summary of the primary data inputs to CSIRO's 2001 stock assessment models. Commission for the Conservation of Southern Bluefin Tuna CCSBT-SC/0108/21.

Preece, A., Cooper, S., and Hartog, J. 2004. Data post-processing for input to the 2004 stock assessments and comparisons of 2001 and 2004 assessment datasets. Commission for the Conservation of Southern Bluefin Tuna. CCSBT-ESC/0409/27.

Table 1. Summary of the southern bluefin tuna tag release and return data from 1991 to 1997 (after data screening).

Release year	Release age	Number releases	Number returns by year						
			1991	1992	1993	1994	1995	1996	1997
1991	1	3299	20	40	46	23	13	5	4
1991	2	3127	103	148	59	34	20	7	5
1991	3	810	63	8	16	7	1	5	1
1991	4	1	0	0	0	0	0	0	0
1992	1	2144		1	21	56	37	11	7
1992	2	4646		88	159	101	33	12	8
1992	3	1097		57	18	11	9	3	2
1992	4	1		0	0	0	0	0	0
1993	1	4898			2	41	201	91	58
1993	2	2937			60	68	67	21	11
1993	3	2777			66	78	32	17	15
1993	4	22			2	0	0	1	0
1994	1	9003				4	110	401	364
1994	2	3158				29	167	76	52
1994	3	3640				77	145	30	40
1994	4	111				4	2	0	0
1994	5	4				0	0	0	0
1995	1	8585					0	87	622
1995	2	5899					83	395	363
1995	3	2629					55	103	74
1995	4	101					1	3	1
1995	5	3					0	0	0
1996	1	82						0	3
1996	2	2518						77	339
1996	3	1511						115	201
1996	4	24						1	1
1996	5	1						0	0
1997	1	884							1
1997	2	592							15
1997	3	526							91
1997	4	17							7

Table 2. Southern bluefin tuna total catch at age data from 1991 to 1997 for ages 1 to 7.

Year	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7
1991	48450	76744	176057	51693	21726	12332	9491
1992	7624	33638	150758	77731	32981	16426	8558
1993	404	38414	120232	65802	48640	31324	17285
1994	187	10398	119166	72806	32144	24928	16435
1995	416	30789	133300	61080	39073	27442	20560
1996	422	26276	171859	76136	38646	24743	18972
1997	1965	32025	203883	72177	43001	27398	21673

Figure 1. SBT releases from 1991 to 1997 by region.
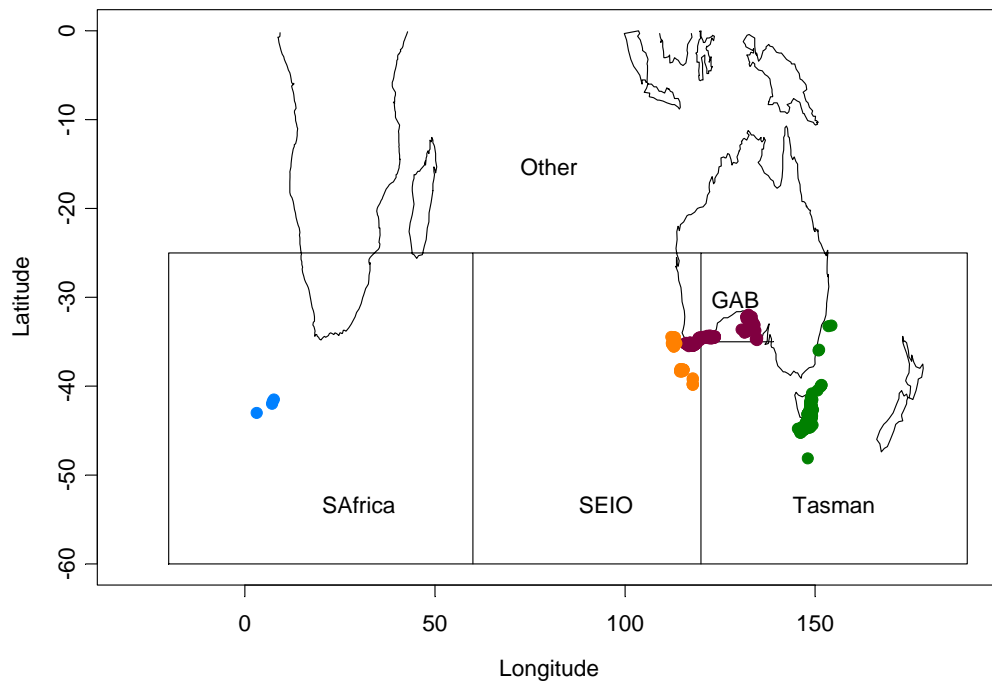
Figure 2. SBT recaptures from 1991 to 1997 by region.

A5-1
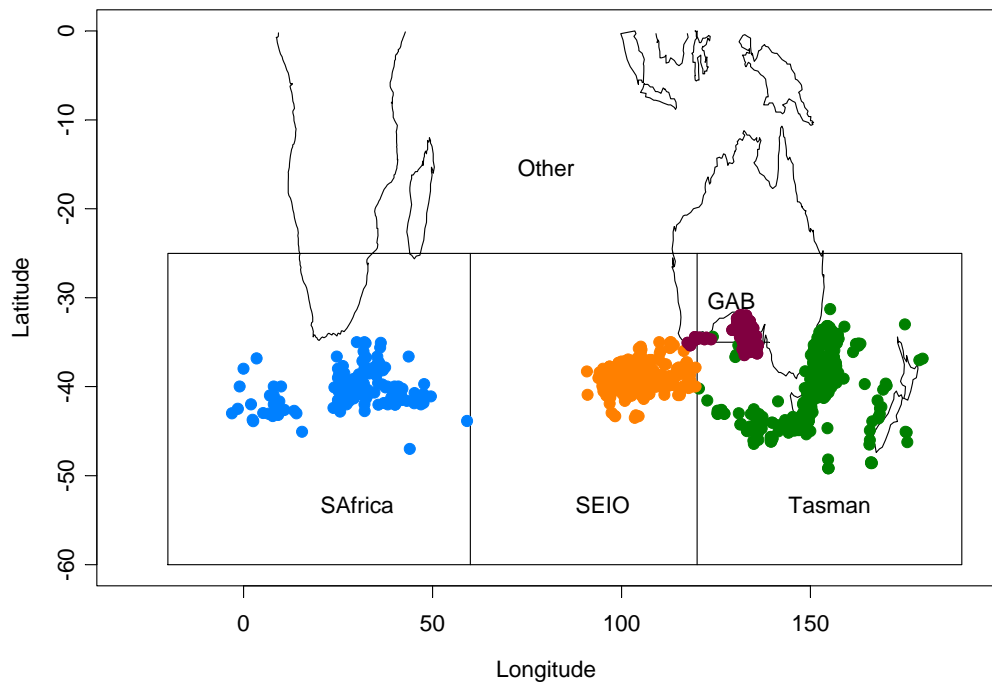
Appendix 5:

An integrated Brownie and Petersen model for estimating mortality rates and population size in a fisheries context with known reporting rates

Tom Polacheck, J. Paige Eveson, Geoff M. Laslett, Ken H. Pollock, William S. Hearn and John M. Hoenig

FRDC Project 2002/015

Abstract

A comprehensive framework for modelling data from multi-year tagging experiments in a fishery context is presented that incorporates catch data into the traditional Brownie tag-recapture model. Incorporation of catch data not only allows for improved estimation of natural and fishing mortality rates (especially fishing) but also allows for direct estimation of population size at the time of tagging. These are the primary quantities required to be estimated in stock assessments; having an approach for directly estimating them that does not require catch rates or data from fishery-independent surveys (which are often unfeasible and provide insufficient data) provides a potentially powerful alternative for augmenting traditional stock assessment methods. Simulations are used to demonstrate the value of directly incorporating catch data into the estimation framework and also to illustrate the trade-off between number of releases and accuracy of catch data in terms of parameter estimation. Results from the simulation scenarios considered suggest that including catch data can improve the precision of fishing mortality rate estimates by up to ~40% and natural mortality rate estimates by up to ~10%. They also show that reducing uncertainty in the catch data can lead to significant improvements in the precision of the population size estimate and the fishing mortality rate estimates (especially at young ages and low numbers of releases), whereas increasing the number of releases can lead to large improvements in the natural mortality rate estimates as well as the fishing mortality rate estimates (especially when the uncertainty in the catch data is high). Finally, the model is applied to southern bluefin tuna (*Thunnus maccoyii*) tag-recapture and catch data collected in the 1990s to provide estimates of natural mortality, fishing mortality and abundance for five cohorts of fish.

Introduction

Mark-recapture experiments represent one of the primary tools for studying the dynamics of wild animal populations, and an extensive literature exists documenting the wide variety of experimental designs and estimation models available (e.g. Seber 1982; Quinn and Deriso 1999; Morgan and Thomson 2002). The attraction of mark-recapture studies is that they have the potential to provide estimates of population abundance and mortality rates (both natural and exploitation). These are fundamental quantities needed for understanding population biology and providing scientific-based management advice, but are extremely difficult to measure directly for most wild populations because conducting a direct census and/or tracking a large number of individuals is often not feasible. There are two basic types of mark-recapture studies – those in which an animal can be captured and released multiple times and those in which only a single recapture is possible because the animal dies in the capture process. In most fishery applications, only single recapture studies are feasible, and it is this type of study considered here.

In fisheries, natural mortality and fishing mortality rates are critical components of the stock assessment process because they form a key component in evaluating stock productivity and density dependent responses of a population; thus, uncertainty in their estimated values can be a major source of uncertainty in the resulting management advice. Direct estimation of fishing and mortality rates has generally been an intractable problem in marine fish populations. Most assessment methods assume that natural mortality is known and is constant with age and time, and values used come from *ad hoc* approaches (e.g. catch curve, life history characteristics, analogy from other stocks). Similarly, the most common approach for estimating fishing mortality rates (e.g. virtual population analysis and related catch-at-age approaches) are dependent upon assumptions about selectivity and require extensive auxiliary relative abundance indices.

Historically, the most common approaches for analysing mark-recapture data from a single release event were based on a Petersen-type model (e.g. Seber 1982). In this case, the primary quantity being estimated is the population size at the beginning of

the experiment. Population size is estimated based on the ratio of the observed number of tags returned within samples taken from the population given the known number of tags released into the population. The Petersen method was developed for use in closed populations (i.e. populations that have no births, deaths, immigration or emigration); however, the method can be applied to give an estimate of abundance at the time of tagging even when there are deaths and emigration, provided these factors apply equally to tagged and untagged animals. While the Petersen approach has been used extensively in wildlife and controlled experimental situations, it has not been widely used in large commercial fisheries because the size of the sample actually examined for tags is often difficult to ascertain, and will almost certainly include error. In addition, if all recaptured tags are not reported, then Petersen models require that reporting rates are either known or estimable from other sources.

A different approach to the Petersen one is the use of multi-year tag release and recovery experiments (Brownie et al. 1985). Brownie models for multi-year tagging data provide estimates of total mortality rates from comparison of the return rates over time from the multiple releases. The power of the approach is that only data on the number of releases and returns (usually by cohort) are required to estimate mortality rates (i.e. there is no need to know the size of the sample examined for tags). Brownie models can provide estimates of total mortality rates even when reporting rates are unknown; however, the separation of natural mortality rates from exploitation rates requires that reporting rates are either known or estimable. In theory, reporting rates can be estimated directly from multi-year tag-return data if they are assumed to be constant over years, but the precision of the estimates is usually unsatisfactory (Hoenig et al. 1998a) and, moreover, reporting rates will vary by year in many situations. Thus additional information is usually required for obtaining reasonable reporting rate estimates.

In recent years, the application of multi-year tagging experiments for estimating mortality rates using general models (Brownie et al. 1985) has been recognized as a powerful approach that can be applied in fishery situations to provide direct estimates of both natural and fishing mortality rates (e.g. Pollock et al. 1991; Polacheck et al. 1996, 1997). A number of papers have further developed these models for application

in particular fishery situations (Hoenig et al. 1998a,b; Hearn et al. 1999; Pollock et al. 2002). In addition, estimates of mortality rates from multi-year tagging programs using a Brownie framework have been incorporated into stock assessments, and the approach underlies the design of a large-scale international tagging program for southern bluefin tuna currently in progress (Anon. 2001a, 2001c, 2002).

The Petersen and Brownie approaches rely on different information from mark-recapture data to estimate fundamentally different quantities (absolute abundance versus mortality rates). In particular, the Petersen method takes advantage of the information content in the proportion of recaptured animals that have a tag, while the Brownie approach uses the information content on the rate at which tags are returned. In the current appendix, we develop a mark-recapture model that combines these two approaches by using both sources of information to provide joint estimates of both abundance and mortality rates. We do this in a fishery context in which catch data are used to provide estimates of the number of animals examined for tags and the uncertainty in these catch statistics is incorporated explicitly (the approach is equally valid in a more controlled experimental situation where exact information is available on the number of animals examined for tags). We demonstrate that there is a synergy in combining the two approaches in that the addition of catch data to the Brownie model not only allows for estimation of abundance but also improves estimation of mortality rates. We also explore the relative trade-off between increased tagging efforts (through more tag releases) and improved catch information (for example, through increased sampling) in terms of accuracy of parameter estimates. Finally, we apply the combined model to southern bluefin tuna tag-recapture and catch data to provide mortality rate and abundance estimates for five cohorts of fish from the early 1990s.

Material and Methods

Basic Dynamic Model

The general population dynamics equations commonly used in fisheries form the basic model underlying the analysis of multi-year tagging experiments used here. These equations involve exponential and competing natural and fishing mortality

rates. The model can be easily translated into alternative formulations for the mortality dynamics, such as a net natural survival rate and annual harvest rates from a pulse fishery (see Hoenig et al. 1998a). For a cohort of animals of a given age, the deterministic equations for the number of animals that survive to the next age and the number caught are (e.g. Hilborn and Walters 1992, p. 378):

$$P_{i+1,t+1} = P_{i,t} \exp\left(-F_{i,t} - M_{i,t}\right) \tag{1}$$

$$C_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} P_{i,t} \left(1 - \exp\left(-F_{i,t} - M_{i,t}\right)\right) \tag{2}$$

where

$P_{i,t}$ = the number of individuals of age i at time t

$C_{i,t}$ = the catch of individuals of age i at time t

$F_{i,t}$ = the instantaneous fishing mortality rate for individuals of age i at time t

$M_{i,t}$ = the instantaneous natural mortality rate for individuals of age i at time t.

In many contexts, $M_{i,t}$ is assumed to be constant with time, although multi-year and multi-cohort tagging programs can provide year- and age-specific natural mortality rates to a limited extent (as discussed later). In the current appendix we focus on a multi-year tagging experiment involving a single cohort, so we can simplify the notation by dropping the t subscript and expressing everything in terms of age. The generalization of all equations to more than one cohort is straightforward. Later we will discuss some advantages of extending a tagging experiment to multiple cohorts, and we will also apply the integrated tag and catch model to multiple cohorts of southern bluefin tuna data.

In the context of a tagging experiment, the above equations provide the basis for predicting the expected number of returns, assuming that the tagged fish constitute a representative sample of the population. Let N_a be the number of tag releases of age a fish from a specific cohort and λ_i be the tag reporting rate for fish captured at age i. Define:

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

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

where S_i represents the annual survival rate of age i fish and u_i represents the annual exploitation rate of age i fish. Then, following Brownie et al. (1985), the expected numbers of tags recaptured and returned from a particular cohort at age i from releases at age a are given by the expressions in Table 1. These expressions for the expected number of returns assume complete and instantaneous mixing of tagged fish with the untagged population and no tagging mortality or tag shedding. If these assumptions are not met, additional parameters and potentially additional data will need to be introduced to account for them. Failing to do so will likely lead to biased parameter estimates and overly optimistic estimates of the precision of the parameter estimates (see Discussion).

Equations (1) and (2) can also be used to provide analogous expressions for the expected catches of age i fish from a particular cohort, conditional on the size of the cohort at age 1, P_1 (Table 2). Conceptually, the catch data can be viewed as a tagging experiment in which the number of "releases" (P_1) is an unknown parameter to be estimated. However, unlike a tagging experiment where the number of returns from a particular release event are essentially known quantities, the numbers of fish caught at each age will be estimated quantities, usually derived from a multi-stage sampling of catches for length combined with age/length keys derived from direct aging of hard parts (otoliths in the case of SBT). Because P_1 is unknown, it is not possible from the catch-at-age data alone to derive estimates of the mortality rates[1]. Combining the catch-at-age data with the multi-year tagging data allows for P_1 to be estimated and for additional information on the F's and M's contained in the catch data to be extracted.

---

[1] Even if M is assumed known as in many stock assessments, there are still too many parameters unless F is assumed constant for a cohort; this is the reason that catch-at-age stock assessment models require additional sources of data for "tuning" (see Hilborn and Walters 1992).

Estimation Model

We use a maximum likelihood approach to jointly model the tag-recapture and catch data. We develop independent likelihoods for the tagging data and the catch data that can be multiplied together to give an overall likelihood function. This likelihood function can then be maximized to give estimates of the unknown F, M and P parameters.

We consider the case in which tag reporting rates are assumed to be known. While this is likely an unrealistic assumption in most situations, it provides a straightforward way to examine the potential gain achieved by combining tagging and catch-at-age data. If reporting rates are not known exactly, then they need to be estimated, generally using auxiliary data. This is true for any application of the Petersen model for estimating abundance or the Brownie model for estimating natural and fishing mortality rates. We have not attempted to deal explicitly with the estimation of reporting rates in the current appendix because there are a number of approaches that can be used to obtain such estimates (e.g. tag seeding, high-reward tags, observers monitoring a portion of catches, tagging twice per year) and how to account for the uncertainty induced by estimating reporting rates depends on the method used. For some methods, the estimates can be considered independent of the tagging and related catch data (e.g. tag seeding data). In such cases, a likelihood can be developed for the data used to estimate the reporting rates and simply multiplied by the tagging and catch likelihoods developed below. In other cases, the estimates of reporting rates will use some of the tagging and related catch data (e.g. when based on observers) and the likelihoods for the tagging and catch data would need to be modified (see Appendix 7).

As developed in Brownie et al. (1985), if each tag return is assumed to be independent, then the numbers of returns at age (including those not returned) from any individual release event (i.e. releases for a particular age) are expected to be multinomial, and the likelihood function for the observed numbers of returns from all release events is the product of multinomials:

$$L_R = \text{K} \times \prod_a \left(\left(\prod_{i \geq a} p_{a,i}^{R_{a,i}} \right) \left(1 - p_{a,\bullet} \right)^{N_a - R_{a,\bullet}} \right) \qquad (3)$$

$$\text{K} = \prod_a \frac{N_a!}{\prod_{i \geq a} R_{a,i}! (N_a - R_{a,\bullet})!}$$

where a indexes release age, i indexes recapture age, $R_{a,i}$ is the number of tags returned from an age i fish released at age a, and $p_{a,i}$ is the probability of a tag being returned from an age i fish released at age a. An expression for $p_{a,i}$ can be obtained from the expected number of returns in Table 1 by dividing by N_i. Explicitly,

$$p_{a,i} = \begin{cases} u_i \lambda_i & i = a \\ S_a \cdots S_{i-1} u_i \lambda_i & i > a \end{cases}$$

Note that in equation (3) a dot in the subscript denotes summation over the index it replaces. Also, K is a constant that can be omitted from the likelihood.

Similarly, if we assume the numbers of fish caught at each age are known accurately (and that each fish has an equal probability of being caught), then the catch-at-age data, including those fish from the cohort not caught, are random multinomial, where each fish has a probability of being captured at age i or not captured. Usually, however, the catch-at-age data are not known accurately. For many fisheries, including that for southern bluefin tuna, the age distribution of the catch is determined by taking a sample, estimating the ages of fish in the sample (either from lengths or from direct aging of hard parts), and using the estimated age frequencies of the sample to represent the total catch. We have chosen to model the error in the catch-at-age data that results from such a sampling procedure as Gaussian with a coefficient of variation (CV) that depends on the level of sampling. The CV is intended to capture variability in the catch-at-age data due to non-homogeneous spatial and temporal distribution of fish, as well as different size/age selectivities among vessels (i.e. if these factors are significant, then the CV of the catch-at-age data would be large because the age distributions derived from different samples could vary greatly).

To fit a model with both multinomial "process" error and Gaussian sampling error would require a relatively sophisticated approach, such as a Kalman filter. In most

commercial fishery situations, the number of fish in the cohort from which catches are being taken will be very large such that the multinomial error will be negligible compared to the Gaussian sampling error, and only the latter source of error needs to be considered. This is the approach taken in the current appendix.

The actual error structure in any particular application will depend upon the details of the sampling and age estimation procedures. The choice of a Gaussian error structure provides a general framework for exploring the effects of measurement error in the catch estimates, but other error structures for the catch data that accommodate a particular sampling regime can easily be incorporated as long as an appropriate likelihood function can be developed. We have also assumed that the errors in the catch-at-age estimates between years (i.e. ages for a single cohort) are independent. This is a reasonable assumption for many fisheries, in which sampling and aging data are collected each year. However, in some situations (particularly where age at length is being estimated from a growth curve), covariance in the estimates between years may exist and would need to be accounted for.

Assuming a Gaussian error structure, the likelihood for the catch-at-age data can be expressed as:

$$L_C = \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\left(\frac{C_i - E(C_i)}{\sigma_i} \right)^2 \right) \tag{4}$$

where C_i is the number of fish caught at age i, $E(C_i)$ is the expected catch at age i (as given in Table 2), and σ_i^2 is the variance of the catch at age i.

The overall likelihood for the combined recapture and catch data can be obtained by multiplying likelihoods (3) and (4) together:

$$L = L_R \times L_C \tag{5}$$

It is legitimate to multiply the two likelihoods together provided the tag-recapture data and the catch data are independent. This assumption may not be met if during the

process of sampling the catch data, tags are found and returned; however, we assume that tags are removed at the time of catch, prior to catch sampling. Moreover, if the catch sample is relatively small, then the expected number of tags in the sample will be so small that the independence assumption will not be seriously violated.

Estimates of the unknown model parameters can be obtained by maximizing the likelihood in (5); however, not all parameters introduced thus far are estimable. First, the 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 $i+1$. Thus, in an experiment with n consecutive release years, estimates can only be obtained for $n-1$ of the natural mortality rate parameters (regardless of the number of recapture years). We address this issue by assuming that $M_i = M_{n-1}$ for $i \geq n$, but other constraints could be used, such as imposing a parametric relationship between natural mortality and age. Second, the inclusion of the catch likelihood (4) results in the addition of one more parameter than data points. For example, three years of catch data result in four extra parameters (P_1 and three σ_i's) but only three additional data points (three C_i's). One way to deal with this issue is to assume that the σ_i's are known. In practice, estimates of the catch-at-age are generally derived from a sampling program, and therefore independent information should be available for estimating the σ_i's (and could possibly be included in the model through another likelihood component). Another way to deal with this issue is to put constraints on the variance parameters; however, it transpires that even if σ_i is constrained to be equal at all ages (i.e. $\sigma_i = \sigma$ for all i), there are still difficulties with estimating σ. On closer inspection, we see that when we allow for a separate fishing mortality for each age of recapture, the catch likelihood (4) can always be maximized by allowing σ to approach zero. This is because F_i's can be found that give a perfect fit to the catch-at-age data (i.e. such that $E(C_i) = C_i$ for all i), in which case the exponential term in the catch likelihood equals one (provided $\sigma > 0$) and the catch likelihood approaches infinity as σ approaches zero. Even if the F_i's that fit the catch data perfectly fit the recapture data very poorly, the catch component of the likelihood will still dominate over the tag-recapture likelihood when σ is sufficiently close to zero. In situations where the tagging data are relatively influential compared to the catch data (such as when the

number of tag releases is large and the variability in the catch data is large), there will often be a local maximum away from zero (we see an example of this later when we apply the model to southern bluefin tuna data).   In such situations, a reasonable estimate of σ can be obtained by setting a lower bound on σ sufficiently larger than zero. Note, however, that simulations suggest that even when a local maximum exists, the local estimate of σ tends to have a negative bias. Generating a likelihood profile for σ is a good way to determine when a local maximum exists.

In the simulations results presented below, we have chosen to model the catches as having a constant coefficient of variation υ (i.e. $\sigma_i = \upsilon E(C_i)$). This formulation is convenient for evaluating the trade-off between devoting resources to tagging (through the number of tags released) and devoting resources to catch sampling (assuming that υ will decrease as sampling increases). We assume that the true value of υ is known in order to avoid problems with estimating it (which are analogous to the problems just discussed with estimating a constant σ).

Results

In the previous section, the model was presented in terms of an overall likelihood (given in (5)) and parameter estimation was referred to in terms of maximizing this likelihood. Conceptually, this seemed simplest to interpret; however, in actually fitting the model, we used the standard procedure of obtaining parameter estimates by minimizing the negative log of the likelihood. Therefore, throughout the results section we refer to and present results for the negative log-likelihood.

Simulation Results

Value of incorporating catch data

To investigate the value of incorporating catch-at-age data in terms of the accuracy of the parameter estimates, we simulated multinomial tag-recapture data and Gaussian catch-at-age data. We then compared the parameter estimates obtained using just the tag-recapture data with those obtained using both the tag-recapture data and the catch data.

We generated tagging data for three consecutive release years (tagging the same cohort each year) and three recapture years using the following values:

$$N_a = 250 \quad a = 1,\ldots,3$$
$$F_i = 0.15 \quad i = 1,\ldots,3$$
$$M_i = 0.2 \quad i = 1,\ldots,3$$
$$\lambda_i = 1.0 \quad i = 1,\ldots,3$$

Gaussian catch data were generated corresponding to each recapture age using the above F and M values, an initial population size of $P_1 = 100000$ and a coefficient of variation of $\upsilon = 0.1$.

We generated 500 tag-recapture data sets and corresponding catch-at-age data sets. Parameter estimates were obtained first using just the tagging data to minimize the negative log of likelihood (3), and second using both the tagging and catch data to minimize the negative log of likelihood (5). Although we used constant mortality rates across ages to generate the data, we estimate a separate fishing mortality rate for each age (F_1, F_2 and F_3) and we estimate a natural mortality rate for age 1 (M_1) and age 2 and above (M_2) – recall this is the most we can estimate with three release years. When we use the likelihood that includes catch data, we also get a direct estimate of the initial size of the cohort, P_1. The mean, standard error (SE) and coefficient of variation (CV) of the parameter estimates are summarized in Table 3.

All of the mortality rate estimates were more precise when the catch data were incorporated, with a 40% reduction in the SE for the F_1 estimates, about a 30% reduction for F_2, and about a 10% reduction for M_1 and M_2. Some improvement in precision is to be expected because the catch data constitute additional data. However, the improvement can be substantial, particularly in light of the fact that catch-at-age data are often available without any additional expense (improvements would be expected with more tagging data as well but would come at more cost). Integrating catch data into the model not only improves the precision of the mortality rate estimates, it also provides a direct estimate of the initial population size, P_1, which

cannot be obtained from the tag-recapture data alone. In our simulations, P_1 was estimated quite accurately, with a CV of about 10% (Table 3).

We note that all of the mean parameter estimates obtained using the model with only tagging data are slightly greater than two SE's above their true value, suggesting they have a small positive bias. When the model with both tagging and catch data is used, all of these biases, except that for F_3, disappear. It is important to keep in mind that all simulations results are random realizations and that repeating the simulations will lead to slightly different results. For example, small biases that appear to be significant in one set of simulations can become insignificant in a different set of simulations, and vice versa (e.g. compare results in Table 3 with Figure 2a). Increasing the number of runs will reduce the amount of variation between two sets of simulations, but 500 runs were adequate to give consistent results with respect to key features, including these small bias effects.

Trade-off between number of releases and accuracy of catch data

We have established that incorporating catch data can improve the precision of the parameter estimates. The degree of improvement will depend on the amount of variability in the catch data as well as the number of tags released. For designing a tagging experiment, it would be very useful to know whether resources would be better spent on tagging large numbers of fish or on reducing the uncertainty in the catch-at-age data (through more port sampling, more on-board observers, collection of otoliths for aging, et cetera).

To address this question, we carried out simulations in which we varied the number of releases (N) from 250 to 2000 and the catch CV (υ) from 0.05 to 0.50. For each value of N we generated 500 tag-recapture data sets, and corresponding to each tag-recapture data set, we generated a catch data set for each value of υ. We assumed three consecutive release years with an equal number of releases in each year, three recapture years, and the same values for the F's, M's and P_1 as in our previous simulations. We estimated the mortality rate and initial population size parameters by minimizing the negative log of the joint tagging and catch likelihood (5), and we also estimated the mortality rate parameters by minimizing the negative log of the tagging likelihood (3) only.

We chose to evaluate the parameter estimates using the root mean squared error (RMSE) because this statistic is a combined measure of both the bias and the precision of the estimates. As the name suggests, RMSE is calculated by computing the mean of the squared differences between the parameter estimates and the true value (which is equivalent to computing the variance plus the average bias squared), and taking the square root of this value. The results are summarized in Figure 1a-f.

First concentrate on the estimates of the fishing mortality rates (Figure 1a-c). Increasing the number of releases resulted in decreases in the RMSE of all the fishing mortality rate estimates (F_1 to F_3), with the rate of decrease slowing considerably after about 1000 releases. Reducing the variability in the catch data (i.e. decreasing υ) also reduced the RMSE of the fishing mortality rate estimates; however, the response lessened as the age of the fish increased such that by age 3, the RMSE of the F_3 estimates was essentially unaffected by the variability in the catch data. There was also a clear interaction between the number of releases and the variability of the catch in the estimation of F_1 and, to a lesser extent, F_2. In particular, for these two parameters, the gain from increasing the number of releases became greater as the catch CV (υ) increased; i.e., releasing more tags improved the parameter estimates a lot more when the catch data were not very informative than when the catch data were known precisely.

Similar to the fishing mortality estimates, increasing the number of releases resulted in decreases in the RMSE of the natural mortality rate estimates (M_1 and M_2), with the rate of decrease slowing as the number of releases increased (Figure 1d-e). However, the natural morality rate estimates showed much less response to changes in the CV of the catch data.

For the initial population size (P_1), the variability in the catch data had a large influence on the precision of the P_1 estimate (Figure 1f). Decreasing the CV of the catch data resulted in an almost linear decrease in the RMSE of P_1 over the range of catch CV values considered. On the contrary, the number of releases had relatively little effect on the estimation of P_1 (especially at higher catch CV's). At first this

seems counter-intuitive because fishing and natural mortality rates were estimated quite poorly at low release numbers. However, it is the probability of catching a fish that is important for estimating P_1, not the fishing and natural mortality rates per se, and this probability can be estimated well even if fishing and natural mortality are not. For example, if natural mortality is overestimated then the probability of catching a fish will be underestimated, but if fishing mortality is also overestimated then the probability of catching a fish will be overestimated; the two counteract each other such that the probability of catching a fish is estimated without bias. A similar argument holds if natural and fishing mortality are both underestimated. We can expect natural and fishing mortality to both be overestimated or both be underestimated in many cases because positive correlations exist between these parameters (see Discussion).

When designing an actual tagging experiment and considering the trade-offs between tagging more fish and improving the catch data, the relative cost of the two data sources would need to be taken into account. These costs will be fishery specific and we have not attempted to incorporate them in our model, but the results presented here should still provide guidance on determining a reasonable balance.

We note that in our simulations there were only a few cases in which small biases were present in some of the parameter estimates, and even in these cases, the contribution of the bias to the RMSE was negligible compared to the variance. Nevertheless, the fact that biases were present at all was somewhat surprising and led us to investigate further.

Biases in parameter estimates

In the section on the value of incorporating catch data, we saw evidence of some small biases in the parameter estimates obtained using only tagging data. We also saw that incorporating the catch data reduced these biases. These results were obtained using 250 tag releases in each year. The biases disappeared when the number of releases each year was increased to 500 or more, even for the model with just tagging data (Figure 2a). Thus, in terms of biases of the mortality rate estimates, the number of tag releases was much more influential than the incorporation of catch

data. In any case, the biases were relatively small, even with small numbers of releases.

Although the biases in the fishing mortality estimates were small, they appeared to increase with age (Table 3; Figure 2a). Thus, we ran simulations with two additional recapture years, keeping the number of release years at three, and we found that the pattern persisted. Namely, the bias in the fishing mortality rates increased exponentially with age, and was non-trivial by age 5; for example, with only 250 releases in each year, the bias in the mean age 5 fishing mortality rate estimate was over 20% (Figure 2b). The biases again decreased as the number of releases increased; however, the biases in the age 4 and especially age 5 fishing mortality rate estimates did not disappear as quickly. In fact, there was still evidence of a small positive bias in the age 5 estimate with 5000 releases per year. The bias results were similar when the model with both tagging and catch data was used.

Although the reason for these biases is not completely understood, they stem from the fact that the natural mortality and fishing mortality rate estimates at close ages are highly correlated combined with the fact that the natural mortality rates are constrained to be the same for ages 2 and above. Recall that this constraint was required because with three release years, the model can only provide estimates of two natural mortality rate estimates, regardless of the number of recapture years. If we fix the natural mortality rates for ages 3 and above at their true values and re-run the simulations, then the biases in the fishing mortality rate estimates generally disappear (for any number of recapture years).

In the model with both tagging and catch data, the initial population size tended to be estimated with a small negative bias. This bias increased as the CV of the catch data increased, but unlike the biases in the mortality rate estimates, did not diminish as the number of releases increased (Figure 3). It can be shown theoretically that this bias is expected, and that it is caused by modelling the variance of the catch data in terms of a known CV and thus as a function of the mean (Appendix 6). In any case, the bias was always small enough that it is not of concern (~5% for a catch CV of 0.5, and less than 2% for catch CV's of 0.2 and below).

Effect of mortality rate values on simulation results

To investigate how the values chosen for *M* and *F* affect the results, we ran additional simulations using the model with both tagging and catch data in which we varied these values, keeping all other parameters constant. We assumed the same experimental design and parameter values as in the section on the value of incorporating catch data, except the levels of natural mortality and fishing mortality were varied (i.e. 3 release years, 3 recapture years, $N_a = 250$, $\lambda_i = 1.0$, $P_1 = 100000$, $\upsilon_i = 0.1$, $F_i = F$ and $M_i = M$ for constants *F* and *M*). For each scenario, we ran 500 simulations. First, we fixed *M* at 0.2 and varied *F* from 0.05 to 0.15 to 0.25. As fishing mortality increased, the CV of all the parameter estimates decreased, with the largest decrease for the natural mortality rate estimates (the increased precision presumably resulting from more tags being recovered) (Figure 4a). Second, we fixed *F* at 0.15 at all ages and varied *M* from 0.1 to 0.2 to 0.3. As natural mortality increased, the CV of the fishing mortality rate estimates and the initial population size parameter increased, although only very slightly (the reduced precision presumably resulting from less tags being recovered), whereas the CV of the natural mortality rate estimates decreased (the increased precision resulting from there being more contrast in the data with which to separate the sources of mortality) (Figure 4b). To check for interaction effects between fishing and natural mortality, we ran simulations for the remaining combinations of the above *F* and *M* values. The results behaved predictably, with no signs of any significant interactions (i.e. when we fixed *F* at 0.05 instead of 0.1 and varied *M*, the results looked the same as Figure 4b except shifted vertically upwards by an approximately equal amount for all parameters).

Note that we chose to evaluate the results using the CV of the parameter estimates rather than the RMSE. This is because as the true value of a parameter increases, we expect the absolute variance and bias of the parameter estimates to also increase. We are more interested in the relative error of the parameter estimates than the absolute error, and the CV provides such a measure. Alternatively, we could have used the RMSE relative to the true parameter value. In our case, these options are almost equivalent because the bias is a negligible component of the RMSE.

Effect of assuming an incorrect coefficient of variation for the catch data

Throughout our simulations, we assumed that we knew the catch CV correctly for inputting into the model. We ran additional simulations to investigate the consequences of inputting the wrong catch CV on the estimates of the mortality and abundance parameters. We generated tag-recapture and catch data sets using the same experimental design and parameter values as in the section on the value of incorporating catch data (i.e. 3 release years, 3 recapture years, $N_a = 250$, $F_i = 0.15$, $M_i = 0.2$, $\lambda_i = 1.0$ and $P_1 = 100000$), with the only difference being we let the true catch CV equal 0.2. We then fit the joint tagging and catch model assuming that the catch CV was equal to a range of values from 0.05 to 0.5. For each CV value considered, we ran 500 simulations. The RMSE of the parameter estimates tended to be smallest when the true value of the catch CV was used, but the effect of using the wrong value was negligible (Figure 5). Although these results are specific to the situation being considered, we would expect even less effect if the tagging component of the likelihood was more influential (i.e. if there were more tag releases or if the catch data had more variability).

Application to southern bluefin tuna

Southern bluefin tuna (SBT) are a long-lived pelagic species found throughout much of the southern oceans. Adults of about 8 years and older spawn in the northeast Indian Ocean between Indonesia and the northwest coast of Australia, generally between the months of September and April. Juveniles (ages 1 to 4) tend to spend their austral summers in southern coastal waters, with age 0 to 2 fish commonly found south of Western Australia (WA) and age 1 to 4 fish commonly found off South Australia in the Great Australian Bight (GAB). During the winter juveniles tend to be found in deep oceanic waters, particularly after age 1. The proportion of the global stock of each age class found off WA and in the GAB during the summer months is not known; however, it is thought to be relatively high but diminishing with age. SBT are harvested by a number of different fishing fleets and countries. In particular, juveniles are harvested primarily by Australian purse seiners operating within the GAB during the summer, and to a lesser extent by various Japanese, Korean and Taiwanese longline fleets operating on the high seas (i.e. the Tasman Sea, southeast Indian Ocean and waters off South Africa) mainly during the winter.

Extensive tagging operations were carried out on SBT in the 1990s in which juvenile fish were caught, tagged, and released primarily in the coastal waters south of WA and in the GAB[2]. All fish were double-tagged. The tag numbers and length of each fish were recorded, together with the location and date of release. Multiple age classes, generally ages 1 to 3, were tagged in each year. A fish's age at tagging was estimated based on its length using the method currently adopted by the Commission for the Conservation of Southern Bluefin Tuna (CCSBT) for its stock assessments (Anon. 2001b). Fortunately, SBT grow rapidly as juveniles so there is good separation between length distributions of the ages being tagged, and therefore the number of aging errors should be small. All tagging was done between November and April, so the ages were adjusted in order that fish caught in November or December from a given year-class/cohort were placed in the same age grouping as those tagged after December. This adjusted age is used throughout the appendix, for releases as well as recaptures. Recaptures occurred in the commercial fisheries throughout the geographical distribution of SBT, ranging in longitude from 0 to 180°E and in latitude from 30 to 50°S. Upon recapture, the finder measured the caudal fork length of the fish and recorded this length as well as the tag number, the date and location of recapture, and sometimes the weight of the fish. This information along with one or both of the tags was sent to CSIRO. Recapture age was determined using the estimated release age and the time between release and recapture. Further details regarding the tagging operations and the data compilation procedures used for our analysis can be found in Appendix 4.

Not all tagged fish that were recaptured will have been reported. Before we can apply the joint tagging and catch model to the SBT data we need to know, or have estimates of, reporting rates. Estimating reporting rates for SBT is complicated because of the nature of the SBT fishery, which comprises multiple components with varying reporting rates. Data collected by observers, which were present on a portion of the vessels in some components of the fishery from 1991 to 1997, can be used to provide information about reporting rates, but observer coverage varied between components

---

[2] Relatively small numbers of fish were tagged in other areas, primarily in waters off eastern Tasmania and from longline vessels. These releases have been excluded from the analyses presented here.

and years. Because different fishery components catch different age-compositions of SBT (due to different gear types and different fishing locations), observer data from one component are not very informative about reporting rates in other components. Nevertheless, the available observer data have been used, along with a number of alternative assumptions, to provide year- and age-specific estimates of reporting rates for SBT from 1991 to 1997 (see Preece et al. 2001 for details). The reporting rate estimates that we use in our current analysis come from Table 4, option 1, of Preece et al. (2001). We treat these values as being known without error. By doing so, we are able to illustrate the application of the estimation model described above to real data. In a more rigorous analysis for the purpose of evaluating the SBT stock, we would incorporate uncertainty in the reporting rate estimates and also consider various alternatives.

Table 4 summarizes the SBT tagging data (releases, recaptures and reporting rates) and corresponding catch data used in our analysis. The catch data were compiled from the data used in CSIRO's 2001 stock assessment for SBT (see Preece et al. 2001). Although recapture and catch information exist beyond 1997, reporting rate estimates are not available beyond this year so we cannot include more years in our analysis. It should be noted that release year and recapture year refer to the period from 1 November of the previous year to 31 October of the given year, as opposed to usual calendar year. This year definition is preferred because it is consistent with the age definition being used, so that year and age give equivalent information when referring to fish from a given cohort. Moreover, this year definition corresponds with the Australian purse seine fishing season (which generally operates from November to April). Because the purse seine fishery catches the majority of juvenile SBT, which are the fish of primary interest in our analysis, it made sense that our fishing mortality estimates correspond to fishing seasons rather than to calendar years.[3]

---

[3] The numbers in Table 4 are not the same as those used in many previous analyses for SBT (e.g. Polacheck et al. 1997; Pollock et al. 2002). There are two main reasons for the differences. First, previous analyses have almost always used the adjusted (Nov to Oct) year and age definitions for tallying releases but the usual calendar year and age definitions for tallying recaptures and catches. Second, the method used by the CCSBT to estimate age from length was updated in 2001 to incorporate a new growth curve, which changed the age estimate of many young fish.

Tag-recapture data exist for multiple cohorts of SBT, therefore we can apply the joint tagging and catch model to each cohort separately (using the model equations directly as they were presented), or we can apply the model to all cohorts simultaneously (by modifying the equations to have an additional subscript for year). These two options are equivalent if each cohort has a unique set of parameters (i.e. fishing mortality rates and natural mortality rates are allowed to differ between cohorts). However, for SBT (and for most species), natural mortality rates are assumed constant over time when conducting stock assessments. We have made this assumption in the analysis presented here. Thus, we still allow for natural mortality to differ between age 1 and older fish, but not between cohorts. To implement this restriction requires all cohorts to be modelled simultaneously, and this was the approach taken in the analysis presented here.

As in our simulations, we assume the catch-at-age is Gaussian distributed with a constant CV across ages and also years (since we have multiple cohorts). We expect the uncertainty in the catch-at-age data for SBT to be fairly large due to errors in the age distribution (from sampling variability and errors in the age estimates) and due to errors in the total catch numbers (from under-reporting as well as difficulties in compiling catch data in a fishery comprising multiple nations and fleets). In the results presented we assumed a CV of 0.3.

In applying the model to SBT data, we modified the tag-recapture component to allow for fishing mortality to differ between tagged fish in the year of tagging and untagged fish in that same year (following the model presented in Hoenig et al. 1998b). This is to allow for the fact that newly tagged fish will not be fully mixed with the untagged population immediately after tagging, and for the fact that tagging generally occurs during the fishing season so tagged fish are only vulnerable for part of the season. We assume that tagged and untagged fish are fully mixed by the year following release (recall all tagging of SBT occurred between November and April so this allows several months for mixing to occur). The modified return probabilities are given by:

$$p_{a,i} = \begin{cases} u_i^* \lambda_i & i = a \\ S_a^* u_i \lambda_i & i = a+1 \\ S_a^* S_{a+1} \cdots S_{i-1} u_i \lambda_i & i > a+1 \end{cases}$$

where

$$S_i^* = \exp\left(-F_i^* - M_i\right)$$

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

$F_i^* =$ the instantaneous fishing mortality rate for age i fish tagged at age i

All other parameters are defined as before. Here S_i^* and u_i^* represent the annual survival and exploitation rates of fish in the year that they are tagged, whereas S_i and u_i represent the annual survival and exploitation rates of fish in the years following tagging. Note that the model proposed by Hoenig et al. (1998b) cannot provide an estimate of F_1, only F_1^*; by incorporating catch data, both parameters can be estimated.

The maximum likelihood parameter estimates and their estimated standard errors are presented in Table 5. Standard error estimates were obtained from the inverse negative Hessian matrix (evaluated at the maximum likelihood estimates) by taking the square root of the diagonal entries. The estimates of M_1 and M_2 are essentially equal and suggest that these two parameters could be combined into a single non-age-specific natural mortality rate. However, the estimate of M_2 seems too high and may reflect not only natural mortality but also underestimation of reporting rates for some ages as well as emigration of some age 2 fish out of the main fishing area (and thus not available for recapture at age 3). The estimates of cohort size decrease over time, but the CV's of the estimates are large (28-56%) and likely underestimated (see Discussion). Also, not fully accounting for changes in reporting rates may be a confounding factor. Before drawing any conclusions, a detailed consideration of alternative reporting rate models as well as residual patterns would need to be undertaken.

A profile of the negative log-likelihood for the catch CV revealed that, as expected from previous discussions, the global minimum in the negative log-likelihood occurs near zero, but that a local minimum also occurs around 0.26 (Figure 6). This, combined with our previous simulation-based observation that the variance parameter tends to be underestimated even when a local minimum in the negative log-likelihood (i.e. maximum in the likelihood) exists, suggests that our initial choice of 0.3 for the catch CV was reasonable.

In the model presented for SBT, there are a large number of parameters being estimated (47) in relation to the number of data points (85). This brings up the issue of model selection and whether or not the parameter set could be reduced. For example, are the fifteen additional F^* parameters to allow for an initial period of non-mixing necessary? Is it necessary to have age-specific natural mortality rates? Instead of having year- and age-specific fishing mortality rates, can fishing mortality be modelled as a separable function with independent year and age factors (i.e. $F_{a,y} = F_a F_y$)? Alternatively, can fishing mortality be modelled using a selectivity function? Model selection can be performed using likelihood ratio tests or, more commonly in wildlife applications, using the Akaike information criterion (AIC) (Akaike 1974). While we stress that performing model selection to find the most parsimonious model is an important part of any analysis, it is not the focus of this appendix and so we did not carry out an exhaustive model selection procedure for the SBT data. We did, however, test the necessity of including separate F's in the year of tagging. The AIC value for the model including separate F's was 47784.6 compared to 47988.6 for the model without separate F's. These results suggest that an initial period of non-mixing does in fact exist and should be allowed for in the model. The residuals (not shown) were much improved in the model with separate F's as well. We also confirmed that M_1 and M_2 can, at least statistically speaking, be replaced with a single M parameter – fitting the model with a single M resulted in no change to the negative log-likelihood value to the first decimal place, and thus a reduction in the AIC of approximately 2. However, our suspicion that the M_2 estimate may encompass more than just natural mortality should be investigated before reaching any conclusions. Alternative parameterizations for the fishing mortality rates would also be worthwhile exploring.

Discussion

In the current appendix, we have developed a comprehensive framework for modelling data from multi-year tagging experiments in a fishery context that incorporates catch data. Integrating multi-year tagging data and catch data results in a synergy between the two data sets in that it provides an estimate of the initial population size, which cannot be obtained from the multi-year tagging data alone (i.e. from a Brownie-type experiment), and it provides estimates of mortality rates, which cannot be obtained from single release tagging data and catch data alone (i.e. from a Petersen-type experiment). Additionally, incorporating catch data into the tagging model improves the mortality rate estimates (especially fishing mortality rates). Our simulation results suggest that including catch data can improve the accuracy of fishing mortality rate estimates by up to ~40% and natural mortality rate estimates by up to ~10%. Of course, the degree of improvement depends on the situation and is greatest when the catch data are known precisely and the number of tag releases is relatively small. Our results also show that improving the precision of the catch data can lead to significant improvements in the accuracy of both the population size estimate and the fishing mortality rate estimates (especially at young ages and low numbers of releases), whereas increasing the number of releases can lead to large improvements in the natural mortality rate estimates as well as the fishing mortality rate estimates (especially when the accuracy of the catch data is low).

The simulation results presented were for a tagging experiment involving a single cohort. In practice, it is likely that two or more cohorts would be tagged in a given year, as was done in the case of southern bluefin tuna tagging experiments. If any parameters are thought to be the same between cohorts, then tagging multiple cohorts can improve the information available for estimation and inference about not only these parameters, but the other model parameters as well. For example, we ran some multi-year, multi-cohort simulations allowing for the fishing mortalities to differ by year and age but for natural mortalities to differ only by age (and still differ only between age 1 and age 2 plus). For each cohort, we assumed the same experimental design and parameter values as in the section on the value of incorporating catch data

except we assumed a true catch CV of 0.2. We found that having data from two cohorts reduced the RMSE on the estimates of M_1 and M_2 by about 25%, and having three cohorts reduced the RMSE on these estimates by 30-40%. The improvement in the fishing mortality estimates increased at older ages. In particular, for the F_1 estimates there was almost no difference in the precision from having more cohorts, but for the F_3 estimates having two cohorts reduced the RMSE by 25-30% and having three cohorts resulted in a further reduction of about 5%. Having either one or two additional cohorts led to about a 10% improvement in the RMSE of the P_1 estimates. Having additional cohorts may also allow for more reliable estimation of the catch CV if constraints are put on the fishing mortality rates between cohorts.

Another advantage of a multi-cohort, multi-year tagging experiment is that it allows for testing of assumptions that are frequently assumed to be true in fishery stock assessments because of lack of data to do otherwise. For example, in our application to southern bluefin tuna data, we could have tested for significant differences in the age-specific natural mortality rates between cohorts to see if our assumption that they were constant over time was reasonable. Similarly, we could have tested whether fishing mortality could have been modelled as a separable process (e.g. a multiplicative combination of a year and age effect). Constant age-specific natural mortality rates and separable fishery mortality rates (at least for a restricted period of years) are frequently assumed in stock assessments. If these assumptions are found to be reasonable, then the number of parameters to be estimated can be greatly reduced, with resulting gains in precision.

High correlations were present between many parameter estimates, and are expected with the model formulation being used. In particular, positive correlations between the natural mortality and fishing mortality estimates, especially at older ages, are expected because an increase in natural mortality means that less fish are still alive in the population; thus, in order to achieve a particular level of catch, fishing mortality must increase (i.e. the percentage of the population caught must increase) as natural mortality increases. For the same reason, the fishing mortality rates between ages and fisheries are often highly positively correlated. For example, if fishing mortality at a given age increases, then there are less fish alive at subsequent ages; thus, in order to

achieve a particular level of catch, fishing mortality must increase at subsequent ages in response to an increase in fishing mortality at a younger age. Finally, negative correlations between the initial population size and the fishing mortality estimates are often present because in order to have achieved a particular level of catch, the population size must have been larger if the fishing mortality had been low than if it had been high.

The estimates presented on the precision of the parameter estimates (in both the simulations and the application to southern bluefin tuna data) are likely to be optimistic for a number of reasons. First, we assumed that the reporting rates were known without error and this assumption is unlikely to be true. Second, if tagging mortality or tag shedding exist then there would be additional variance introduced in the parameter estimates, as well as bias if they are not accounted for. Third, the variance in the number of recaptures is likely to be overdispersed relative to a multinomial distribution due to incomplete mixing and heterogeneity in the capture probabilities of fish. All of these factors need to be addressed in the design and analysis of any tagging experiment.

With regard to the first issue, making the assumption that reporting rates were known without error provided the most straightforward way of evaluating the potential gain achieved by incorporating catch data in the tag-recapture model. However, in rigorous applications of the model, uncertainty in the reporting rates should be accounted for. A Bayesian approach could be used in cases where one has reporting rate estimates and associated uncertainties, but not the raw data used to obtain these estimates. However, if the raw data are available, it would be preferable to incorporate these data directly in the model, either as a multiplicative likelihood component if the data are independent of the tag-recapture and catch data or through modifications to the tagging and catch likelihoods presented if they are not. The use of observer data to estimate reporting rates is clearly a case where the data are not independent of the tagging and catch data. In Appendix 7 we have extended the model presented in this appendix to incorporate the estimation of reporting rates through observer data.

With regard to the second issue, if tag mortality and tag shedding cannot be assumed to be essentially zero, then auxiliary information needs to be collected to allow for their estimation. There has been extensive work on approaches for estimating these quantities, especially tag shedding from double-tagging experiments (e.g. Beverton and Holt 1957; Kirkwood and Walker 1984; Barrowman and Myers 1996). Once estimated, immediate (often referred to as Type I) tag shedding and tagging mortality are essentially treated as a reduction in the number of releases, whereas subsequent (Type II) tag shedding and tagging mortality are generally assumed to occur at a constant rate over time and treated the same way as natural and fishing mortality. As such, if these quantities are not estimated with a reasonably high level of precision, they will introduce substantial uncertainty into the abundance and mortality rate estimates. We also note that if heterogeneity exists among taggers, then using a simple average rate for tag shedding and tagging mortality over all taggers can induce substantial biases to the parameter estimates of interest (Hearn et al. 1991).

The third issue of incomplete mixing of tagged animals within the full population and heterogeneity in the capture probabilities is the most challenging problem confronting the use of tagging experiments. These factors can introduce increased variance as well as biases and are probably the biggest limitation in the application of tagging experiments in fisheries. Strategies can and should be used to avoid and minimize the effect of incomplete mixing both in the design of the releases and in the analysis of the return data. For example, to the greatest extent possible, it is important to spread tagging effort throughout the area where fish are found. In cases where tagging occurs around the same time as peak catches (which is likely since that is usually the time when fish are most available), it is important to allow for a period of non-mixing. A method for incorporating non-mixing in the first year has been developed for Brownie-type models (Hoenig et al. 1998b) and, as we illustrated in our application to southern bluefin tuna data, can easily be extended to the Brownie model with catch data developed here. If the period of non-mixing is less than a year, then precision in the parameter estimates can be gained by modifying to model to only allow for non-mixing in part of the first year. When incomplete mixing results from a high degree of site fidelity within the geographic range of the population, it may be necessary to include a spatial component within the analysis. In such cases, there is no universal

solution and the most appropriate estimation approach will depend upon the specifics of the population. We are currently in the process of developing a spatial version of the joint tagging and catch model for southern bluefin tuna. A special case of incomplete mixing occurs if fish have a high degree of school fidelity, in which case the probabilities of recapturing fish tagged from the same school are not independent. When school fidelity exists, more realistic variance estimates could perhaps be obtained using a bootstrap approach and resampling the data using schools as the basic sampling unit. A more general approach would be to directly incorporate the overdispersion into the model using, say, a Dirichlet-multinomial distribution as opposed to a multinomial distribution (see Appendix 9).

Overall the results from this appendix demonstrate the potential for combining multi-year tagging data with catch-at-age data to provide reliable estimates of the population parameters of interest for stock assessments. Reasonable levels of precision can be obtained for all parameter estimates as long as there is an appropriate level and balance of effort devoted to tag releases and estimation of the catch-at-age data.

Similarities exist between the methods and likelihood function developed here and those developed for use in integrated statistical catch at age/size stock assessment models involving tagging data (e.g. Fournier et al. 1998; Kolody and Polacheck 2001; Butterworth et al. 2003). However, we are not aware of any integrated stock assessment models that have incorporated a Brownie-type estimator for the tagging data (although it would be relatively straightforward to do so). Even so, there are important distinctions between these integrated stock assessment models and the combined tagging and catch model presented here. Integrated stock assessment models attempt to estimate the entire age structure and history of a population since the beginning of exploitation. The models are over-parameterized and various assumptions (particularly with respect to catchability and selectivity) are required to yield an identifiable set of parameters. These can involve various "penalty" terms, and the results are frequently sensitive to alternative assumptions and to relative weights given to the penalty terms and data components of the objective function. Simply improving the precision of the data will not necessarily improve the reliability of the estimates or the structural features (e.g. penalty functions) of the model. On the

other hand, the integrated tagging and catch model allows for all parameters to be estimated (with the exception of natural mortality at the older ages) without requiring any assumptions about selectivity and catchability. As a result, the model can be used to test a suite of assumptions with regard to the parameters, such as whether fishing mortality can adequately be described using a selectivity function or whether certain parameters are common between ages or years. Furthermore, the integrated tagging and catch model can be used to directly evaluate the trade-off in the amount of resources placed into the different data collection components. Of course, a disadvantage is that estimates are only possible for cohorts and ages for which tagging data are available. In this context, we see the integrated tagging and catch model as complementary to stock assessment models.

Another feature of the integrated tagging and catch model is that it does not rely on the use of catch rate (CPUE) or standard effort data. These data form an essential component of many stock assessments (i.e. those without fishery-independent surveys or tagging), but are also one of the more uncertain and unreliable components. Tagging can provide a useful alternative to catch rate and effort data, and are perhaps the only viable alternative in a number of fisheries where fishery-independent surveys are not possible. Admittedly, tagging experiments involve a number of complications and assumptions of their own. Ensuring that the tagging experiment (both the release and recovery components) are well designed and executed is critical in order to avoid situations that would render the data as uncertain and unreliable as catch rate data. However, unlike catch rate data, many of the uncertainties associated with tag-recapture data can be minimized through well-designed and implemented tagging experiments and comprehensive analysis methods. Moreover, the data themselves can be used to test for violations of the critical assumptions.

Ultimately the tagging and catch likelihood components developed here may be incorporated into an integrated statistical stock assessment. However, an independent analysis of only the tagging and catch data is important because it provides estimates of critical parameters that are not confounded by the structural assumptions of the stock assessment models and also allows for testing of these assumptions. Furthermore, it provides a means to check for consistency between the tagging and

catch data themselves as well as between these data sets and other data sets used in the assessment. To this end, it is desirable to carry out a variety of independent analyses using different data components in order to cross check model assumptions and consistency of data sets. Such checks should be done as a standard part of any assessment in order to ensure the robustness of the conclusions. If inconsistencies are found it is important to develop and incorporate a range of hypotheses and not simply "average" out the discrepancies.

In summary, tagging experiments provide critical independent estimates of key population parameters that can be compared with estimates from other sources. Without being a panacea to the estimation problems facing stock assessments, tagging experiments are a potentially powerful tool that we think deserve more consideration within research programs for improving and augmenting traditional stock assessment approaches.

Acknowledgements

Literature Cited

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, **19**: 716-723.

Anon. 2001a. Report of the tagging program workshop. Commission for the Conservation of Southern Bluefin Tuna. 2-4 October. Canberra, Australia.

Anon. 2001b. Report of the fifth meeting of the scientific committee. Commission for the Conservation of Southern Bluefin Tuna. 19-24 March 2001. Tokyo, Japan.

Anon. 2001c. Report of the eighth annual meeting. Commission for the Conservation of Southern Bluefin Tuna. 15-19 October 2001. Miyako, Japan.

Anon. 2002. Report of the seventh meeting of the scientific committee. Commission for the Conservation of Southern Bluefin Tuna. 9-11 2002. Canberra, Australia.

Barrowman, N.J., and Myers, R.A. 1996. Estimating tag-shedding rates for experiments with multiple tag types. Biometrics **52**: 1410-1416.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. 2, Mar. Fish. G.B. Minist. Agric. Fish. Food **19**, 533 pp.

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

Butterworth, D.S., Ianelli, J.N., and Hilborn, R. 2003. A statistical model for stock assessment of southern bluefin tuna with temporal changes in selectivity. S. Afr. J. Mar. Sci. **25**: 331-361.

Fournier, D.A., Hampton, J., and Sibert, J.R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Can. J. Fish. Aquat. Sci. **57**: 1002-1010.

Hearn, W.S., Leigh, G.M., and Beverton, R.J.H. 1991. An examination of a tag-shedding assumption, with application to southern bluefin tuna. ICES J. Mar. Sci. **48**: 41-51.

Hearn, W.S., Polacheck, T., Pollock, K.H., and Whitelaw, W. 1999. Estimation of tag reporting rates in age-structured multicomponent fisheries where one component has observers. Can. J. Fish. Aquat. Sci. **56**: 1255-1265.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Chapman and Hall, London. 570pp.

Hoenig, J.M., Barrowman, N.J., Hearn, W.S., and Pollock, K.H. 1998a. Multiyear tagging studies incorporating fishing effort data. Can. J. Fish. Aquat. Sci. **55**: 1466-1476.

Hoenig, J.M., Barrowman, N.J., Pollock, K.H., Brooks, E.N., Hearn, W.S., and Polacheck, T. 1998b. Models for tagging data that allow for incomplete mixing of newly tagged animals. Can. J. Fish. Aquat. Sci. **55**: 1477-1483.

Kirkwood, G.P., and Walker, M.H. 1984. A new method for estimating tag shedding rates, with application to data for Australian salmon, *Arripis trutta esper* Whitley. Aust. J. Mar. Freshwater Res. **35**: 601-606.

Kolody, D., and Polacheck, T. 2001. Application of a statistical catch-at-age and length integrated analysis model for the assessment of southern bluefin tuna stock dynamics 1951-2000. Commission for the Conservation of Southern Bluefin Tuna. CCSBT-SC/0108/13.

Morgan, B.J.T., and Thomson, D.L. (Editors). 2002. Statistical analysis of data from marked bird populations. J. Applied Statistics **29**(1-4), 669pp.

Polacheck, T., Hearn, W., Millar, C., Whitelaw W., and Stanley, C. 1996. Estimation of natural and fishing mortality rates for juvenile southern bluefin tuna based on multi-year tagging of cohorts. Commission for the Conservation of Southern Bluefin Tuna CCSBT/SC/96/11.

Polacheck, T., Hearn, W., Millar, C., Whitelaw, W., and Stanley, C. 1997. Updated estimates of mortality rates for juvenile southern bluefin tuna from multi-year tagging of cohorts. Council for the Conservation of Southern Bluefin Tuna CCSBT/SC/970726.

Pollock, K.H., Hoenig, J.M., and Jones, C.M. 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.

Pollock, K.H., Hearn, W.S., and Polacheck, T. 2002. 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.

Preece, A., Polacheck, T., Kolody, D., Eveson, P., Ricard, D., Jumppanen, P., Farley, J., and Davis, T. 2001. Summary of the primary data inputs to CSIRO's 2001 stock assessment models. Commission for the Conservation of Southern Bluefin Tuna CCSBT-SC/0108/21.

Quinn, T.J., and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York. 542pp.

Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. Charles Griffin, London. 654 pp.

Table 1. Expected number of returns at age from each release event in a multi-year tagging study involving a single cohort of fish.

Release age	Number releases	Expected number of returns from age class i				
		1	2	3	4	5
1	N_1	$N_1 u_1 \lambda_1$	$N_1 S_1 u_2 \lambda_2$	$N_1 S_1 S_2 u_3 \lambda_3$	$N_1 S_1 S_2 S_3 u_4 \lambda_4$	$N_1 S_1 S_2 S_3 S_4 u_5 \lambda_5$
2	N_2		$N_2 u_2 \lambda_2$	$N_2 S_2 u_3 \lambda_3$	$N_2 S_2 S_3 u_4 \lambda_4$	$N_2 S_2 S_3 S_4 u_5 \lambda_5$
3	N_3			$N_3 u_3 \lambda_3$	$N_3 S_3 u_4 \lambda_4$	$N_3 S_3 S_4 u_5 \lambda_5$

Table 2. Expected number of fish caught at age from a cohort of fish with age 1 abundance of P_1.

Size of cohort	Expected catch from age class i				
	1	2	3	4	5
P_1	$P_1 u_1$	$P_1 S_1 u_2$	$P_1 S_1 S_2 u_3$	$P_1 S_1 S_2 S_3 u_4$	$P_1 S_1 S_2 S_3 S_4 u_5$

Table 3. Summary of parameter estimates obtained from 500 simulations using the model with tag data only and using the model with both tag and catch data. Results obtained using 250 tag releases per year and, in the model including catch data, a coefficient of variation for the catch data of 0.1.

	Parameter	F_1	F_2	F_3	M_1	M_2	$P_1/10^5$
Model	True value	0.15	0.15	0.15	0.20	0.20	1.0
Tag only	Mean	0.1536	0.1533	0.1562	0.2242	0.2173	--
	SE	0.0013	0.0012	0.0018	0.0073	0.0084	--
	CV	0.1919	0.1763	0.2572	0.7284	0.8630	--
Tag & catch	Mean	0.1501	0.1512	0.1556	0.2121	0.2099	1.0092
	SE	0.0008	0.0008	0.0018	0.0065	0.0077	0.0052
	CV	0.1132	0.1241	0.2527	0.6806	0.8232	0.1147

Table 4. Southern bluefin tuna tag-recapture, catch and reporting rate data used in the joint tagging and catch model.

Cohort	Release year	Release age	Number releases	Number tagged fish recaptured						
				1991	1992	1993	1994	1995	1996	1997
1990	1991	1	3299	20	40	46	23	13	5	4
	1992	2	4646		88	159	101	33	12	8
	1993	3	2777			66	78	32	17	15
1991	1992	1	2144		1	21	56	37	11	7
	1993	2	2937			60	68	67	21	11
	1994	3	3640				77	145	30	40
1992	1993	1	4898			2	41	201	91	58
	1994	2	3158				29	167	76	52
	1995	3	2629					55	103	74
1993	1994	1	9003				4	110	401	364
	1995	2	5899					83	395	363
	1996	3	1511						115	201
1994	1995	1	8585					0	87	622
	1996	2	2518						77	339
	1997	3	526							91

Cohort	Number fish caught						
	1991	1992	1993	1994	1995	1996	1997
1990	48111	34651	121008	72260	37713	24143	21962
1991		7500	37418	117982	60187	38009	27739
1992			403	9672	131524	79619	44177
1993				164	29677	168672	72505
1994					408	25129	203016

Cohort	Reporting rate						
	1991	1992	1993	1994	1995	1996	1997
1990	0.985	0.504	0.664	0.601	0.265	0.335	0.546
1991		0.956	0.774	0.706	0.426	0.354	0.553
1992			0.739	0.779	0.735	0.495	0.656
1993				0.962	0.820	0.539	0.767
1994					0.957	0.628	0.867

Table 5. Parameter estimates obtained by applying the integrated model to southern bluefin tuna tagging and catch data from several cohorts. Standard error estimates are given in parentheses below the point estimates. Note that the figures for the initial population size, P_1, are in millions. Results were obtained using a coefficient of variation for the catch data of 0.3.

M_1	M_2
0.455	0.463
(0.031)	(0.039)

Cohort	F_1	F_2	F_3	F_4	F_5	F_6	F_7	P_1
1990	0.014	0.046	0.100	0.111	0.169	0.113	0.129	4.013
	(0.004)	(0.008)	(0.009)	(0.012)	(0.032)	(0.029)	(0.040)	(0.561)
1991	0.003	0.023	0.089	0.213	0.122	0.137		3.412
	(0.001)	(0.004)	(0.009)	(0.023)	(0.023)	(0.033)		(0.472)
1992	0.000	0.019	0.173	0.205	0.207			2.259
	(0.000)	(0.003)	(0.013)	(0.023)	(0.036)			(0.327)
1993	0.000	0.030	0.300	0.448				1.714
	(0.000)	(0.003)	(0.021)	(0.058)				(0.276)
1994	0.000	0.032	0.336					1.830
	(0.000)	(0.004)	(0.024)					(0.356)

Cohort	F_1^*	F_2^*	F_3^*
1990	0.008	0.047	0.046
	(0.002)	(0.005)	(0.006)
1991	0.001	0.034	0.038
	(0.001)	(0.004)	(0.004)
1992	0.001	0.015	0.036
	(0.000)	(0.003)	(0.005)
1993	0.001	0.022	0.187
	(0.000)	(0.002)	(0.019)
1994	0.000	0.061	0.283
	(0.000)	(0.007)	(0.031)

Figure 1. Effect of changing the number of tag releases and coefficient of variation (CV) of the catch data on the root mean squared error (RMSE) of the fishing mortality rate estimates (a-c), the mortality rate estimates (d-e), and the initial population size estimate (f). Catch CV=NA refers to results obtained using tagging data only. Results obtained using true parameter values $F_1 = F_2 = F_3 = 0.15$, $M_1 = M_2 = 0.2$, and $P_1 = 100000$; 500 simulations for each scenario.

a) F_1

b) F_2

c) F_3

Number of releases

d) M_1
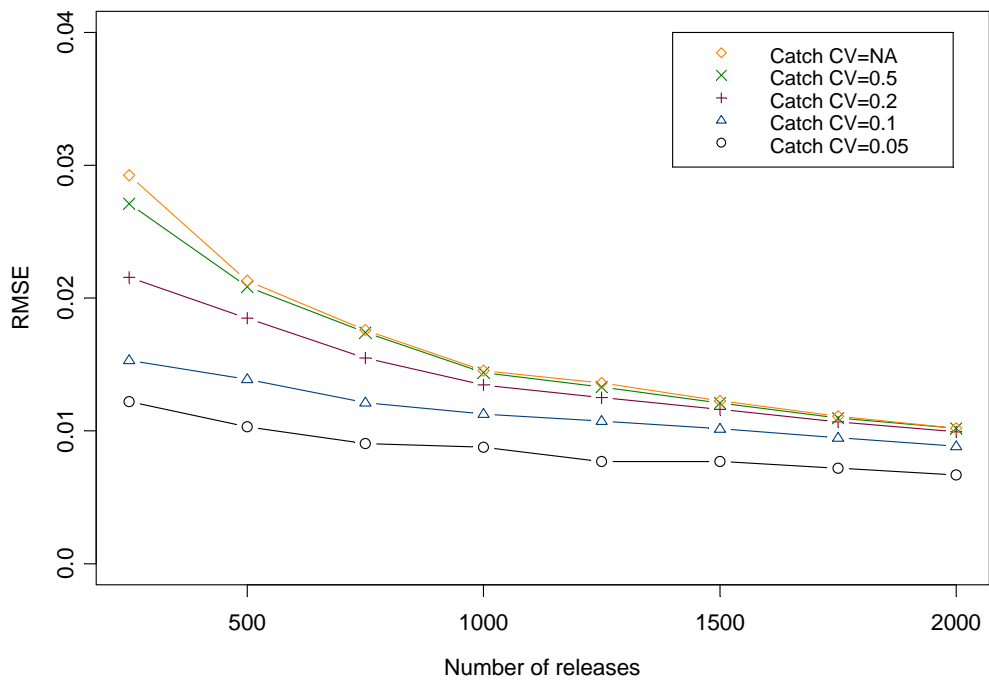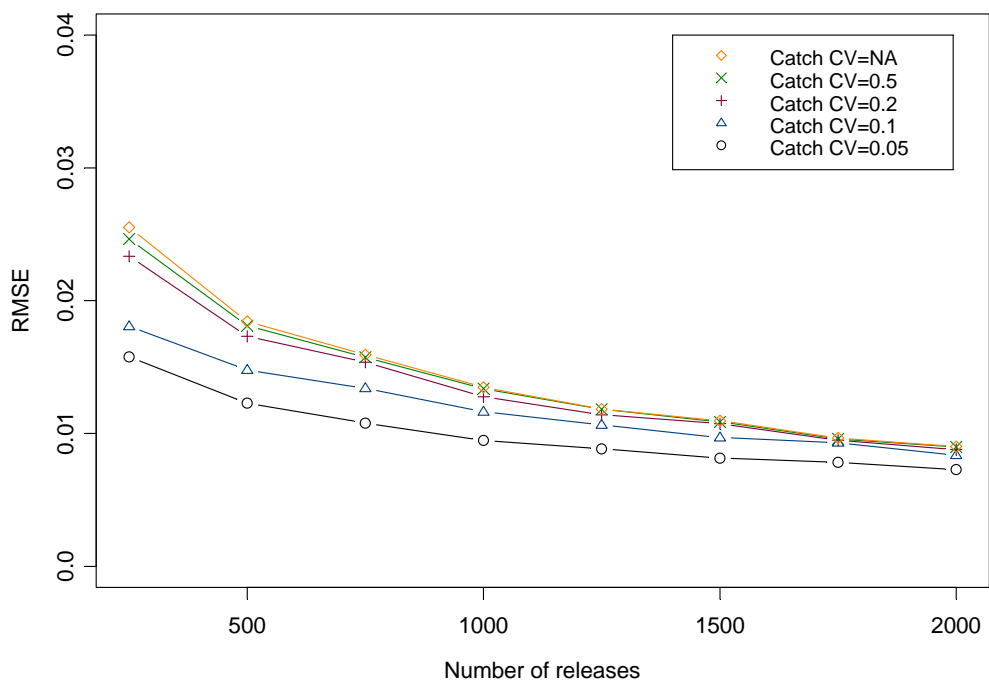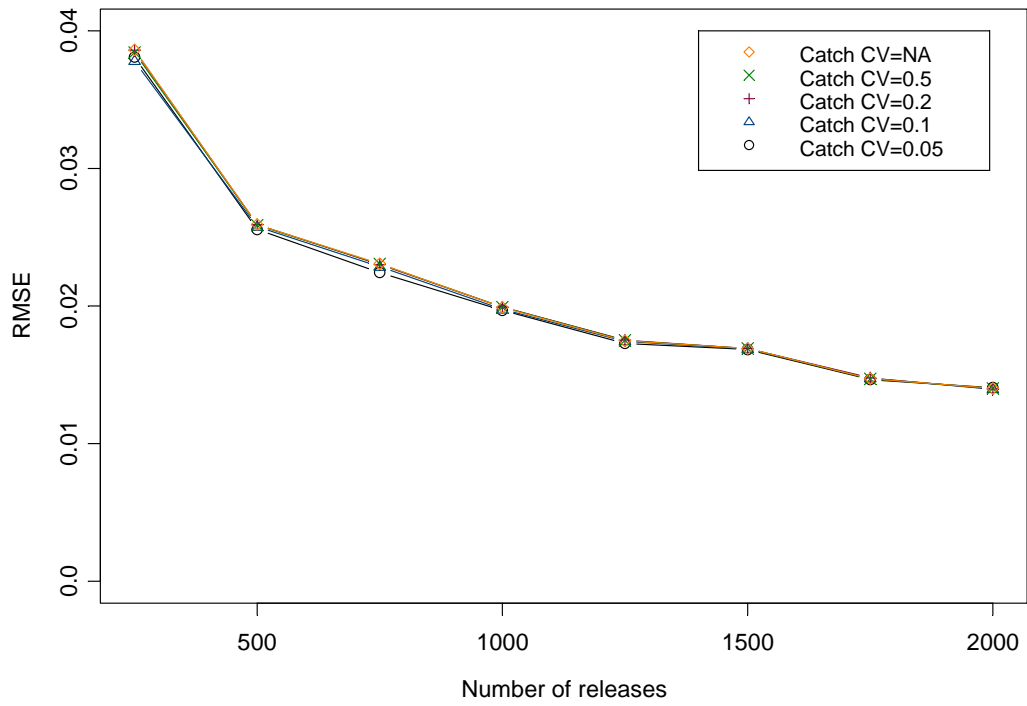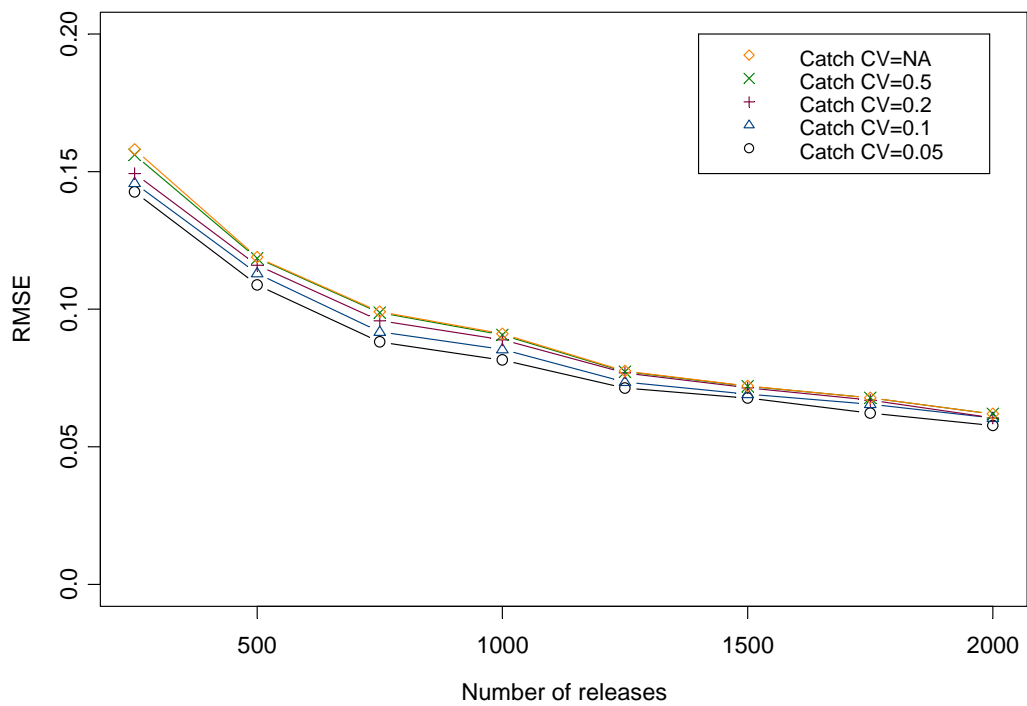
Number of releases

e) M_2
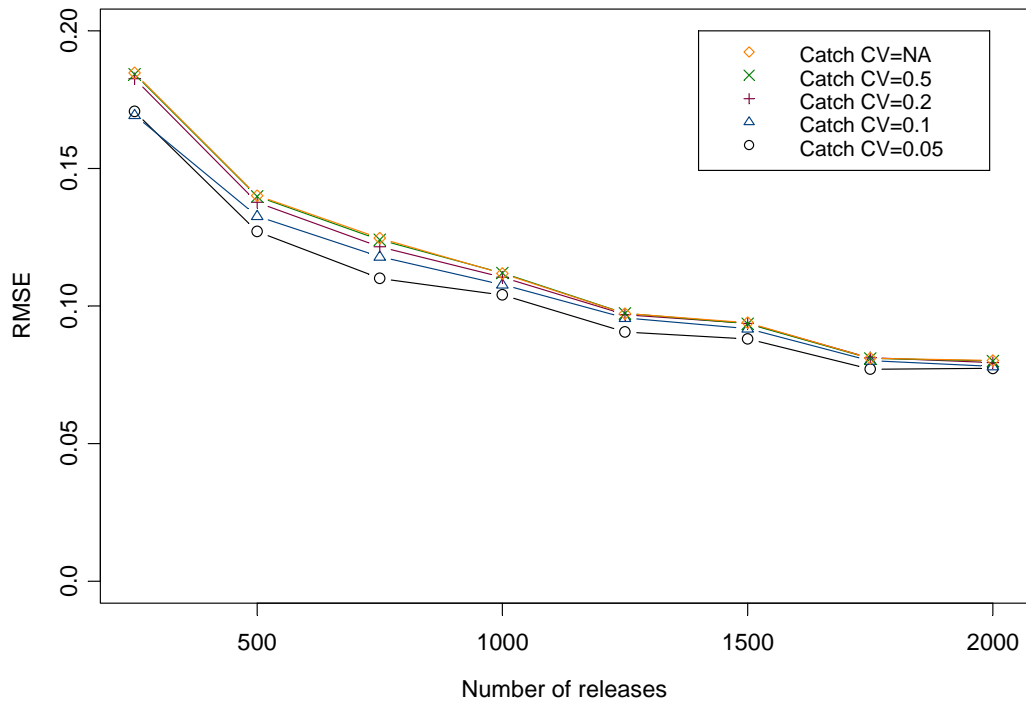
f) P_1 (No estimate available for model with tagging data only.)
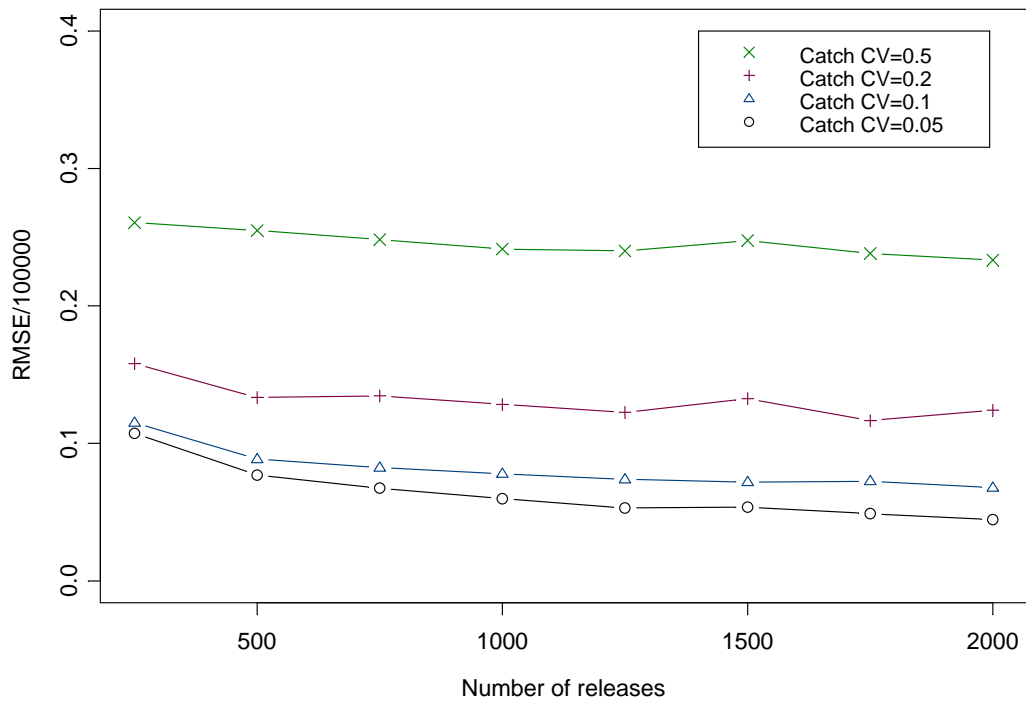
Figure 2. Effect of changing the number of tag releases (N) on the biases of the parameter estimates: a) using 3 release years and 3 recapture years; b) using 3 release years and 5 recapture years. Results shown were obtained using the model with tagging data only using parameter values $F_1 = F_2 = F_3 = 0.15$, $M_1 = M_2 = 0.2$, and $P_1 = 100000$; 500 simulations for each scenario.

a)

b)

A5-42
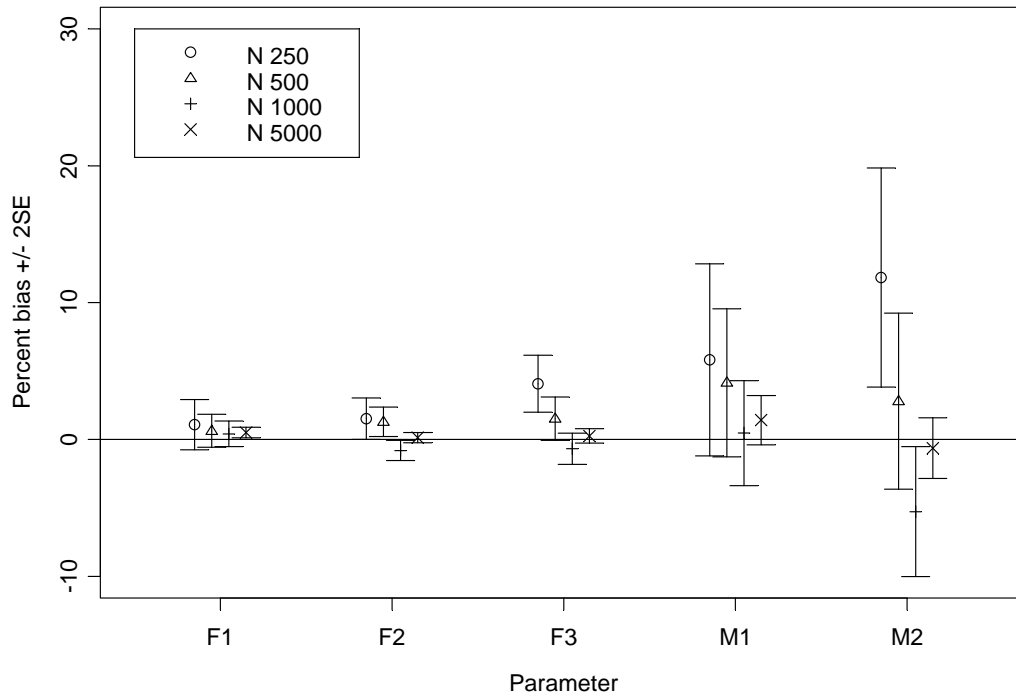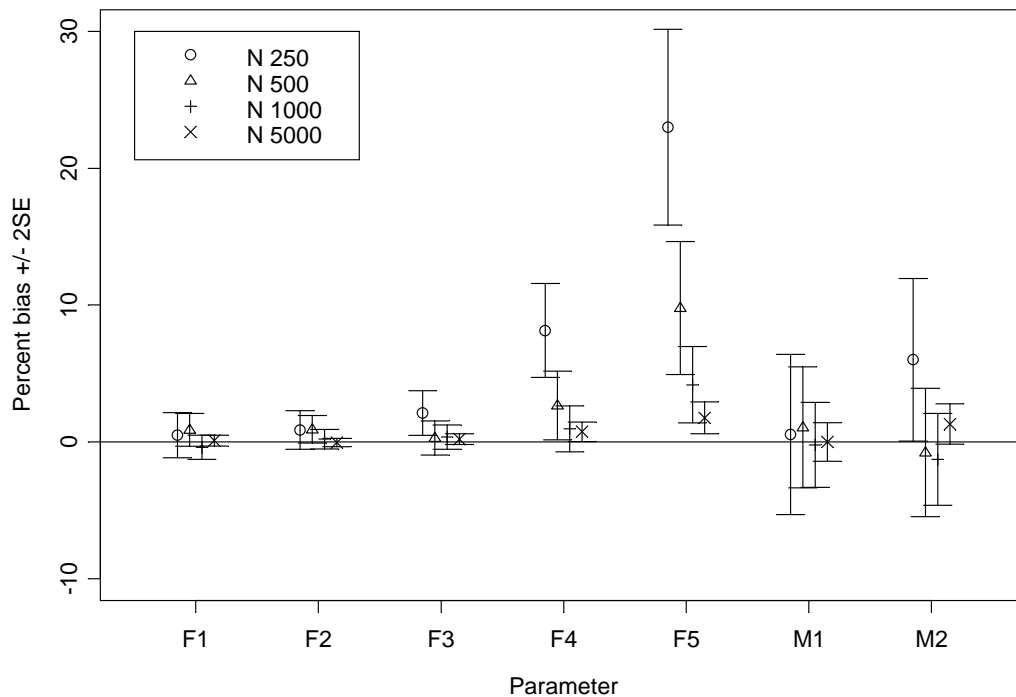
Figure 3. Effect of changing the number of tag releases and the coefficient of variation (CV) of the catch data on the bias of the initial population size (P_1) estimate. Results obtained using the integrated tagging and catch model with parameter values $F_1 = F_2 = F_3 = 0.15$, $M_1 = M_2 = 0.2$, and $P_1 = 100000$; 500 simulations for each scenario.
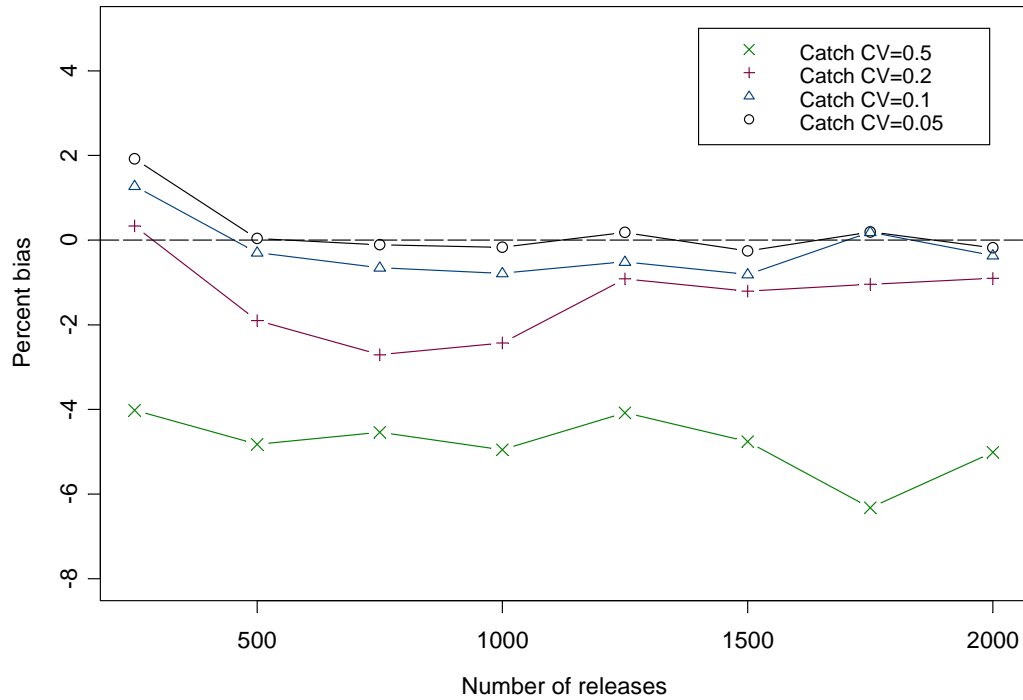
Figure 4. Effect of varying a) the level of fishing mortality (F), and b) the level of natural mortality (M), on the coefficient of variation (CV) of the parameter estimates in the model with both tagging and catch data. Results obtained using 250 tag releases per year, $P_1 = 100000$ and $\upsilon = 0.1$; 500 simulations for each scenario.

a)
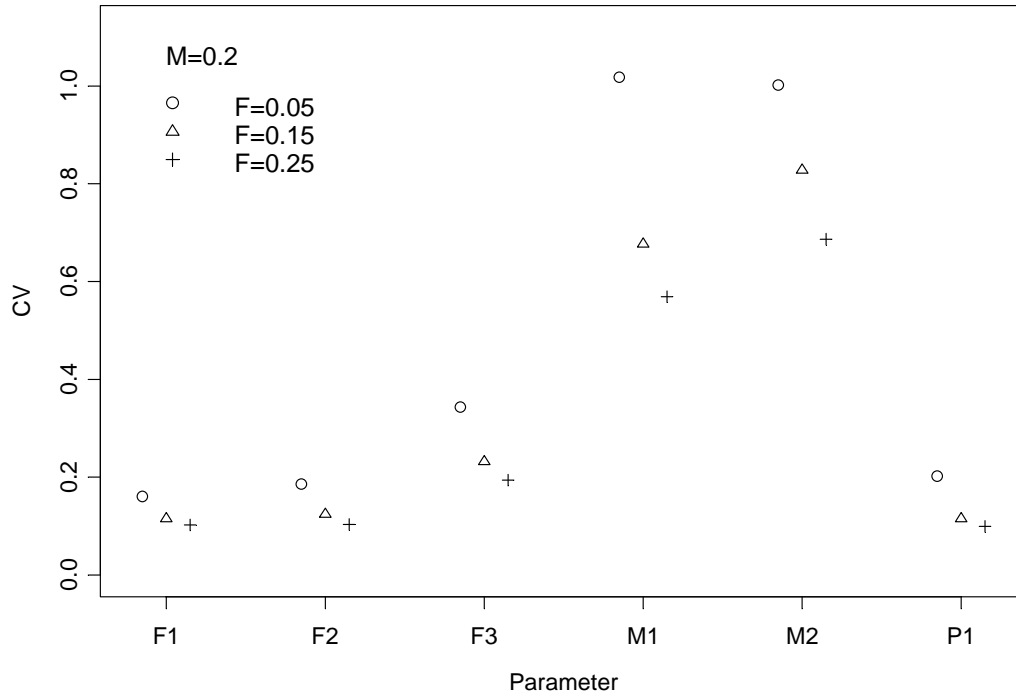
b)
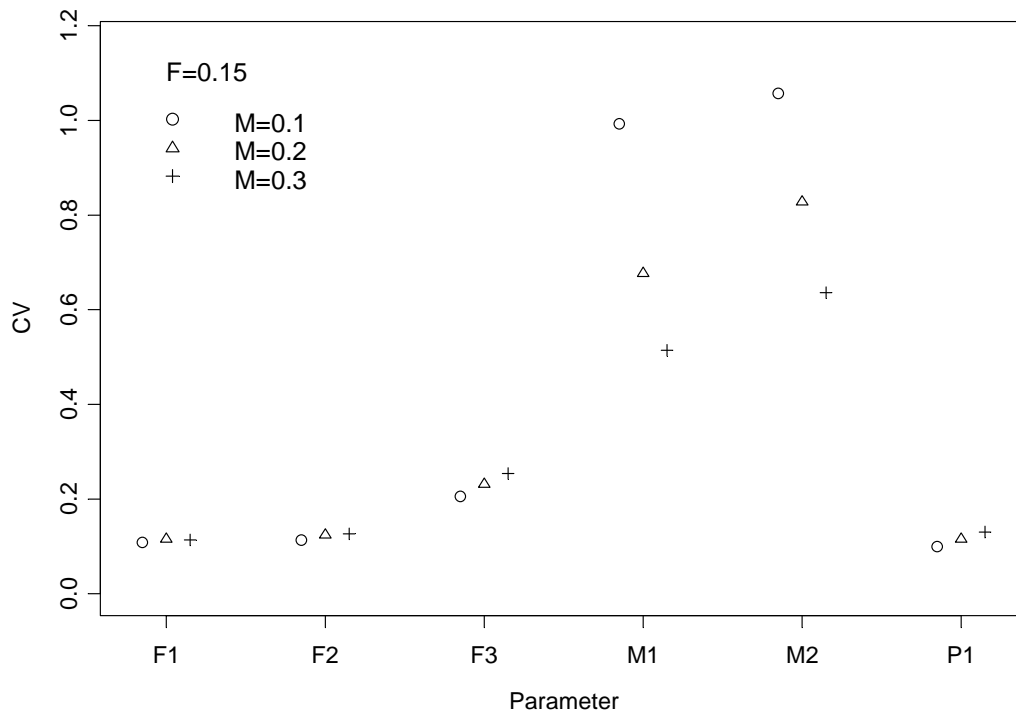
Figure 5. Effect of assuming various values for the coefficient of variation (CV) of the catch data on the root mean squared error (RMSE) of the parameter estimates when the true CV value is 0.2. Results obtained using integrated tagging and catch model with 250 tag releases per year and parameter values $F_1 = F_2 = F_3 = 0.15$, $M_1 = M_2 = 0.2$, and $P_1 = 100000$; 500 simulations for each scenario.
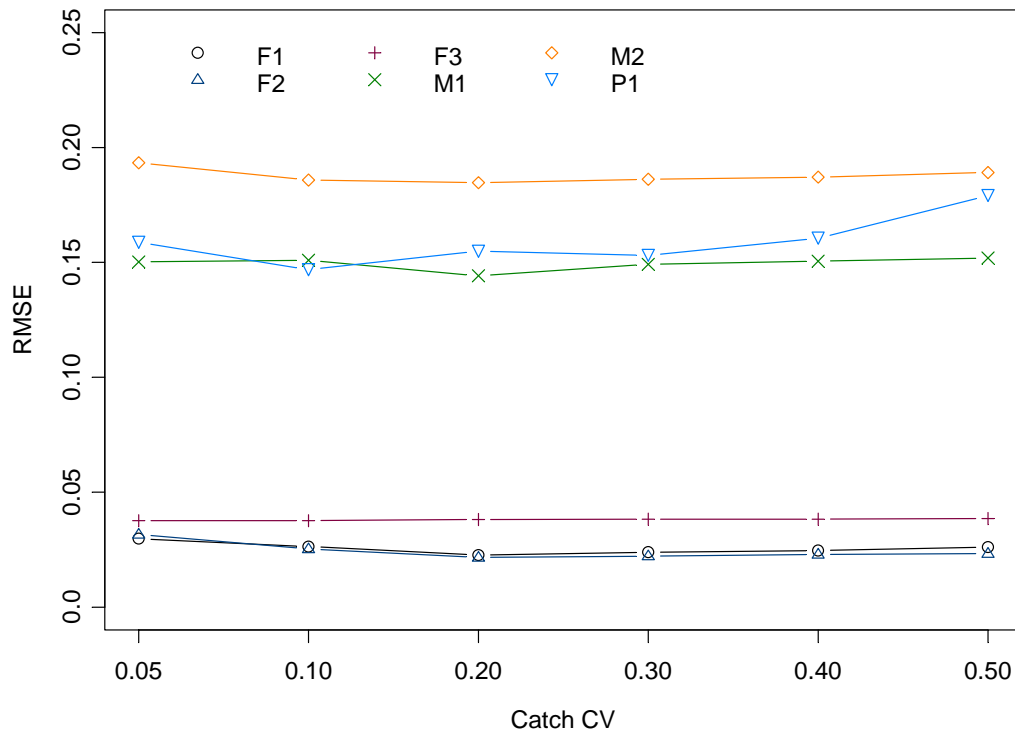
Figure 6. Negative log-likelihood profile for the coefficient of variation (CV) of the catch data for southern bluefin tuna.
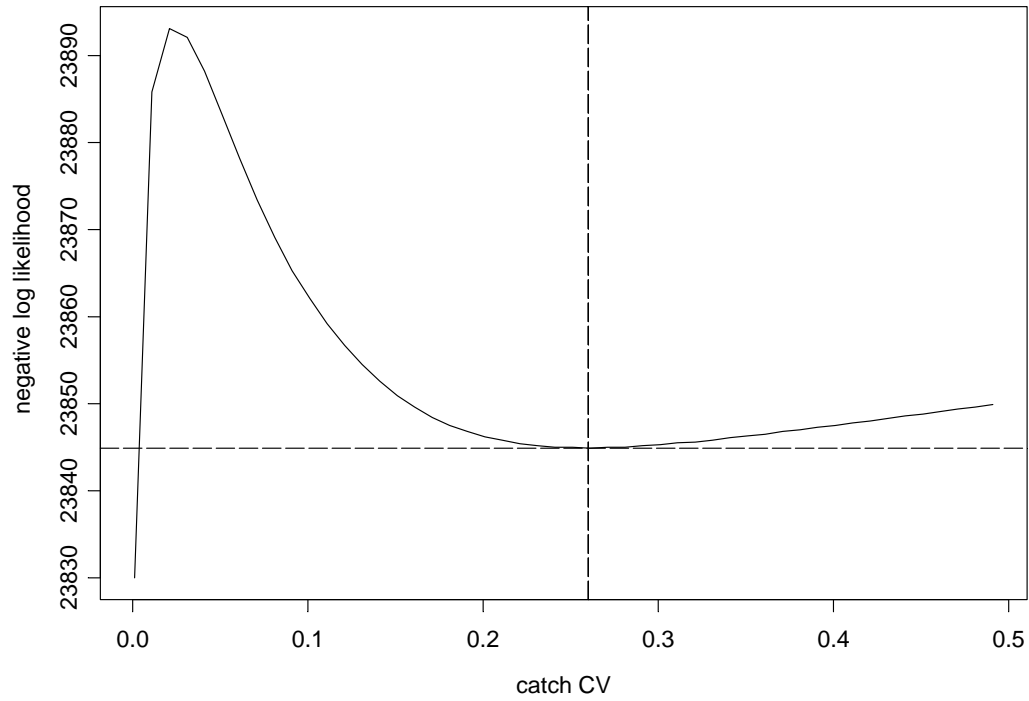
Appendix 6:

Estimating the bias in the estimator of population size in a simple tag-recapture and catch model

Geoff M. Laslett, J. Paige Eveson and Tom Polacheck

1   Introduction

In the simulations we conducted using the integrated tag-recapture and catch-at-age models presented in Appendices 5 and 7, we found that the estimator of population size tends to be negatively biased. In the current appendix, we examine a simpler case and show that this bias is expected.

2   A simple model

We assume that we have tag-recapture and catch data. We tag n fish and recapture m. The probability of capture, p, is to be estimated. We assume that m follows a Binomial(n, p) distribution. Simultaneously, we have N fish in the population, and catch M. We need to estimate N. We assume that

$$M \sim \text{Gaussian}(Np, \sigma^2)$$

where σ^2 is known.

We use maximum likelihood to estimate the parameters. The log-likelihood is (ignoring additive constants)

$$m \log p + (n - m) \log(1 - p) - \tfrac{1}{2} \frac{(M - Np)^2}{\sigma^2} \, .$$

It is immediate that the maximum likelihood estimators are

$$\hat{p} = \frac{m}{n}$$
$$\hat{N} = \frac{Mn}{m}$$

For future reference, note that $\widehat{Np} = M$.

To check for bias, we simulated 100000 sets of data with $n = 1000$, $N = 100000$, $p = 0.5$ and $\sigma = 10000$. Note that the coefficient of variation of M is $\sigma/Np = 0.2$. We found the average estimates of p and N over all simulations to be at $\bar{\hat{p}} = 0.5$ and $\bar{\hat{N}} = 100166$. There appears to be a slight positive bias in the estimator of N, the opposite of what we see in our simulations.

Before moving on to the next section, where we introduce a slightly more realistic model, we offer a rough explanation of the bias seen in the above simulations. Strictly speaking $E(1/m) = \infty$, because there is a small probability that $m = 0$, but in practice we ignore this. For a general random variable X with mean μ and variance σ^2, the expansion

$$f(X) \approx f(\mu) + (X - \mu)f'(\mu) + \tfrac{1}{2}(X - \mu)^2 f''(\mu)$$

yields the approximation

$$E[f(X)] \approx f(\mu) + \tfrac{1}{2}\sigma^2 f''(\mu) \,.$$

Applying this to $f(x) = 1/x$ and $X = m$, we get

$$E[1/m] \approx 1/np + \tfrac{1}{2}2np(1 - p)/(np)^3 = 1/np + (1 - p)/(np)^2 \,.$$

Thus

$$E\left[\frac{1}{m}\right] \approx \frac{1}{np}\left(1 + \frac{(1 - p)}{np}\right)$$

and the factor $1 + (1 - p)/np \to 1$ as n increases. Hence, for large n,

$$E[\hat{N}] \approx N \,.$$

This theory predicts that for our simulations with $p = 0.5$ and $n = 1000$, the expected bias in $\hat{N} = Mn/m$ is a factor of about 1.001. This is roughly what we observed.

3   A slightly more complex model

We now assume the same model, except that σ^2 is not known, but the coefficient of variation, $c = \sigma/Np$, of M is known. This is closer to the actual situation in Appendices 5 and 7. The log-likelihood is now (apart from additive constants)

$$l = m \log p + (n - m)\log(1 - p) - \tfrac{1}{2}\frac{(M - Np)^2}{c^2 N^2 p^2} - \log N - \log p \,.$$

Expanding $(M - Np)^2$, we obtain

$$l = m \log p + (n - m) \log(1 - p) - \log p - \frac{1}{2c^2} \left(\frac{M^2}{N^2 p^2} - \frac{2M}{Np} + 1 \right) - \log N \,.$$

Differentiating l with respect to N and setting to 0, we obtain the equation

$$-\frac{1}{c^2} \left(-\frac{M^2}{x^2} + \frac{M}{x} \right) - 1 \;=\; 0 \qquad\qquad (1)$$

where $x = Np$. Solving for x, we have the two quadratic solutions

$$x = \frac{-M \pm \sqrt{M^2 + 4M^2 c^2}}{2c^2} \,.$$

Now $x > 0$ so we must adopt the unique solution

$$x = \frac{-M + \sqrt{M^2 + 4M^2 c^2}}{2c^2} = M \left(\frac{\sqrt{1 + 4c^2} - 1}{2c^2} \right) \,.$$

We now turn to the estimation of p. Differentiating l with respect to p and setting the result to 0, we obtain

$$
\begin{aligned}
\frac{\partial l}{\partial p} \;&=\; \frac{m}{p} - \frac{(n - m)}{(1 - p)} - \frac{1}{p} - \frac{1}{2c^2} \left(-\frac{2M^2}{N^2 p^3} + \frac{2M}{Np^2} \right) \\
&=\; \frac{m}{p} - \frac{(n - m)}{(1 - p)} - \frac{1}{p} - \frac{1}{c^2 p} \left(-\frac{M^2}{x^2} + \frac{M}{x} \right) \\
&=\; \frac{m}{p} - \frac{(n - m)}{(1 - p)} - \frac{1}{p} + \frac{1}{p} \\
&=\; \frac{m}{p} - \frac{(n - m)}{(1 - p)} \\
&=\; 0
\end{aligned}
$$

where the last term in the second line has been simplified using (1). Hence

$$\hat{p} = \frac{m}{n}$$

and

$$\hat{N} = \frac{Mn}{m} \left(\frac{\sqrt{1 + 4c^2} - 1}{2c^2} \right) \,.$$

We see that, compared with the known variance model, the known coefficient of variation model has left $\hat{p}$ unchanged, but has introduced a factor

$$\frac{\sqrt{1 + 4c^2} - 1}{2c^2}$$

into $\hat{N}$.

What is the effect of the factor? If c is small, we use $\sqrt{1+\epsilon} \approx 1 + \frac{1}{2}\epsilon$ to obtain

$$\frac{\sqrt{1+4c^2}-1}{2c^2} \approx \frac{(1+2c^2)-1}{2c^2} = 1 \,.$$

However, in general, $1 + 4c^2 \leq 1 + 4c^2 + 4c^4 = (1+2c^2)^2$, so that

$$\frac{\sqrt{1+4c^2}-1}{2c^2} \leq \frac{(1+2c^2)-1}{2c^2} = 1 \,.$$

Hence $\hat{N}$ in the known coefficient of variation model is always smaller than $\hat{N}$ in the known variance model. We typically use $c = 0.2$, in which case the factor

$$\frac{\sqrt{1+4c^2}-1}{2c^2} = 0.962 \,.$$

This is a reduction of almost 4%.

For the same simulations as used in section 2, we obtained the estimates $\bar{\hat{p}} = 0.5$ and $\bar{\hat{N}} = 96451$.

The bias in $\hat{N}$ for the constant coefficient of variation model does not disappear as sample size increases. Using the same approach as in section 2, we obtain the results that, for large n,

$$\mathrm{E}[\hat{N}] \approx N \left(\frac{\sqrt{1+4c^2}-1}{2c^2} \right) \,,$$

and the bias factor does not disappear as n increases. We can, however, correct the bias quite trivially by multiplying $\hat{N}$ by the inverse of the bias factor.

Where is the bias coming from? We explore this in the next section.

4   Aside: the Gaussian distribution with known coefficient of variation

Consider a sample $X_1, \ldots, X_n$ from a Gaussian distribution $N(\mu, (c\mu)^2)$, where c is known, and μ is to be estimated. It is easily shown that the maximum likelihood estimator is

$$\hat{\mu} = \frac{\sqrt{4c^2 S^2 + (1+4c^2)\bar{X}^2} - \bar{X}}{2c^2} \,,$$

where $\bar{X} = \sum_{i=1}^{n} X_i/n$ and $S = \sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/n}$. See Khan (1968).

For $n = 1, \ldots, 10$, we simulate 20000 sets of data from this distribution with $\mu = 10$ and $c = 0.2$, and record the mean of the maximum likelihood estimates, which we denote by $\bar{\hat{\mu}}$ (Table 1).

We can also caluculate the expected value of $\hat{\mu}$. From the fact that $\bar{X}$ and S are statistically independent, and that

$$\bar{X} \sim N(\mu, (c\mu)^2/n), \quad nS^2/(c\mu)^2 \sim \chi^2_{n-1}$$

we have

$$\mathrm{E}[\hat{\mu}] = \frac{\mathrm{E}\left[\sqrt{4c^2 S^2 + (1 + 4c^2)\bar{X}^2}\right] - \mu}{2c^2}$$

and

$$\mathrm{E}\left[\sqrt{4c^2 S^2 + (1 + 4c^2)\bar{X}^2}\right] = \int_0^\infty \int_0^\infty \sqrt{4c^4 \mu^2 v/n + (1 + 4c^2)x^2}\, \phi(x)dx\, g(v)dv$$

where $\phi(.)$ is the Gaussian density and $g(.)$ is the χ^2_{n-1} density. In theory, there is a tiny probability that $\bar{X} < 0$, but we have ignored this. For $n \geq 2$, we have computed this integral numerically. The case $n = 1$ must be treated separately:

$$\mathrm{E}[\hat{\mu}] = \frac{\sqrt{1 + 4c^2} - 1}{2c^2}\, \mu\,.$$

We have included the expected values of $\hat{\mu}$ for $n = 1, \ldots, 10$ in Table 1.

Table 1: Bias in the maximum likelihood estimator, $\hat{\mu}$, and in the expected
value of $\hat{\mu}$ in a $N(\mu, (c\mu^2))$ distribution with $\mu = 10$ and $c = 0.2$

n	1	2	3	4	5	6	7	8	9	10
$\bar{\hat{\mu}}$	9.62	9.84	9.89	9.90	9.92	9.94	9.95	9.95	9.97	9.97
$\mathrm{E}[\hat{\mu}]$	9.629	9.818	9.880	9.910	9.928	9.940	9.949	9.955	9.960	9.964

Khan (1968) states that the formula for $\hat{\mu}$ only applies for $n \geq 2$, but we have applied it for $n = 1$. He also states that $\hat{\mu} \to N(\mu, c^2\mu^2/n(1 + 2c^2))$ as $n \to \infty$. Our simulations suggest that, although $\hat{\mu}$ might be asymptotically unbiased, for small n

it shows a distinct negative bias. The case of small n is of interest to us.

Gleser and Healy (1976) have investigated Bayesian point estimators for μ. They use inverse gamma priors (that is, the density of the prior is that of a random variable $1/G$ where G has a gamma distribution), and their preferred point estimator is the expectation of the posterior distribution. This is a one parameter problem, so the posterior is easy to compute. Their approach may not be easy to generalise.

5 Correcting the bias

It is unlikely that in more complex and realistic situations, we can obtain a formula for either $\hat{N}$ or $\mathrm{E}[\hat{N}]$. If we wish to correct the bias for more complicated models such as those in Appendices 5 and 7, we could estimate the parameters by maximum likelihood, simulate many sets of data to determine the size of the bias for the estimated parameter set, and then correct the bias.

An alternative to simulation is to adapt the idea of Khan (1968). We first estimate the parameters by maximum likelihood, then, using the maximum likelihood estimates of the probabilities of capture, we compute an unbiased estimate of the population size.

We first review Khan (1968). Suppose that a set of independent random variables $X_1, \ldots, X_n$,

$$X_i \sim N\left(\mu, (c\mu)^2\right),$$

where c is known. Let

$$d_1 = \bar{X}$$

and

$$d_2 = k\sqrt{\sum_{i=1}^{} (X_i - \bar{X})^2}$$

where

$$k = \frac{1}{\sqrt{2}c} \frac{\Gamma((n-1)/2)}{\Gamma(n/2)}.$$

Khan (1968) notes that, for $0 < \alpha < 1$, $d = \alpha d_2 + (1 - \alpha)d_1$ is an unbiased estimator of μ, and chooses α to mimimise the variance of d.

Now consider the usual 'regression through the origin' case of

$$X_i \sim N(\mu p_i, \sigma_i^2) \, .$$

The maximum likelihood estimator of μ is

$$\hat{\mu} \;=\; \frac{\sum_{i=1}^n X_i/\sigma_i^2}{\sum_{i=1}^n p_i/\sigma_i^2} \, . \tag{2}$$

If the σ_i^2 values are all equal, we get

$$\hat{\mu} \;=\; \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n p_i} \, . \tag{3}$$

Now suppose we have a situation closer to the catch model used in Appendices 5 and 7; that is, suppose we have a set of independent random variables $X_1, \ldots, X_n$, where

$$X_i \sim N\!\left(\mu p_i, (c\mu p_i)^2\right) ,$$

and c and $p_1, \ldots, p_n$ are known. We could follow Khan (1968) and try to use information on μ in the variance directly in the estimator; however, this may be dangerous because we do not really know the coefficient of variation c, and d_2 depends on it. Alternatively, we could use (2), which reduces to

$$\hat{\mu} \;=\; \frac{\sum_{i=1}^n X_i/p_i^2}{\sum_{i=1}^n 1/p_i} \, ?$$

This is also potentially dangerous because, in our case, the p_i values (i.e., the probabilities of capture) are estimated rather than known, and some p_i may be small. For the sake of robustness, we recommend using (3). Out of all unbiased estimators of the form $\sum_{i=1}^n X_i \, p_i^\nu / \sum_{i=1}^n p_i^{\nu+1}$ for a given real ν, it has the weakest dependence on the p_i's, and then they occur only as a linear sum. We can then estimate the standard error of (3) by

$$\text{s.e.}[\hat{\mu}] \;=\; c\mu \frac{\sqrt{\sum_{i=1}^n p_i^2}}{\sum_{i=1}^n p_i} \, .$$

Of course, in our case the p_i must be estimated, so we just substitute the $\hat{p}_i$ into (3), and then $\hat{p}_i$ and $\hat{\mu}$ into the formula for the standard error.

6 An alternative to the Gaussian: the negative binomial

We assume the same model as in sections 2 and 3, except that

$$M \sim \text{Negative Binomial}(P, k)$$

where $k > 0$ is known and $P > 0$ is to be estimated. Thus M has the discrete distribution

$$\Pr\{M = \nu\} = \begin{pmatrix} k + \nu - 1 \\ k - 1 \end{pmatrix} \left(\frac{P}{Q}\right)^{\nu} \left(1 - \frac{P}{Q}\right)^{k}$$

with mean kP and variance $kP(1 + P)$. Here $Q = 1 + P$ and $\nu = 0, 1, 2, \ldots$. In our situation, $kP = Np$, so we parameterise the negative binomial by keeping k and replacing P by Np/k.

The log-likelihood is (ignoring additive constants)

$$m \log p + (n - m) \log(1 - p) + M \log(Np) - (M + k) \log(Np + k) .$$

It is immediate that the maximum likelihood estimators are

$$\hat{p} = \frac{m}{n}$$
$$\hat{N} = \frac{Mn}{m}$$

Hence the properties of $\hat{N}$ are more like those of the model in section 2 ($\hat{N}$ almost unbiased) than the model in section 3 ($\hat{N}$ noticeably biased).

Let c be the coefficient of variation of a negative binomial with mean μ and parameter k. It is immediate that

$$c = \frac{\sqrt{\mu + \mu^2/k}}{\mu} = \sqrt{\frac{1}{\mu} + \frac{1}{k}} .$$

Suppose for the moment we can assume that $\mu \gg k$. Then $c \approx 1/\sqrt{k}$ and the coefficient of variation of the negative binomial is virtually constant for a fixed k and varying μ. For $c = 0.2$, we should choose $k = 25$. Hence, provided $\mu \gg 25$ (say, $\mu > 1000$, which seems realistic in our case) we have a model with virtually constant coefficient of variation and an estimator of N with good properties.

7 References

GLESER, L.J. & HEALY, J.D. (1976) Estimating the mean of a normal distribution with known coefficient of variation. *Journal of the American Statistical Association* **71,** 977-981.

KHAN, R.A. (1968) A note on estimating the mean of a normal distribution with known coefficient of variation. *Journal of the American Statistical Association* **63,** 1039-1041.

A7-1

Appendix 7:

An integrated Brownie and Petersen model for estimating mortality rates and population size in a fisheries context with unknown reporting rates

J. Paige Eveson, Tom Polacheck and Geoff M. Laslett

FRDC Project 2002/015

Introduction

In Appendix 5 we developed a maximum likelihood tag-recapture model that combines a traditional Brownie approach with a Petersen-type approach to provide joint estimates of both abundance and mortality rates. This was done in a fisheries context in which catch data were used to provide estimates of the number of animals examined for tags and the uncertainty in the catch data was incorporated explicitly. We demonstrated that the addition of catch data to the Brownie tagging model not only allows for estimation of abundance but also improves estimation of mortality rates.

The integrated tagging and catch model developed in Appendix 5 assumes that all recaptured tags are reported. When recapture information comes from commercial fisheries, we do not expect the reporting rates to be 100% nor do we expect them to be known; thus, they must be estimated. Although tag-recapture data by themselves contain some information about reporting rates, the information is generally weak and insufficient to distinguish non-reporting from natural mortality and fishing mortality without making some fairly restrictive assumptions (see Hoenig et al. 1998). Auxiliary data for estimating reporting rates can be obtained through a variety of methods (Pollock et al. 2001 gives a brief review of the possible methods). The data then needs to be incorporated into the tagging model. For some methods, such as tag seeding experiments, the data for estimating reporting rate estimates can be considered independent of the tagging and related catch data. In such cases, a likelihood can be developed for the data and simply multiplied to the tagging and catch likelihoods developed in Appendix 5 (we illustrate a model using tag seeding data in Appendix 12). For other methods, such as placing observers on vessels, the reporting rate estimates use some of the tagging and/or related catch data. In such cases, the likelihoods for the tagging and catch data need to be modified.

In the current appendix, we expand the integrated tagging and catch model developed in Appendix 5 to incorporate the estimation of reporting rates for the situation in which observers monitor a portion of the catches. We first describe the expanded model, detailing the modifications to the tag-recapture and catch likelihoods that are necessary to incorporate observer data. We then apply the model to simulated data to investigate experimental design issues; in particular, we explore the effect of releasing more tags and increasing observer coverage on the precision and bias of parameter estimates. Resources are usually limited so the question of whether to devote more resources to tagging fish or observing catches is important to address. Finally, we conclude with some general discussions.

Methods

Basic dynamic model

Underlying the analysis of multi-year tagging experiments used here are the general population dynamic equations commonly used in fisheries. The equations were presented in Appendix 5 but are repeated here for readability. For a cohort of animals of a given age, the number that survives to the next age and the number caught are given by:

$$P_{i+1,t+1} = P_{i,t} \exp\left(-F_{i,t} - M_{i,t}\right) \tag{1}$$

$$C_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} P_{i,t} \left(1 - \exp\left(-F_{i,t} - M_{i,t}\right)\right) \tag{2}$$

where
$P_{i,t}$ = the number of individuals of age i at time t
$C_{i,t}$ = the catch of individuals of age i at time t
$F_{i,t}$ = the instantaneous fishing mortality rate for individuals of age i at time t
$M_{i,t}$ = the instantaneous natural mortality rate for individuals of age i at time t.

These equations are expressed in terms of instantaneous and competing fishing and natural mortality rates, and assume that deaths due to fishing and natural causes occur coincidently at constant rates throughout the year. The model can easily be expressed in terms of alternative formulations, such as annual harvest rates from a pulse fishery (see Hoenig et al. 1998).

In the context of a tagging experiment, the above equations provide the basis for predicting the expected number of tagged fish in the population and tag returns by age and year, assuming that the tagged fish constitute a representative sample of the population. Analogous to equations (1) and (2), we have:

$$N_{i+1,t+1} = N_{i,t} \exp\left(-F_{i,t} - M_{i,t}\right) \tag{3}$$

$$R_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} N_{i,t} \left(1 - \exp\left(-F_{i,t} - M_{i,t}\right)\right) \tag{4}$$

where
$N_{i,t}$ = the expected number of tagged individuals of age i at time t
$R_{i,t}$ = the expected number of tag returns from fish captured at age i at time t
$\lambda_{i,t}$ = the tag reporting rate for fish captured at age i at time t

As in Appendix 5, we will focus on a multi-year tagging experiment involving a single cohort. For a given cohort, year and age provide equivalent information so we can simplify the notation by dropping the t subscript and expressing everything in terms of age i. The generalization of the model to more than one cohort is straightforward.

Estimation Model with Reporting Rates Estimated From Observer Data

We develop an estimation model for a multi-year tagging experiment in which fish from a single cohort have been tagged in consecutive years. Fish are subsequently caught in a fishery and a percentage of the tags that are recaptured are returned. Observers monitor catches on a portion of the fishing vessels, and we assume that 100% of recaptured tags are returned from the observed component of the fishery. We also assume that, on average, fishing mortality is the same for all vessels (observed and unobserved), so the relative return rate between observed catches and

unobserved catches provides an estimate of the reporting rate in the unobserved component. Data required for the model are the numbers of fish caught by age in the observed and unobserved components and the number of tags returned by age in the observed and unobserved components. We develop a joint likelihood model for the tag-return and catch data that is analogous to the model in Appendix 5 except we modify the likelihoods to incorporate observer information for estimating reporting rates.

Before proceeding we introduce the following notation:

Table 1. Data (to be inputted into the model).

N_a = the number of tag releases of age a fish from a particular cohort

$R_{a,i}^o$ = the number of tag returns of age i fish that were tagged at age a from the observed component of the fishery

$R_{a,i}^u$ = the number of tag returns of age i fish that were tagged at age a from the unobserved component of the fishery

C_i^o = the number of age i fish from the cohort of interest caught in the observed component of the fishery

C_i^u = the number of age i fish from the cohort of interest caught in the unobserved component of the fishery

δ_i = the proportion of age i fish from the cohort of interest in the observed component of the fishery

υ_i = the coefficient of variation of age i catch data

Table 2. Parameters (to be estimated in the model).

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 at age 1 of the tagged cohort

λ_i = the tag reporting rate for fish captured at age i

Now define:

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

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

S_i represents the annual survival rate of age i fish and u_i represents the annual exploitation rate of age i fish.

First consider the tag-recapture component of the model. The probability of a tag being returned from the observed component of the fishery from an age i fish released at age a is:

$$p_{a,i}^{o} = \begin{cases} \delta_i u_i & i = a \\ \delta_i S_a \cdots S_{i-1} u_i & i > a \end{cases} \tag{5}$$

Similarly, the probability of a tag being returned from the unobserved component of the fishery from an age i fish released at age a is:

$$p_{a,i}^{u} = \begin{cases} (1 - \delta_i) u_i \lambda_i & i = a \\ (1 - \delta_i) S_a \cdots S_{i-1} u_i \lambda_i & i > a \end{cases} \tag{6}$$

For tags released at a particular age, the numbers of returns at age from the observed component and from the unobserved component (as well as those not returned from either component) are expected to be multinomial with probabilities given in (5) and (6). Thus, the likelihood equation for all the returns is:

$$L_R = \gamma \times \prod_a \left(\left(1 - p_{a,\bullet}^{o} - p_{a,\bullet}^{u} \right)^{N_a - R_{a,\bullet}^{o} - R_{a,\bullet}^{u}} \prod_{i \geq a} \left(p_{a,i}^{o}{}^{R_{a,i}^{o}} p_{a,i}^{u}{}^{R_{a,i}^{u}} \right) \right) \tag{7}$$

where

$$\gamma = \prod_a \frac{N_a!}{\left(N_a - R_{a,\bullet}^{o} - R_{a,\bullet}^{u} \right)! \prod_{i \geq a} \left(R_{a,i}^{o}! \, R_{a,i}^{u}! \right)}$$

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

Next consider the catch component of the model. We assume that all fish caught in the observed component of the fishery are sampled for length and/or age, but that no fish from the unobserved component are sampled. Thus, no age information is available for the unobserved catches, and only catch-at-age data from the observed component can be included 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:

$$\pi_i^{o} = \begin{cases} \delta_i u_i & i = 1 \\ \delta_i S_1 \cdots S_{i-1} u_i & i > 1 \end{cases} \tag{8}$$

If each fish has an equal probability of being caught and the numbers of fish caught at each age in the observed component of the fishery are known accurately, then the observer catch-at-age data would be random multinomial with each fish in the cohort having a probability of being captured at age i (given by equation (8)) or else not captured. Usually, however, all fish do not have an equal probability of being caught because of non-homogeneous spatial and temporal distributions, as well as different size/age selectivities among vessels. Furthermore, the numbers of fish caught at each age are usually not known precisely because the ages are estimated either from lengths or from direct aging of hard parts. Following the approach in Appendix 5, we have chosen to model the error in the catch-at-age data as Gaussian; however, another distribution could be have been used if it was considered more appropriate.

Specifically, we assume that the observed catch of age i fish, C_i^o, has a Gaussian distribution with coefficient of variation (CV) υ_i. Thus, the likelihood for the observed catch data is:

$$L_C^* = \prod_i \frac{1}{\sqrt{2\pi}\left(\upsilon_i E(C_i^o)\right)} \exp\left(-\frac{1}{2}\left(\frac{C_i^o - E(C_i^o)}{\upsilon_i E(C_i^o)}\right)^2\right) \tag{9}$$

where $E(C_i^o) = P_1 \pi_i^o$.

The CV of the catch data will in part be determined by the level of observer coverage since this determines the level of sampling. We would expect the CV to decrease as the level of observer coverage increases but the exact relationship will depend on the situation.

The overall likelihood for the combined recapture and catch data can be obtained by multiplying likelihoods (7) and (9) together:

$$L^* = L_R^* \times L_C^* \tag{10}$$

In the current formulation with just one cohort, there is not enough information to estimate the proportion of fish of each age in the observed component of the fishery (i.e. the δ's). To do so we would need to know 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 assume that the total catches are known well enough that the δ's can be treated as known without error.

The reporting rates (λ's) as well as the F, M and P parameters are all estimated by maximizing (10). As explained in Appendix 5 for the model without reporting rates, in a tagging experiment with n consecutive release years, estimates can only be obtained for $n-1$ of the natural mortality rate parameters (regardless of the number of recapture years). Thus we assume that $M_i = M_{n-1}$ for $i \geq n$. Furthermore, the catch CV parameter υ is not estimated reliably so we assume that it is known without error.

Simulation Results

Trade-off between number of releases and observer coverage

When reporting rates are estimated from observer information, we can look at the trade-off between increasing the number of releases and increasing the level of observer coverage on the precision of the parameter estimates. Increasing observer coverage will improve the parameter estimates by improving the reporting rate estimates and also by improving the precision of the catch-at-age data. To best evaluate the trade-off between tag releases and observer coverage, we assumed that the number of releases was the same for all release ages/years (i.e. $N_a = N$ for all a) and that the level of observer coverage was the same over all recapture ages/years (i.e. $\delta_i = \delta$ for all i).

We carried out simulations in which we varied the number of releases (N) from 250 to 2000 and the level of observer coverage (δ) from 0.05 to 0.5. For each combination of N and δ, we generated 500 simulated tag-recapture and corresponding catch data sets. To generate the data, we assumed three consecutive release years and three recapture years. We set the fishing mortality rate to be 0.15 and the natural mortality rate to be 0.2 for all ages (i.e. $F_i = 0.15$ and $M_i = 0.2$ for all i), we set the reporting rate in the unobserved component to be 0.25 for all ages (i.e. $\lambda_i = 0.25$ for all i), and we set the population size at age 1 to be 100000 (i.e. P_1 =100000). The coefficient of variation used to generate the catch data was determined by the level of observer coverage according to the hypothetical relationship $\upsilon = 0.75 * (0.05)^{\sqrt{\delta}}$ (illustrated in Figure 1). Note that even with 100% observer coverage, the CV does not to go to zero because there will still be variability in the catch process and errors in the estimated ages. Assuming this relationship, the levels of observer coverage that we considered of 0.05, 0.1, 0.2 and 0.5 correspond to catch CV's of 0.38, 0.29, 0.20 and 0.09 respectively. Although this relationship is rather arbitrary, it seemed reasonable for our investigative purposes.

For each of the 500 simulated data sets, we obtained parameter estimates of F_1, F_2, F_3, M_1, M_2, λ_1, λ_2, λ_3 and P_1 by maximizing the joint tagging and catch likelihood given in (10). Recall that we are assuming υ is known without error and that we constrain natural mortality to be the same for ages 2 and above (by necessity since we have only three release years). We then looked at how the precision of the parameter estimates (as measured by the CV of the estimates) changed as the number of releases and level of observer coverage changed (Figure 2).

First concentrate on the fishing mortality rate results (Figure 2a-c). Increasing the number of releases increased the precision (i.e. decreased the CV) of the fishing mortality rate estimates; the improvement was greater at low release numbers and also for older ages (i.e. for F_3 than F_1). This makes sense because if the number of tag releases is quite small then there would be relatively few tag returns from older fish, as many of the tagged fish would have died earlier due to natural mortality or fishing. Increasing the level of observer coverage also improved the precision of the fishing mortality rate estimates; the gain in improvement was largest at low levels of observer coverage (i.e. the F estimates improved more when observer coverage increased from 0.05 to 0.1 than when it increased from 0.2 to 0.5). There was also an interaction between the two variables in which the improvement in the F estimates that resulted from increasing the level of observer coverage diminished as the number of releases increased, and vice versa. In general, larger gains in the precision of the F estimates could be achieved by increasing the level observer coverage from 0.05 to 0.5 than from increasing the number of releases from 250 to 2000, especially for the age 1 and 2 estimates.

Next consider the natural mortality rate estimates (Figure 2d-e). The precision of the M_1 and M_2 estimates was significantly improved both by increasing the number of releases and by increasing the level of observer coverage. There was little interaction between the two variables, as indicated by the almost parallel lines in the figures. Contrary to the F estimates, larger gains in the precision of the M estimates could generally be achieved by increasing the number of releases from 250 to 2000 than

from increasing the level of observer coverage from 0.05 to 0.5. Relative to the other parameters, M_1 and M_2 were estimated quite poorly – even in the best case scenario with 2000 releases and observer coverage of 0.5, the CV of both M estimates is over 0.5, and in the worst case scenario with 250 releases and observer coverage of 0.05, the CV's are over 1.0.

For the population size parameter, P_1, increasing the level of observer coverage resulted in fairly large increases in the precision of the parameter estimate (Figure 2f). Increasing the number of releases also improved the estimates of P_1, especially at low release numbers, but the improvement was generally not as large as could be achieved by increasing the level of observer coverage. There was also a small interaction between the two variables in which increasing the level of observer coverage improved the precision of the P_1 estimate less as the number of releases increased, and vice versa.

While the reporting rate estimates are not of primary interest, these results are still shown for completeness (Figure 2g-i). The precision of the reporting rate estimates improved significantly as the number of releases increased, and the improvement was similar for all ages (i.e. for λ_1 to λ_3). Increasing the level of observer coverage from 0.05 to 0.2 also led to large improvements in the λ estimates, but the precision of the estimates actually started to degrade with further increases (from 0.2 to 0.5). This is because the reporting rate is estimated not only from the number of tags returned from the observed component but also from the unobserved component. When the level of observer coverage is high, the number of returns from the unobserved component will be small (especially if the reporting rate is low) and therefore more uncertain.

Effect of reporting rate on parameter estimates

Throughout the trade-off simulations we assumed a constant reporting rate of 0.25. To explore the effect of changing the reporting rate on the results, we re-ran the full set of simulations using a constant reporting rate value of 0.75 (Figure 3). As we would expect, all parameters were estimated more precisely when the reporting rate was high (0.75) than when it was low (0.25); this was true for any number of releases and level of observer coverage. However, the magnitude of the improvement depended on the parameter, as well as the number of releases and level of observer coverage. In particular:

1) The precision of the F estimates improved more at lower release numbers and, as a result, became less affected by the number of releases, especially for ages 1 and 2. Also, the interaction between the number of releases and the level of observer coverage lessened.
2) The precision of the M estimates improved more at lower levels of observer coverage and, as a result, became less affected by the level of observer coverage.
3) The precision of the P_1 estimate became almost unaffected by the number of releases due to large improvements at low release numbers.
4) The largest improvements in precision were for the reporting rate estimates, and the improvements were greatest at low release numbers.
5) The precision of all parameter estimates improved more at lower levels of observer coverage (as observer coverage decreases, the number of tagged fish caught in the unobserved component increases and the reporting rate has a larger effect).

Roughly speaking, increasing the reporting rate has a similar effect to increasing the number of releases because it results in more tag returns. The number of extra tag returns (and thus the effective increase in the number of releases) depends on the level of observer coverage – if the observer coverage is close to zero, then tripling the reporting rate in the unobserved component will result in almost triple the total tag returns (i.e. effectively a tripling of the number of releases), whereas if the observer coverage is close to one, then tripling the reporting rate in the unobserved component will have almost no effect on the total tag returns (i.e. effectively no change in the number of releases). Of course, it is not quite that simple because increasing the number of releases and increasing the reporting rate can lead to equivalent increases in the total number of returns, but the proportion of returns coming from the unobserved and observed components will not be affected the same (increasing the number of releases leads to proportional increases in tag returns in both components, whereas increasing the reporting rate only leads to increases in tag returns in the unobserved component).

Biases in parameter estimates

Until now we have only considered the precision of the parameter estimates and how it can be affected by the number of releases and level of observer coverage. In Appendix 5 we saw that biases were sometimes present in the parameter estimates obtained using the integrated tagging and catch model with known reporting rates; generally these biases were small and diminished as the number of releases increased. In the current appendix, the parameter estimates obtained using the integrated tagging and catch model extended to incorporate the estimation of reporting rates through observer data also contained biases in some situations. As in Appendix 5, the biases decreased as the number of releases increased, but the results are more complex because the biases also depend on the level of observer coverage. Furthermore, the nature and magnitude of the biases depended on the true reporting rate value (compare Figures 4 and 5).

For a constant reporting rate of 0.25 (Figure 4), the estimates of F_1 and F_2 were unbiased, and so were the estimates of P_1 (except with 250 releases). For all other parameters (i.e. F_3, M_1, M_2, λ_1, λ_2 and λ_3), the estimates showed some significant biases when both the number of releases the level of observer coverage were low, but these biases diminished as either variable increased. In particular, when the observer coverage was 0.5, the biases were less than 5% for all parameters with any number of releases over 500. Conversely, when the number of releases was 1500 or more, the biases were less than 5% for all parameters regardless of the level of observer coverage (except for λ_1 with observer coverage of 0.05).

When the true reporting rate was increased from 0.25 to 0.75, the bias results changed in somewhat unpredictable ways (Figure 5). We would expect the biases to decrease, and for some parameters this was true. Namely, the λ estimates became unbiased, plus the biases in the M estimates mostly disappeared (note that these were the parameters that had the largest biases with a reporting rate of 0.25). On the contrary, biases were introduced in the F_1 and F_2 estimates. Although the biases were small (generally <10%) and diminished as either the number of releases or the level of observer coverage increased, they are not intuitive. In addition, the P_1 estimates became negatively biased at low levels of observer coverage. In our simulations for

the model with known reporting rates in Appendix 5, we also obtained negative biases in the P_1 estimates when the CV of the catch data was high (which corresponds to low levels of observer coverage in the current model). A theoretical explanation for negative bias in P_1 is outlined in Appendix 6 for a simpler model that uses only total releases, recaptures and catches; the bias stems from the fact that the variance of the Gaussian catch data is being modelled in terms of a CV, and therefore as a function of the mean. It would seem reasonable that the arguments made in Appendix 6 could be extended to explain the negative biases in the P_1 estimates obtained using the more complex multi-year tagging and catch model, both with known reporting rates (Appendix 5) and with estimated reporting rates (current appendix). However, it is unclear why in the model with estimated reporting rates the biases were only seen when the true reporting rate was 0.75 and not when the true reporting rate was 0.25.

In Appendix 5 we saw that the biases in the F estimates increased with age for the model with known reporting rates. The same holds true for the model with reporting rates estimated from observer data. We re-ran the tradeoff simulations using five recapture years instead of three, and again found that the biases in the F estimates increased exponentially with age and were non-trivial by age 5. For example, using a true reporting rate of 0.25, the bias in F_5 was over 60% when the number of releases was 250 and the observer coverage was 0.1 or less. Although the bias diminished as the number of releases and level of observer coverage increased, it only became less than 5% when the number of releases was at least 1750 and the observer coverage was 0.5. As discussed in Appendix 5, the biases appear to stem from the fact that natural mortality and fishing mortality are highly correlated and we are constraining natural mortality to be the same for ages 2 and above. Recall that this constraint was required because with three release years, the model can only provide estimates of two natural mortality rate parameters, regardless of the number of recapture years.

Discussion

In the current appendix we have extended the integrated tagging and catch model developed in Appendix 5 to incorporate the estimation of reporting rates through observer data. When tag returns come from a commercial fishery it is almost certain that reporting rates will not be 100% and unlikely that they will be known; therefore, it is important to have a method of estimating them and incorporating the estimates into the broader tag-recapture model. Although several methods can be used for estimating reporting rates, the one we focussed on in this appendix is the use of observers in a portion of the fishery to monitor catches and ensure tag returns.

Tagging and observer programs can be costly to run and resources available for doing so are usually limited. Thus, we applied our model to simulated data to explore the effect of releasing more tags versus increasing observer coverage on the precision and bias of the parameter estimates. In particular, we were interested in the estimates of natural mortality, fishing mortality, and abundance; reporting rate estimates are also provided by the model but these are usually of secondary interest. We found that increasing the number of releases improved the precision of all parameters, especially the fishing mortality estimates at older ages and the natural mortality estimates. The improvement tended to be greatest at low release numbers such that increasing the number of releases beyond 1000 resulted in only marginal gains for most parameters

(natural mortality rates being a bit of an exception). Increasing the level of observer coverage also improved the precision of all parameters, especially the fishing mortality estimates and the population size estimate. Although the improvement per unit increase in observer coverage was greatest at low levels (i.e. in going from 0.05 to 0.1), substantial gains in precision were still achieved in all parameters (except reporting rates) by increasing observer coverage from 0.2 to 0.5. The results depended on the reporting rate value assumed, with the precision of the parameter estimates being greater when the reporting rate was high; however, the above general observations held true regardless. With regard to biases, any biases in the parameter estimates diminished as the number of releases increased and as the level of observer coverage increased. However, biases were not generally an issue of concern except when both the number of releases was very low (500 or less) and the levels of observer coverage was low (0.1 or less). Even then, in the case where the reporting rate was high, the biases with 250 releases and observer coverage of 0.05 were never more than 12% for any of the parameter estimates. An unexpected observation regarding biases was that the age 1 and 2 fishing mortality estimates and the population size estimate showed almost no signs of biases in the simulations with a reporting rate of 0.25 but did show small biases in the simulations with a reporting rate of 0.75; the reason for this is unclear.

Although we would expect our general observations to remain the same, the results will depend on the relationship assumed between the level of observer coverage and coefficient of variation of the catch data. For example, if we had assumed a flatter relationship where increasing observer coverage did not reduce the variability in the catch data as much, then the level of observer coverage would have had less effect on the precision and bias of the parameter estimates. As an extreme case, we could assume that the CV of the catch data is independent of the level of observer coverage, in which case increasing the level of observer coverage would only improve the parameter estimates through improving the reporting rate estimates; however, it is difficult to envisage a situation where this would be plausible. A point to consider in determining an appropriate relationship is that it will depend on the nature of the species and the fishery. If fish are not distributed homogeneously in space or time, or if there is large variability in the size/age selectivities of vessels, then the reduction in the catch variability from increasing the observer coverage will depend on how the increased coverage is achieved. If all of the additional observer data comes from only a few vessels or cruises, then the gain will be much less than if it comes from a large number of vessels or cruises operating over a wide geographic and temporal range.

The model we presented assumes complete mixing of tagged and untagged fish and that the fate of each tagged fish is independent of the fate of other tagged fish (i.e. every fish has an equal probability of being caught). It also assumes no mortality due to tagging (either immediate or subsequent) and no tag shedding. These assumptions are common to most tag-recapture models and the consequences of violating them as well as ways of testing for violations have been discussed at length in the literature (see Pollock et al. 2001 and references given therein). As such, a full discussion will not be repeated here; however, we do note that if any of the assumptions are violated, then the estimates presented on the precision and bias of the parameter estimates would be overly optimistic. Furthermore, increasing the number of releases may have more benefit than the results presented suggest. For example, if substantial tag shedding or tagging mortality exists, then having larger numbers of tag releases would

be necessary to obtain reasonable return numbers. Also, if incomplete mixing or non-independence between fish due to, say, high school fidelity is an issue, then tagging more fish throughout a wider geographic and temporal range could be very beneficial.

The simulation results presented were for a tagging experiment involving a single cohort. In practice, it is likely that two or more cohorts would be tagged in a given year. If any parameters are thought to be the same between cohorts, then tagging multiple cohorts can improve the information available for estimation and inference about not only these parameters, but the other model parameters as well. We ran some multi-year, multi-cohort simulations allowing for the fishing mortality rates and reporting rates to differ by year and age but constraining the natural mortality rates to differ only by age, not year (and still differ only between age 1 and age 2 plus). Having additional cohorts resulted in large improvements in the precision of the natural mortality rate estimates and also in the fishing mortality rate estimates at older ages. For example, we ran simulations using two and three cohorts, and for each cohort we assumed the same experimental design and parameter values as in our trade-off simulations for one cohort (i.e. 3 release years and 3 recapture years, F_i = 0.15, M_i = 0.2, and λ_i = 0.25 for all ages i, and P_1 =100000). With 1000 releases for each age and cohort and with observer coverage of 0.1 in all years, we found that having data from two cohorts reduced the CV of the M_1 and M_2 estimates by 0.18 and 0.20 respectively, and having data from three cohorts reduced the CV of both the M_1 and M_2 estimates by 0.28. For the F_3 estimates, having either two or three cohorts reduced the CV by about 0.07-0.09. There was very little difference in the precision of the F_1, F_2 and P_1 estimates regardless of the number of cohorts.

In summary, the simulation results presented may provide useful insight into design issues for those starting up new or modifying current tagging and observer programs for the purposes of estimating mortality and abundance.

References

Hoenig, J.M., Barrowman, N.J., Hearn, W.S., and Pollock, K.H. 1998`. Multiyear tagging studies incorporating fishing effort data. Can. J. Fish. Aquat. Sci. **55**: 1466-1476.

Pollock, K.H., Hoenig, J.M., Hearn, W.S., and Calingaert, B. 2001. Tag reporting rate estimation: 1. An evaluation of the high-reward tagging method. N. Am. J Fish. Manage. **21**: 521-532.

Figure 1. The assumed relationship between level of observer coverage and the coefficient of variation (CV) of the catch-at-age data for the longline fishery.

Figure 2. Effect of changing the number of releases and level of observer coverage on the **coefficient of variation** (CV) of the fishing mortality rate estimates (a-c), mortality rate estimates (d-e), initial population size estimate (f), and reporting rate estimates (g-i). All results obtained using 500 simulations and a constant **true reporting rate of 0.25**.

a) F_1

b) F_2

c) F_3

d) M_1

e) M_2

f) P_1

g) λ_1

h) λ_2
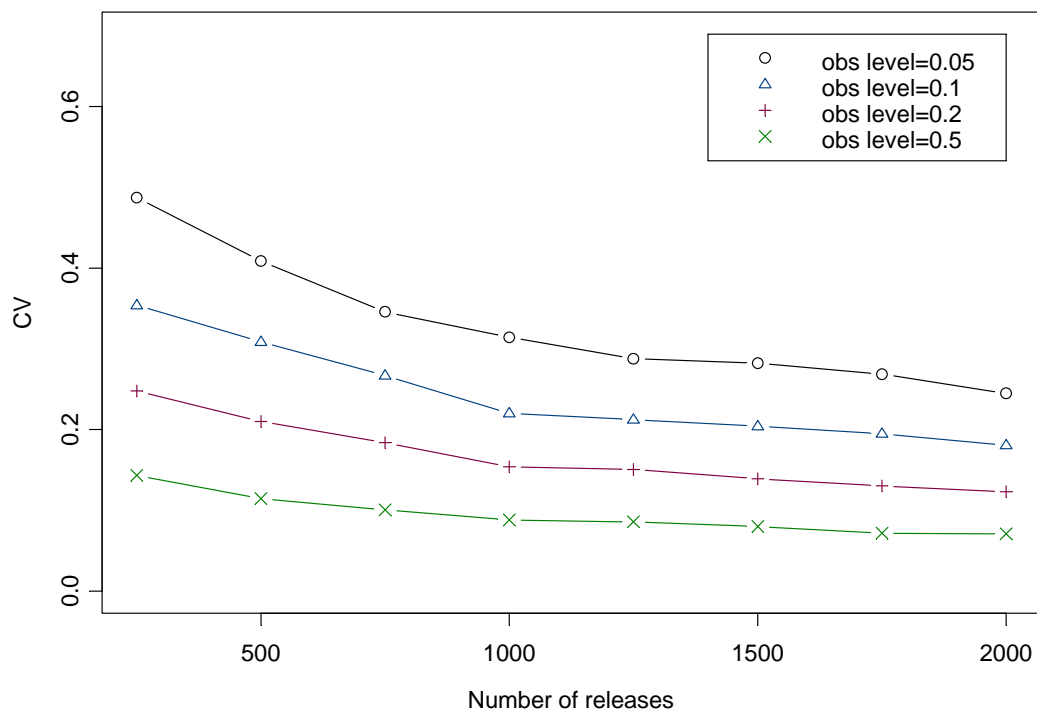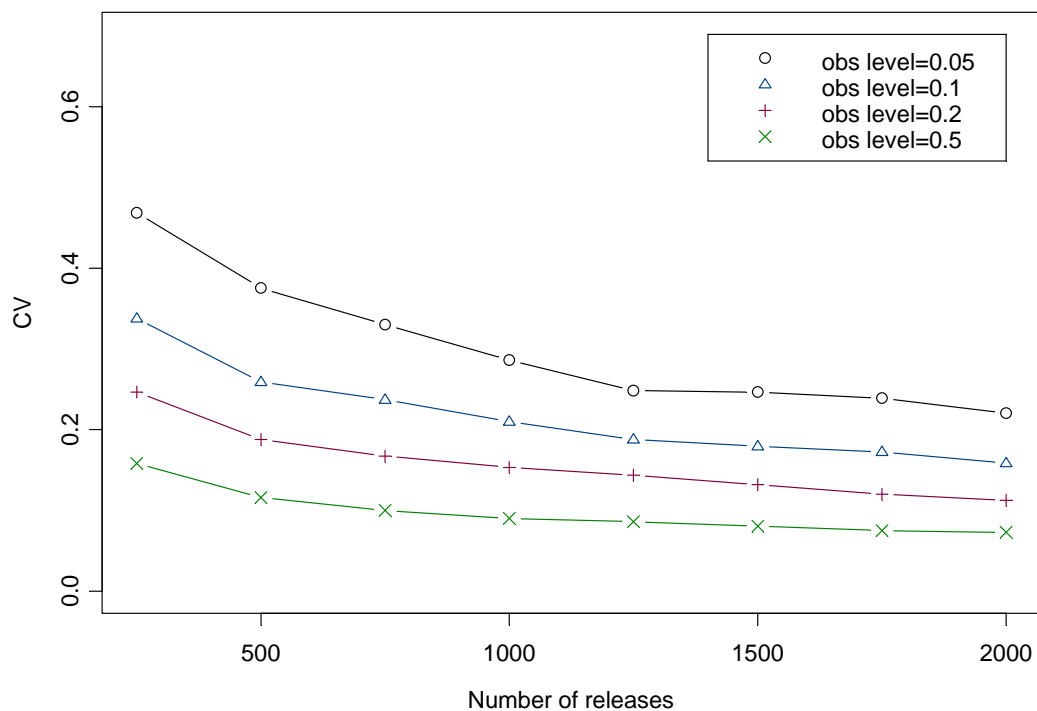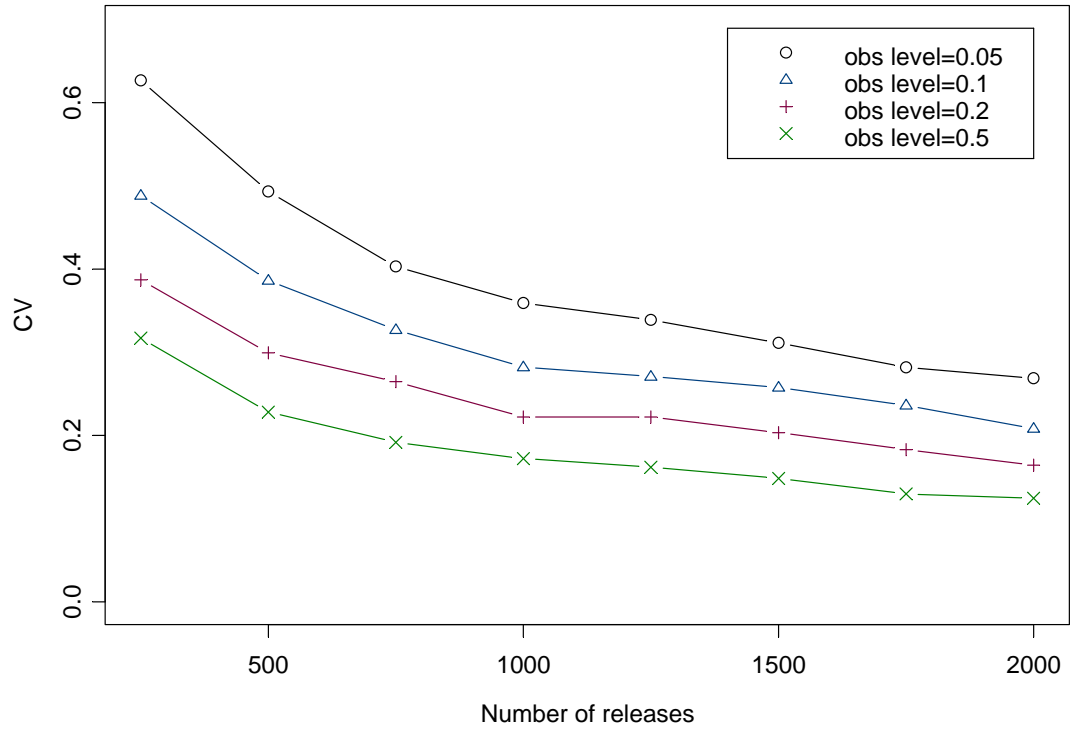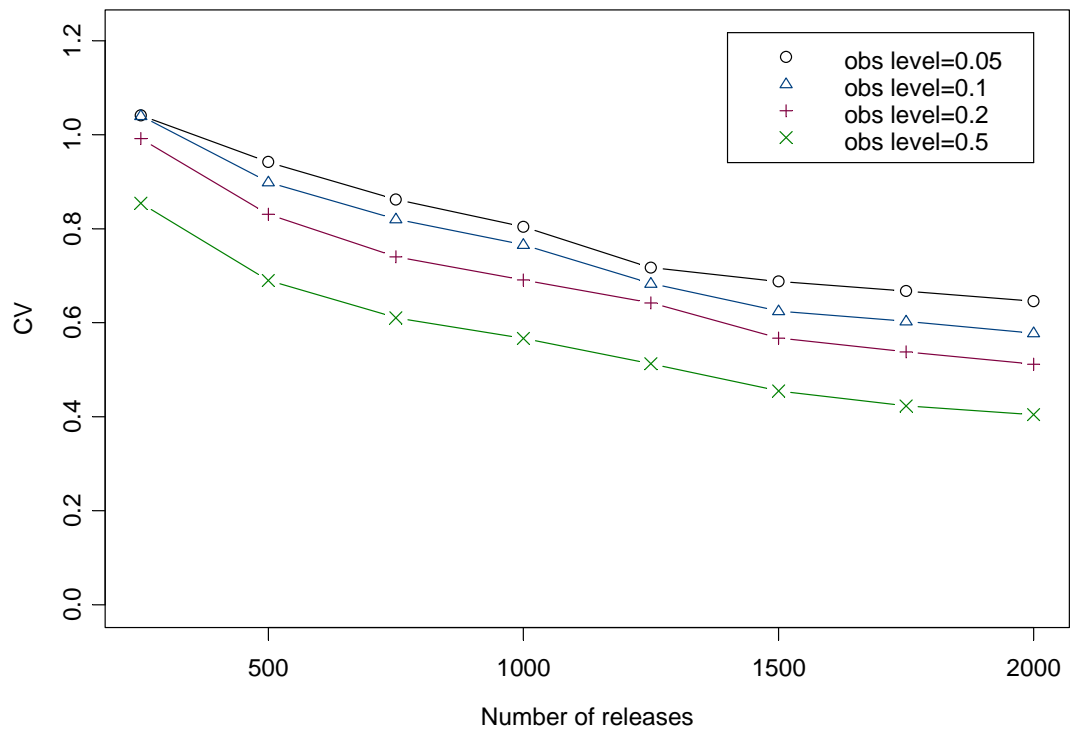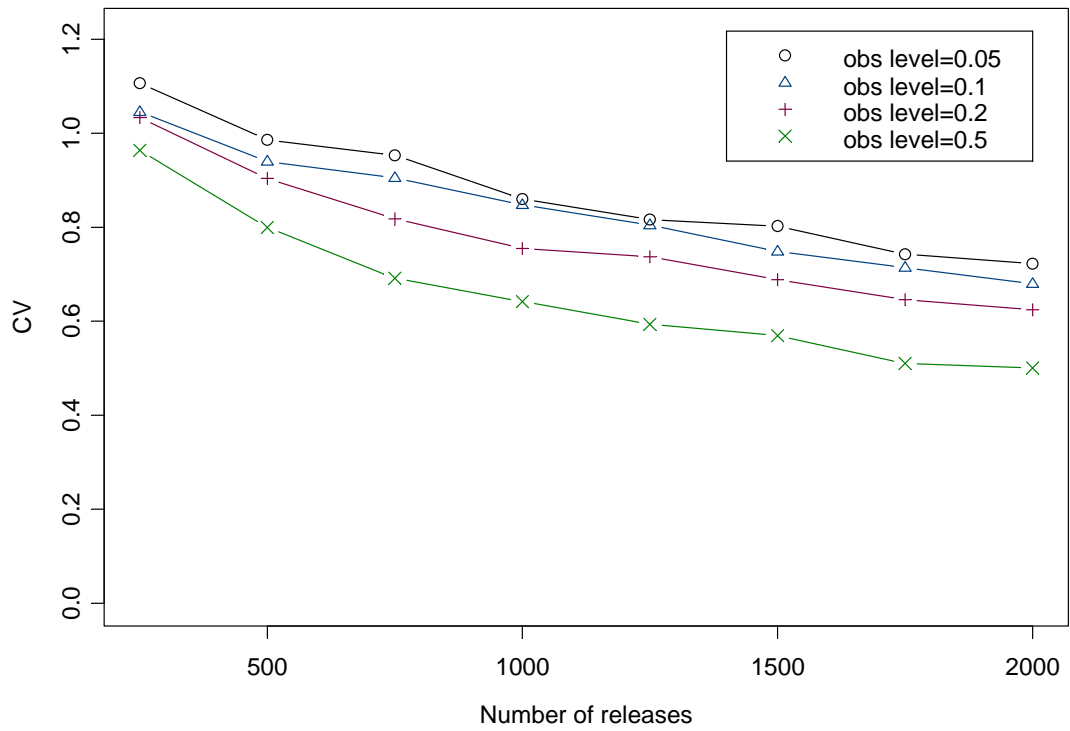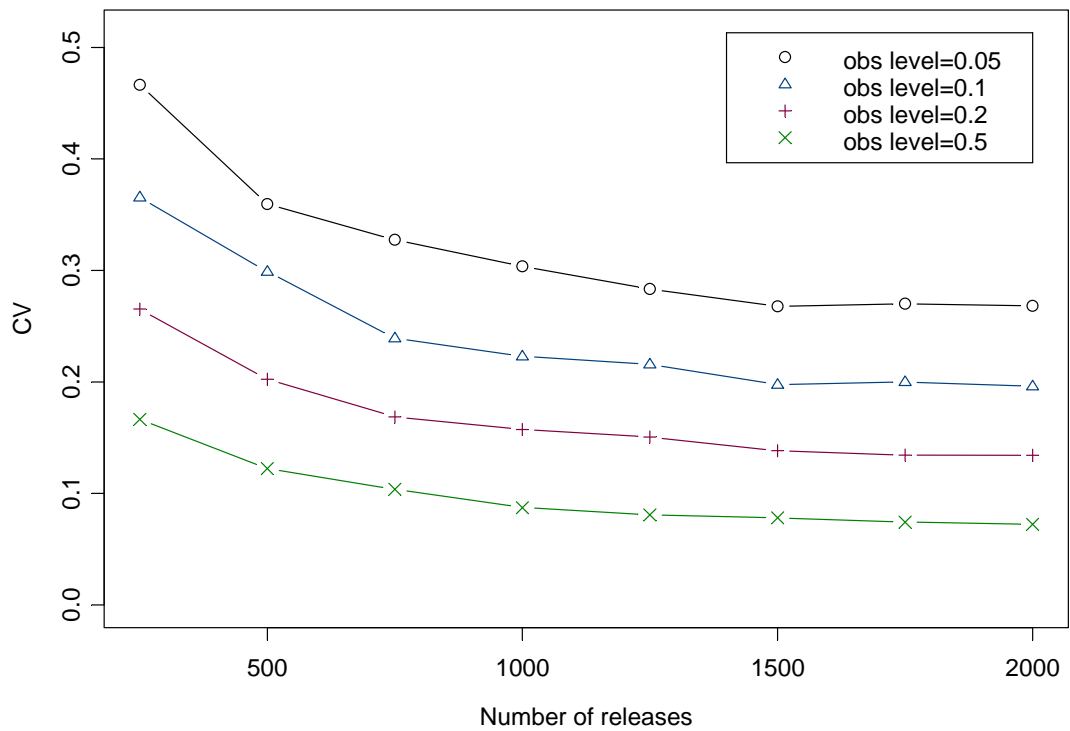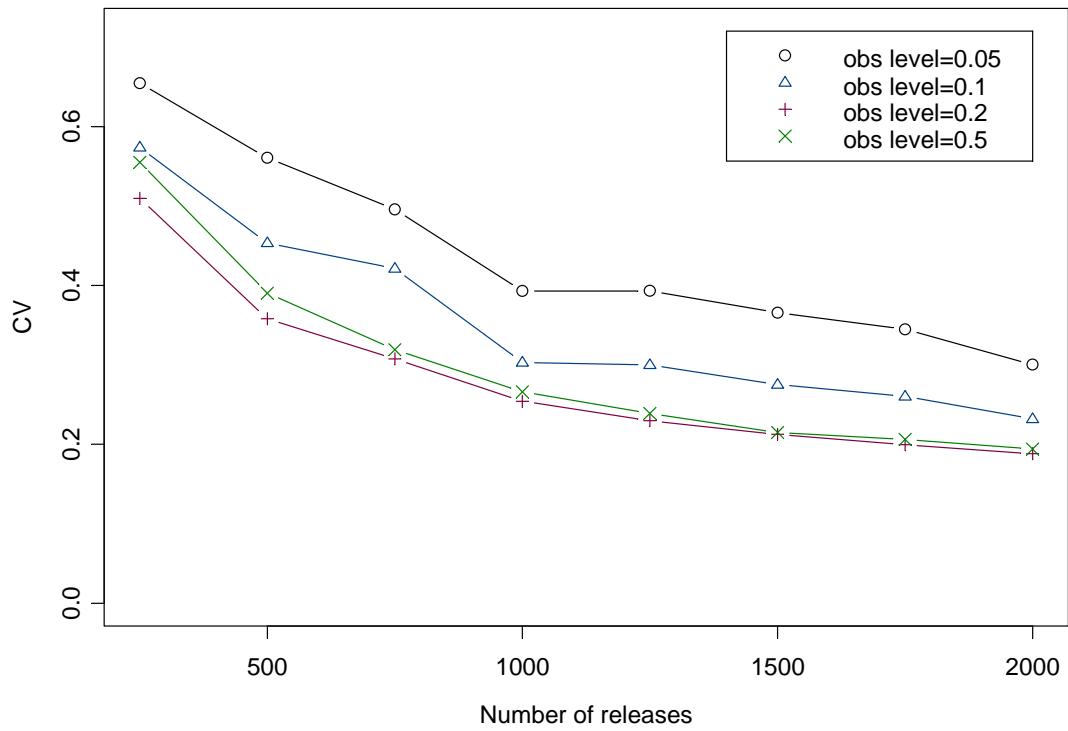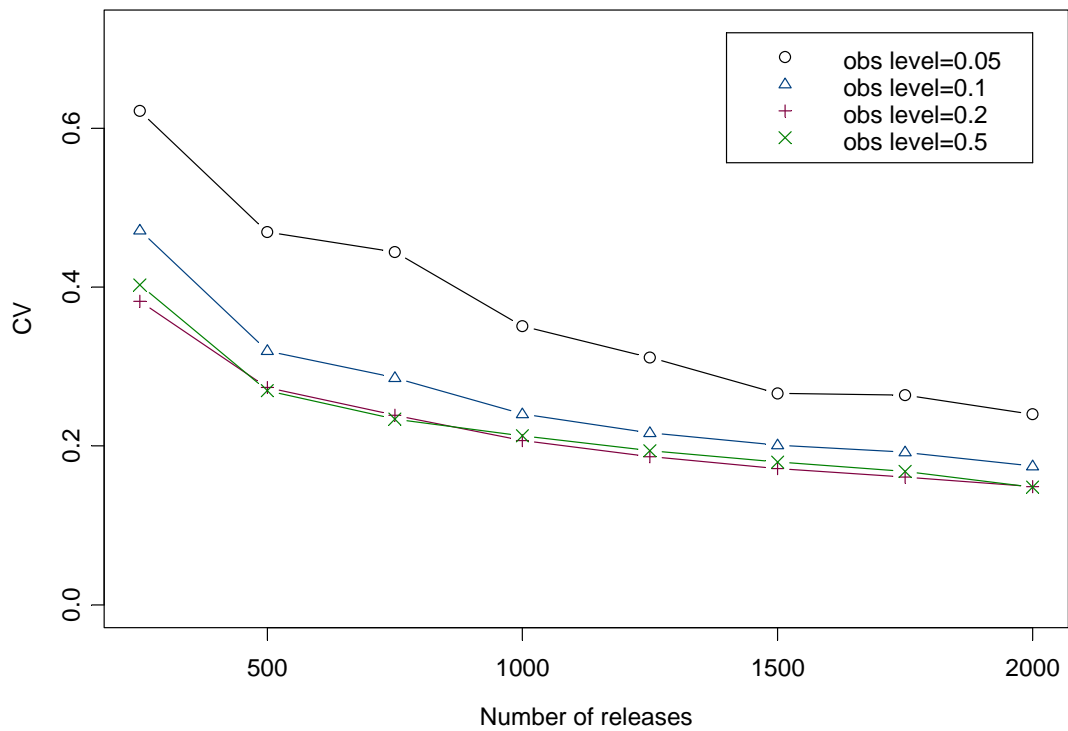
i) λ_3

Figure 3. Effect of changing the number of releases and level of observer coverage on the **coefficient of variation** (CV) of the fishing mortality rate estimates (a-c), mortality rate estimates (d-e), initial population size estimate (f), and reporting rate estimates (g-i). All results obtained using 500 simulations and a constant **true reporting rate of 0.75**.

a) F_1

b) F_2

c) F_3

d) M_1

e) M_2

f) P_1

g) λ_1

h) λ_2

i) λ_3

Figure 4. Effect of changing the number of releases and level of observer coverage on the **percent bias** of the fishing mortality rate estimates (a-c), mortality rate estimates (d-e), initial population size estimate (f), and reporting rate estimates (g-i). All results obtained using 500 simulations and a constant **true reporting rate of 0.25**.

a) F_1

b) F_2

c) F_3

d) M_1

e) M_2

f) P_1

g) λ_1

h) λ_2
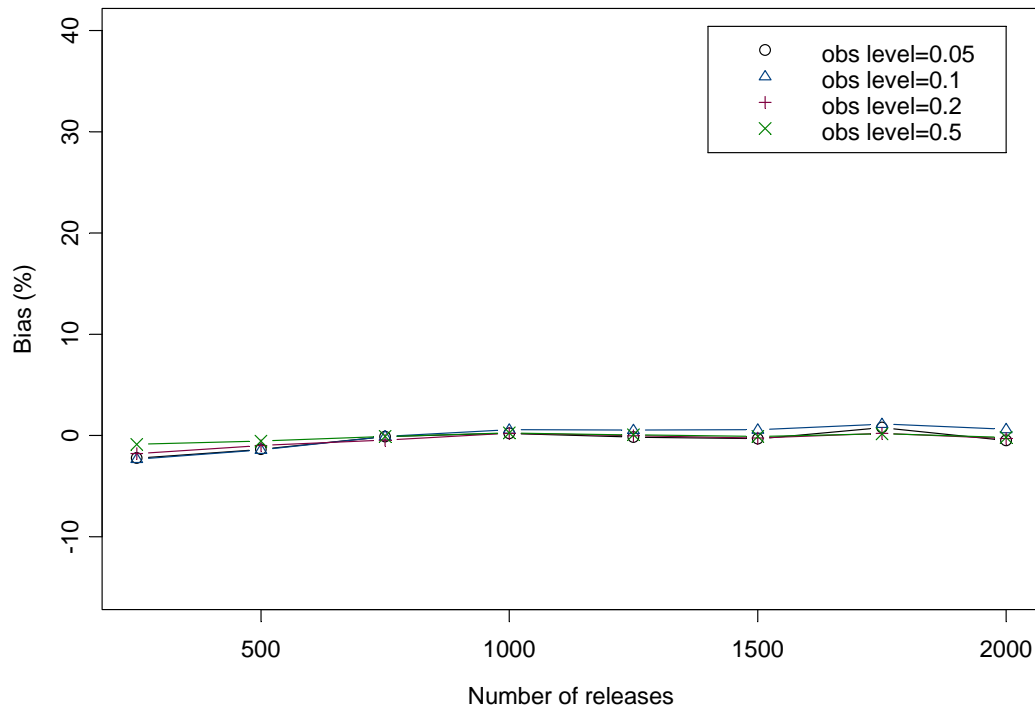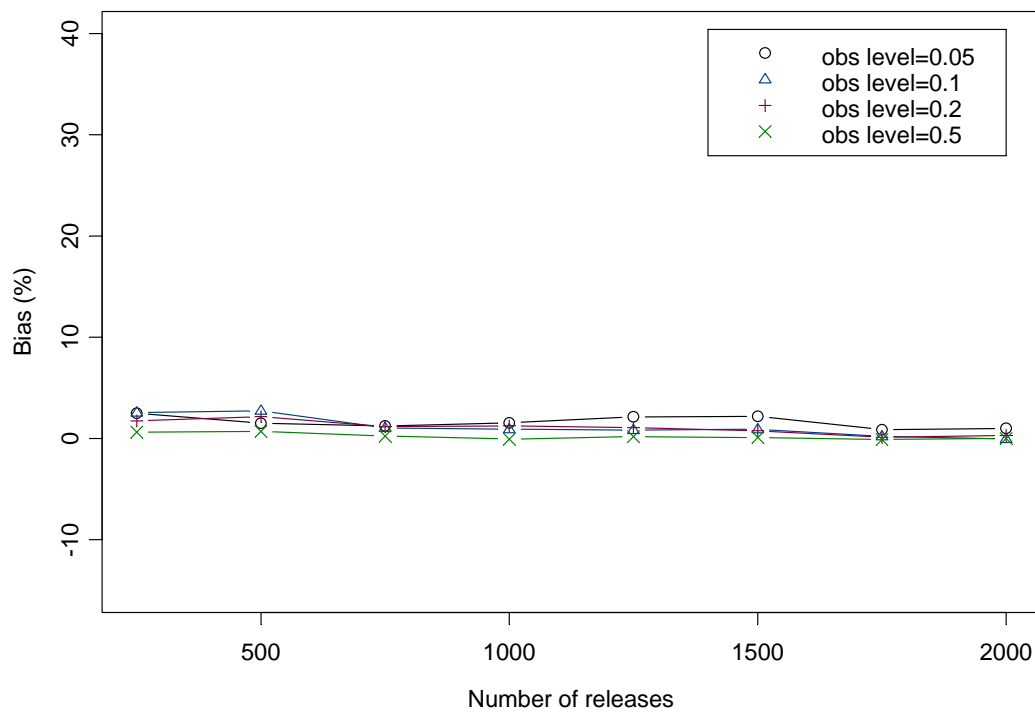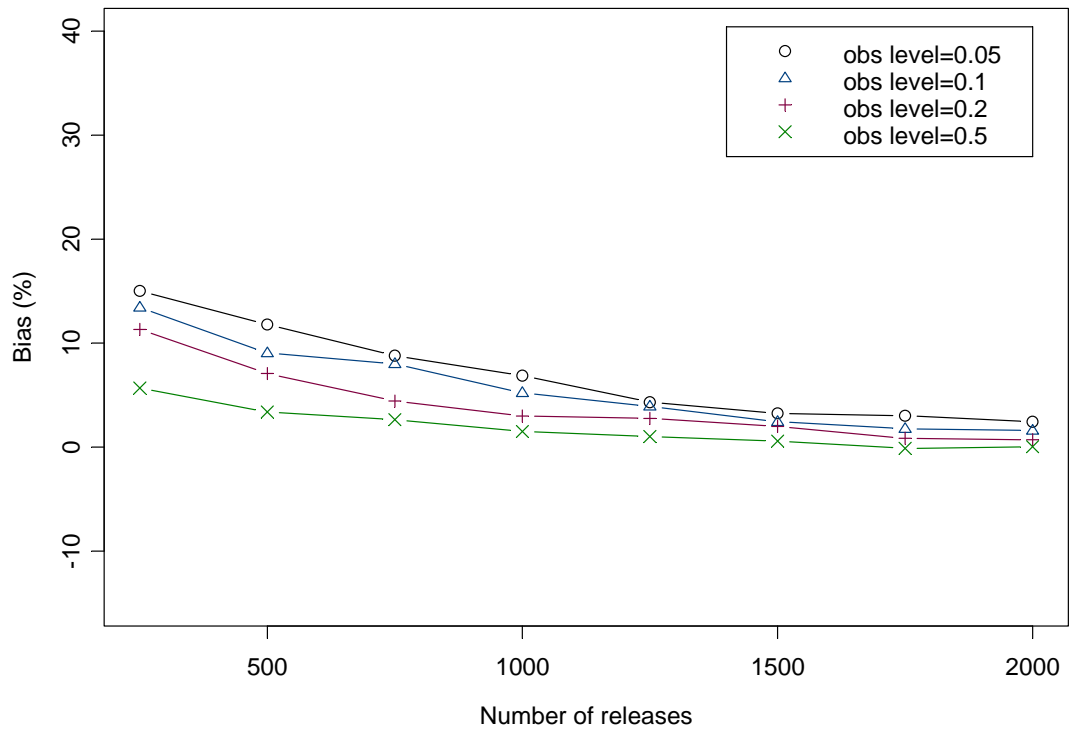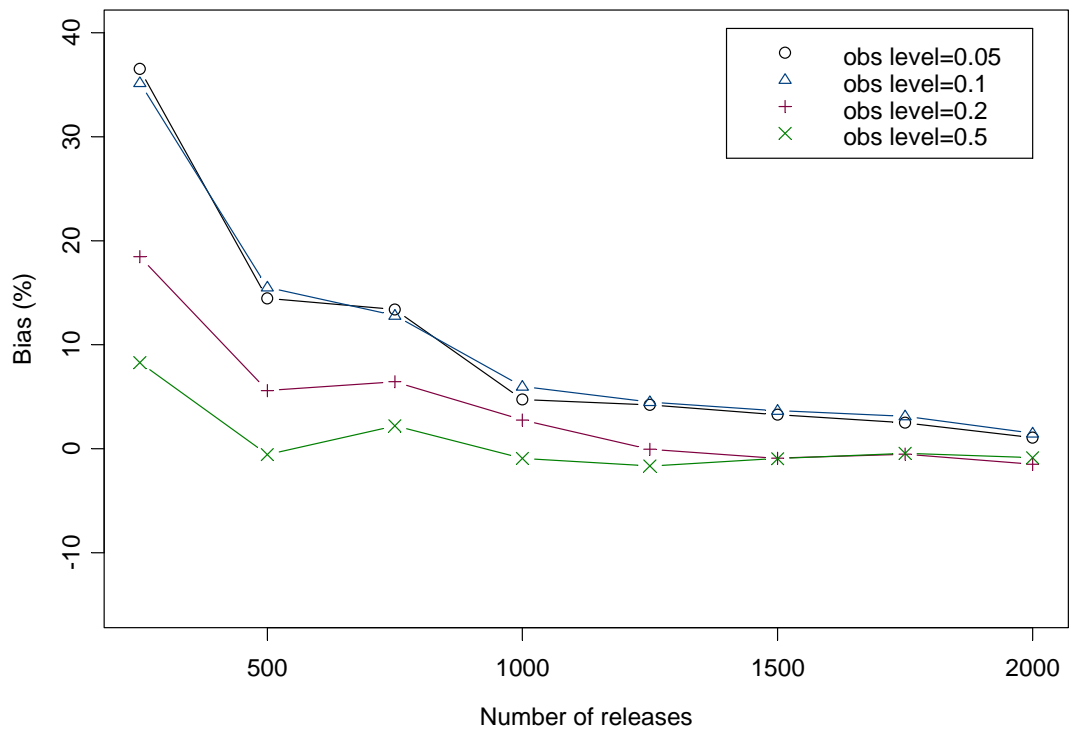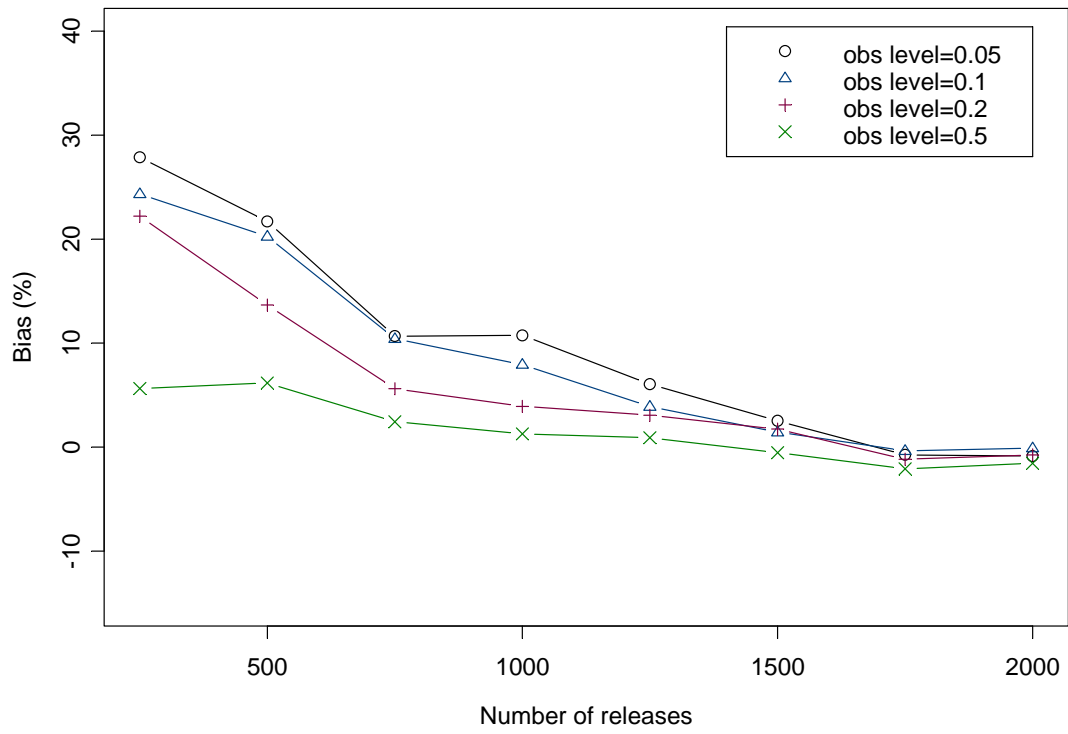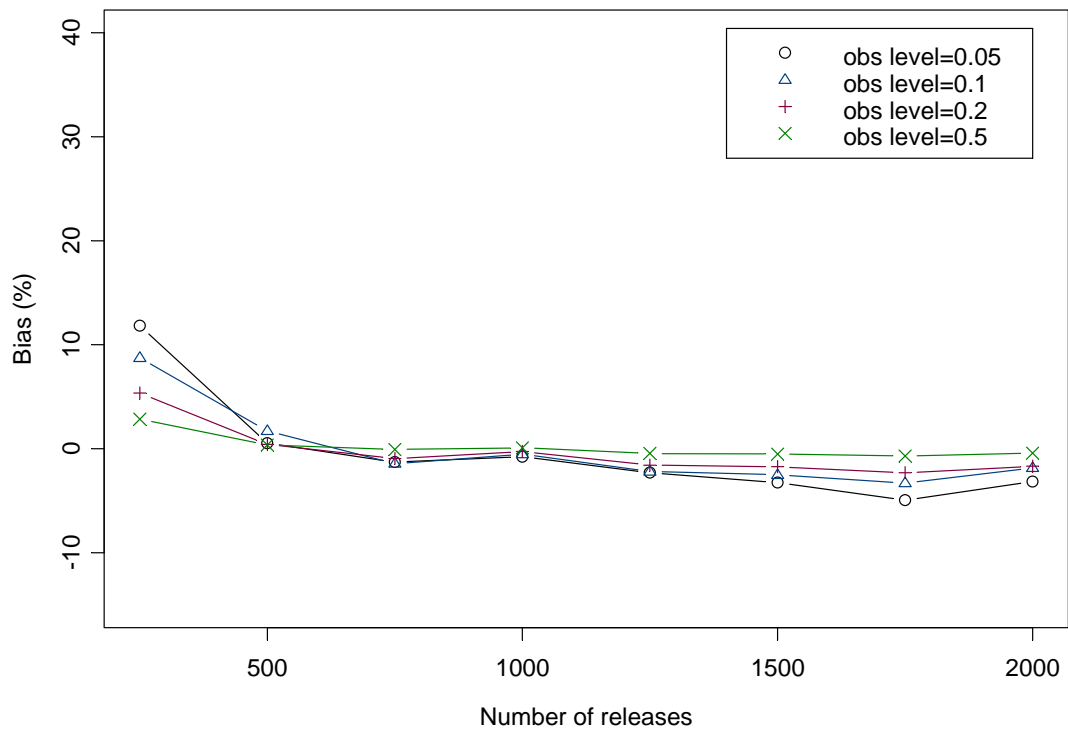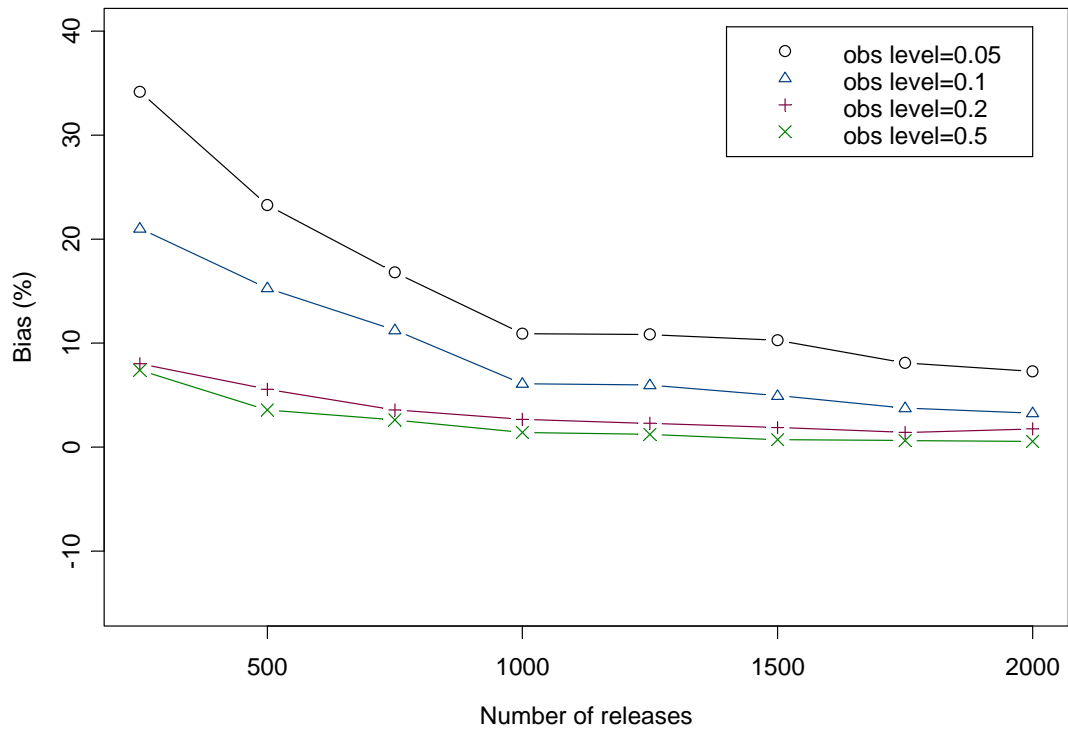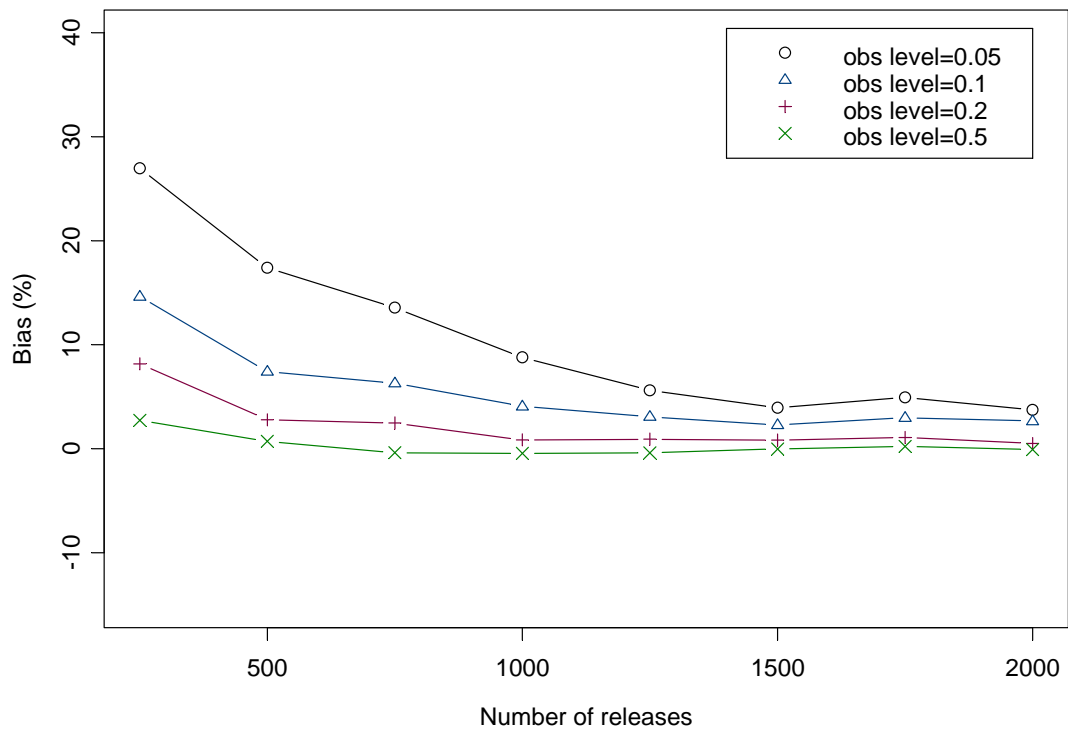
i) λ_3
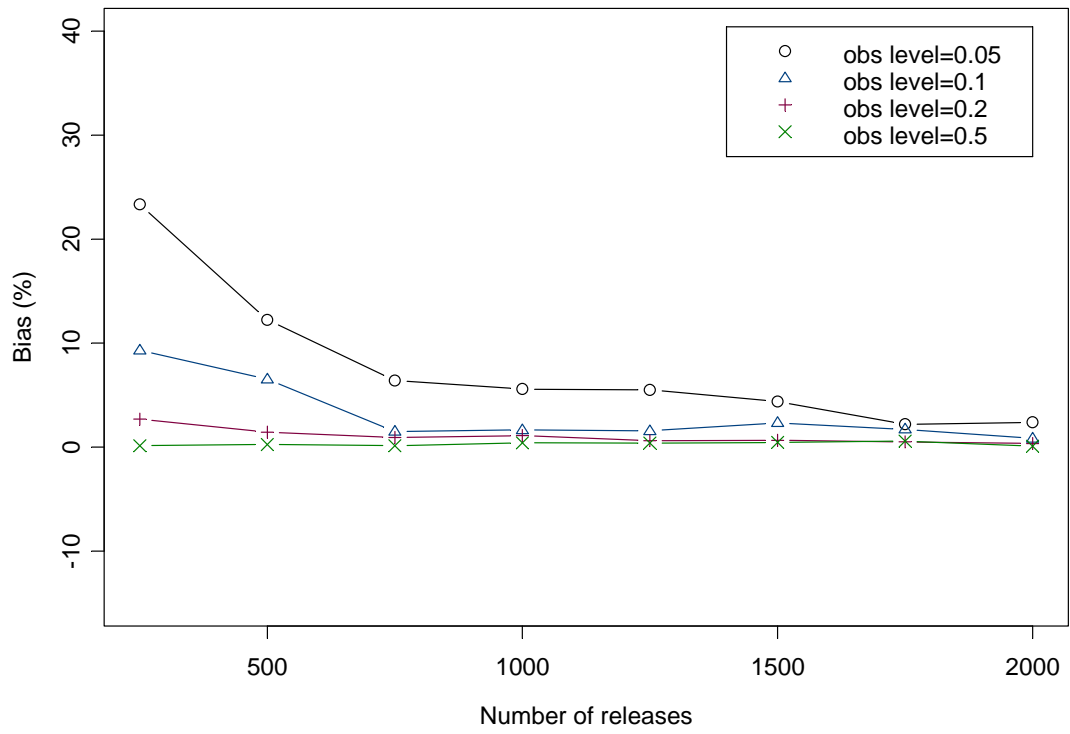
Figure 5. Effect of changing the number of releases and level of observer coverage on the **percent bias** of the fishing mortality rate estimates (a-c), mortality rate estimates (d-e), initial population size estimate (f), and reporting rate estimates (g-i). All results obtained using 500 simulations and a constant **true reporting rate of 0.75**.

a) F_1

b) F_2

c) F_3

d) M_1

e) M_2

f) P_1

g) λ_1

h) λ_2

i) λ_3

A8-1
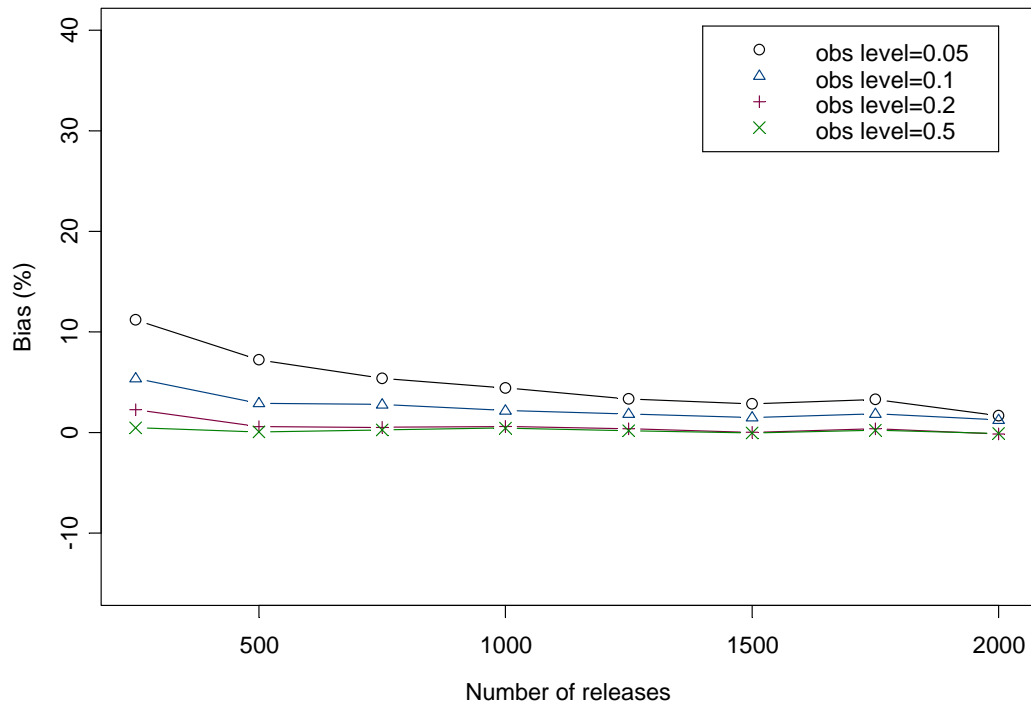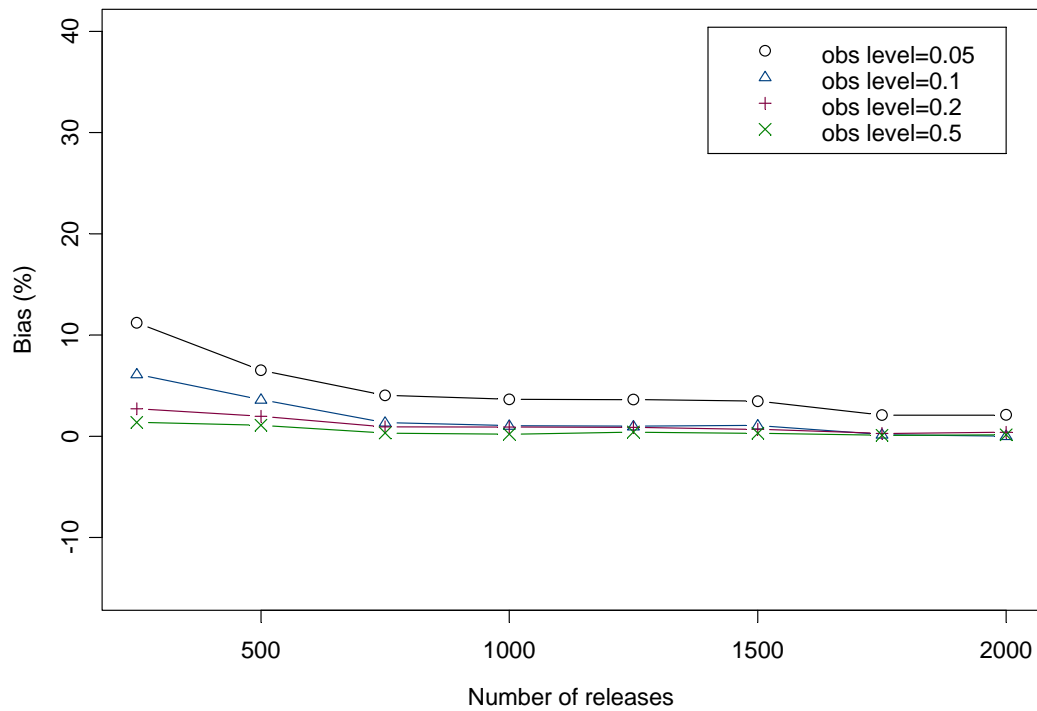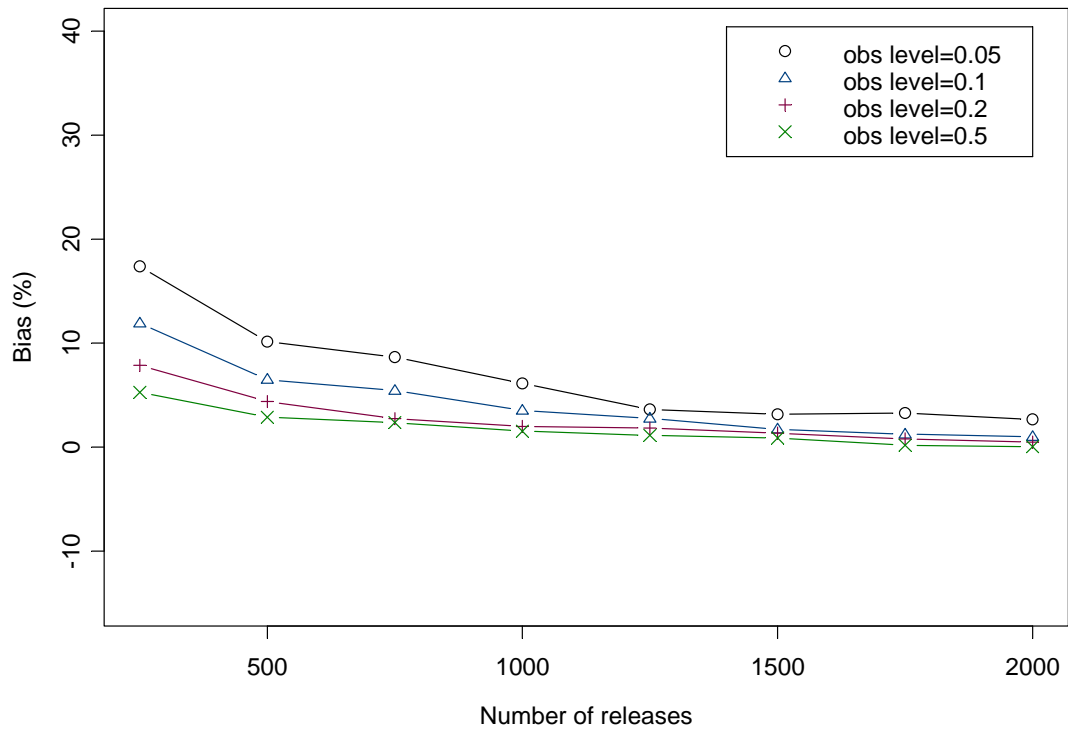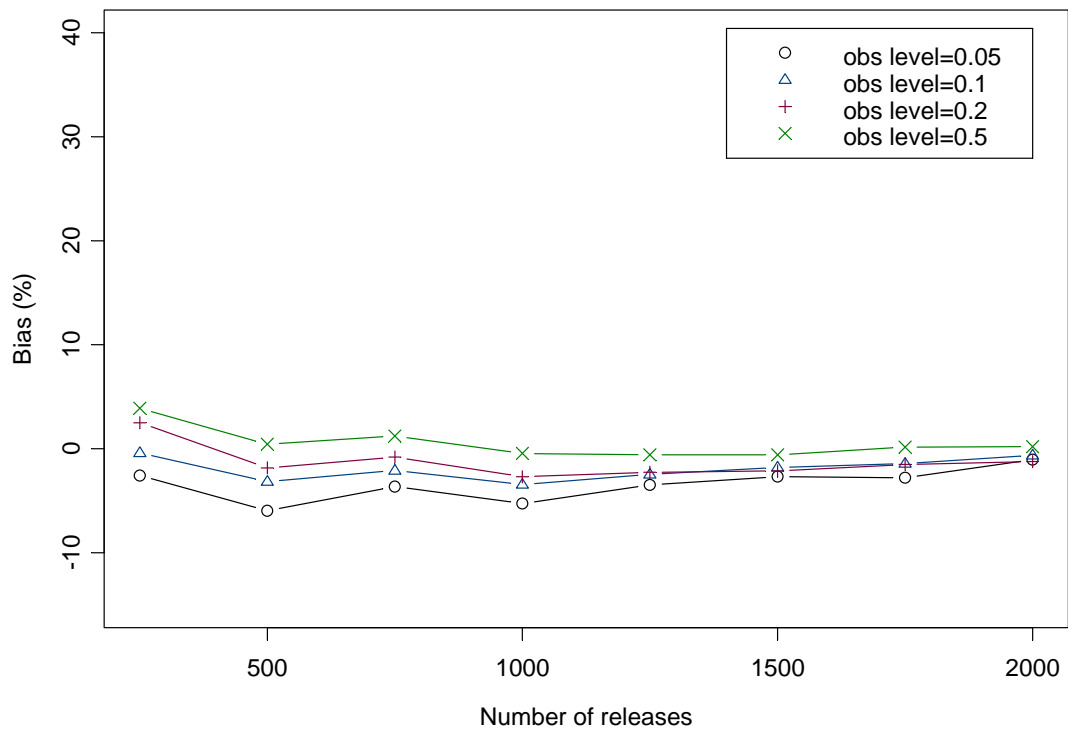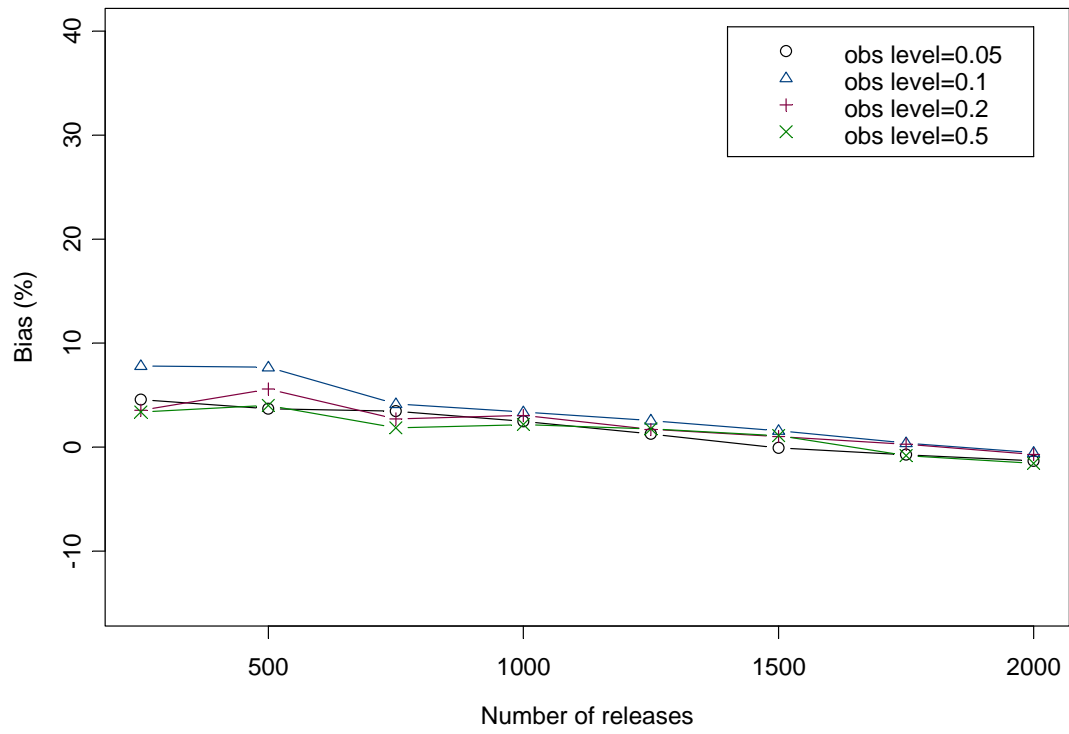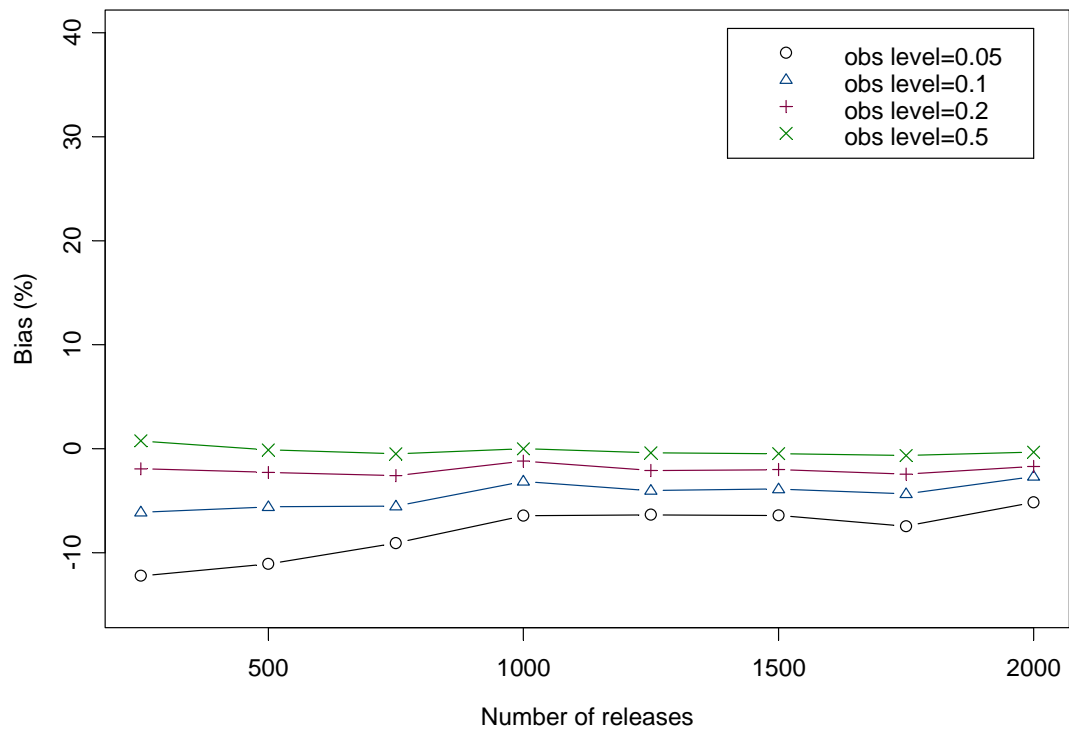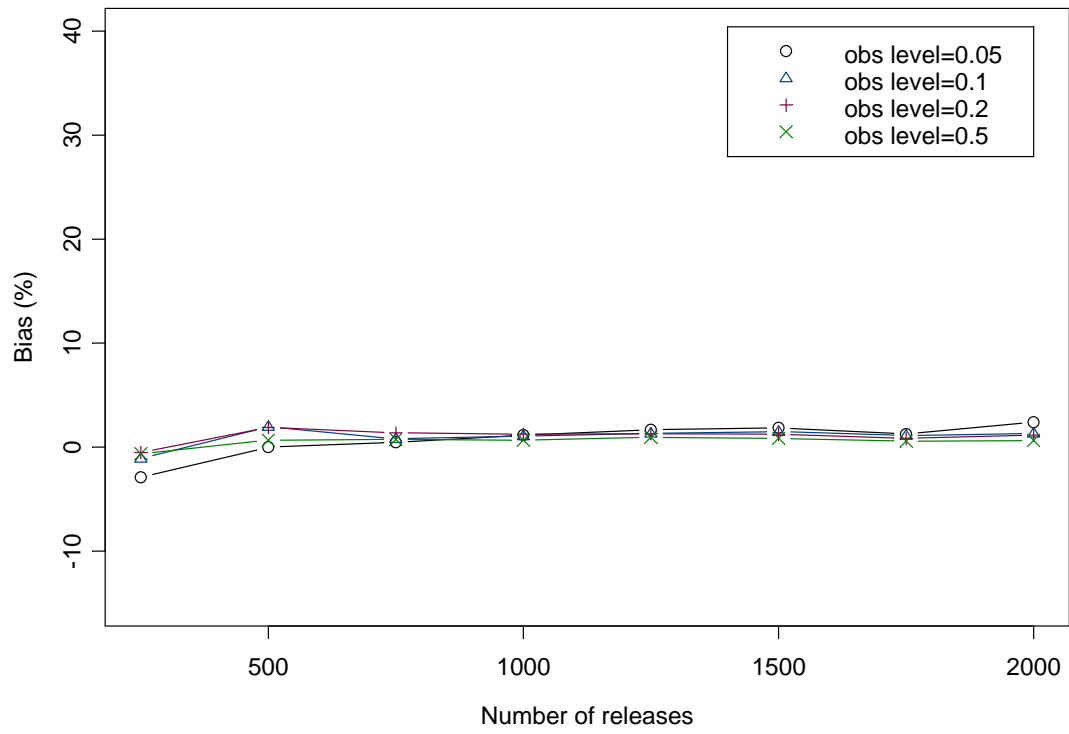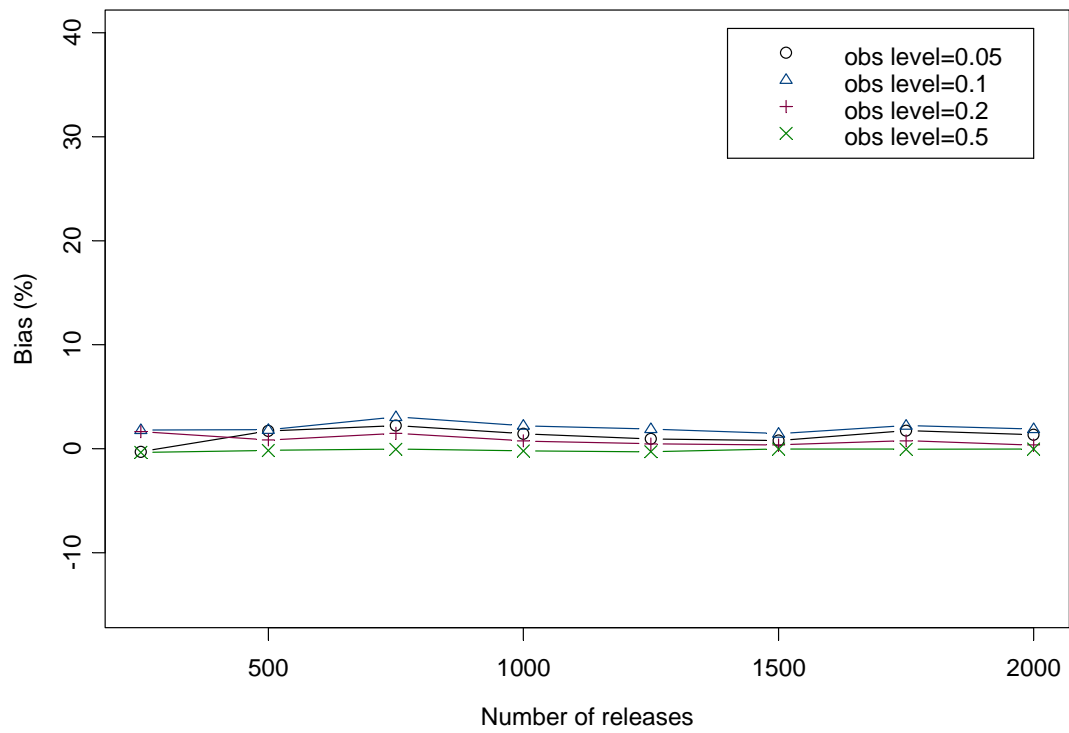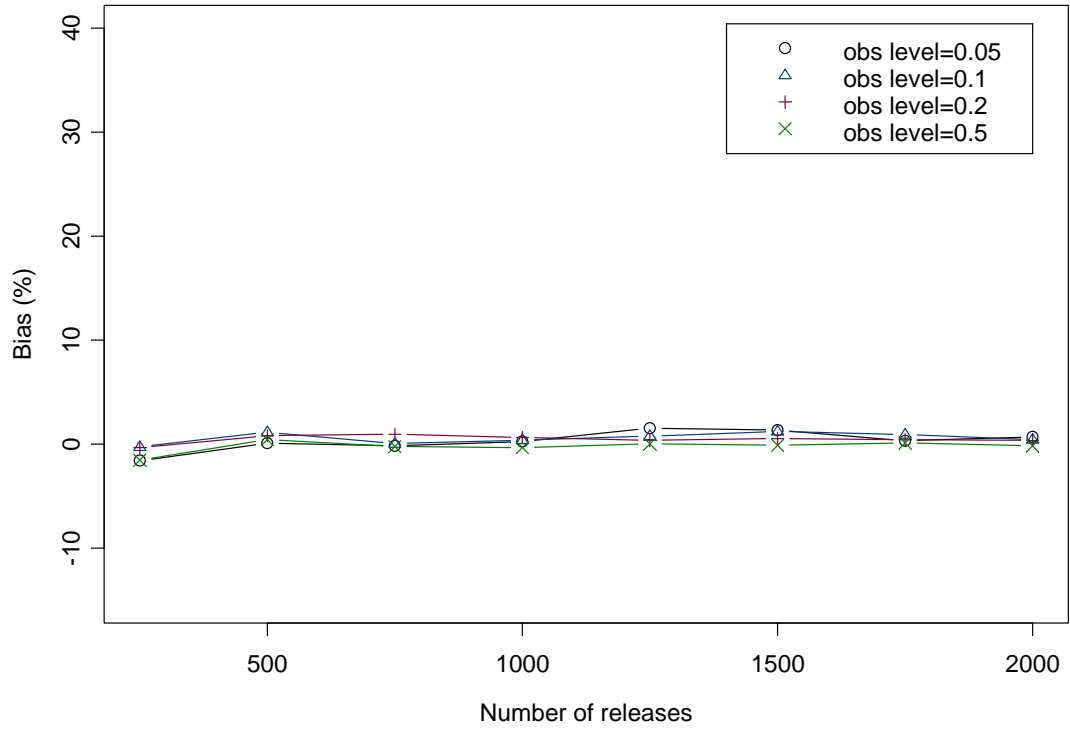
Appendix 8:

An evaluation of the consequences of aging errors in the context of a Brownie tag-recapture model

J. Paige Eveson, Tom Polacheck and Geoff M. Laslett

FRDC Project 2002/015

Introduction

Brownie et al. (1985) developed a range of models for estimating survival rates for data from multi-year tagging studies. Among these are models that can provide year- and age-specific estimates of survival and exploitation rates. One of the assumptions of Brownie models that allow for age-specific parameter estimates is that the age of individuals at the time of tagging is correctly determined, and therefore the number of individuals tagged at each age is known accurately. Here age refers to age class, and thus errors in the age at release will be integers. Brownie et al. (1985) did not concentrate on age-specific models because their models were developed for bird banding experiments, and age determination for most bird species is not possible. However, applications of Brownie models that allow for year- and age-specific estimates in a fishery context have been develop and applied (e.g., Pollock et al. 1991; Hoenig et al. 1998a, b; Polacheck et al. 1996, 1998).

A number of approaches exist for estimating the age of fish. The most accurate and reliable are those based on counts of annuli in hard parts (e.g., scales, otoliths, vertebrae). However, not all species produce reliable annuli for aging, and validated methods for age estimation using hard parts do not exist for many species. Moreover, age determination using hard parts whose collection requires injury or death to the fish, such as otoliths or vertebrae, can obviously not be directly used to estimate the age of fish to be tagged and released. The simplest and most easily implemented approach for estimating the age at release for tagged fish is to estimate age based on length and an estimated growth curve (which we refer to as "cohort slicing"). In fact, the tag-return data from the experiment can be used as the basis for estimating or updating the growth curve (although this would entail some delay before ages could be assigned to tagged fish). Cohort slicing has been the method used to age released fish in applications of the Brownie model to data from southern bluefin tuna (SBT) tagging experiments. However, aging errors will occur when cohort slicing from a length curve is used to assign ages to tagged fish.

In all of the analyses presented in this report, and in most published analyses of multi-year tagging data using Brownie-type models, the numbers of fish released at each age are treated as exact. As such, when cohort slicing is used to estimate the age of

tagged fish, it is important to know whether aging errors, if ignored, would induce substantial errors into the population dynamics parameter estimates derived from tagging models.

The goal of this appendix is to see how the accuracy of the mortality rate estimates obtained using a traditional Brownie model is affected when the ages at release contain errors that are not accounted for. To do so, we simulated "true" release and recapture data with correct release ages and corresponding "estimated" release and recapture data using length-based release age estimates. Note that the age of an animal at recapture is determined from the dates of release and recapture and the age at release; thus, an error in the age at release will be propagated through to the age at recapture (we assume the dates of release and recaptures are known quite accurately so that if the age at release is correct then the age at recapture is also correct). We fit a Brownie model to both data sets to obtain age-specific estimates of natural and fishing mortality, repeated this multiple times and compared the parameter estimates.

The results will depend on the number of fish whose ages are misclassified, which will depend on the shape and variability of the age-length relationship and the relative numbers of fish at each age being tagged. The results will also depend on the true parameter values – if mortality rates are very similar across ages, then age misclassifications will not have serious consequences. As such, we looked at results over a range of scenarios, including a scenario that is representative of the situation for SBT in order to get an idea of how much errors in aging are likely to have affected the parameter estimates for SBT presented in this report (Appendix 15 in particular, as well as the applications to SBT in Appendices 5 and 7).

Methods

The model

We present the model in terms of multi-year tag and recapture data for a single cohort of fish. This is how the Brownie model is generally formulated, and in our simulations to investigate the consequences of aging errors we will apply the model to data from a single cohort. Extension of the model to more cohorts is straightforward. Hoenig et al. (1998a) re-parameterized the traditional Brownie models in terms of

instantaneous rates of natural mortality and exploitation (i.e. fishing mortality in a fishery context) instead of rates of survival and tag recovery, and this is the parameterization used here, as well as throughout the report. For simplicity, we assume that all recaptured tags are reported. This assumption should not influence our general findings about how aging errors affect the mortality rate parameter estimates.

For a single cohort of fish tagged at ages 1, 2 and 3, the expected numbers of returns at age corresponding to each release age are given by:

Release year	Release age	Number releases	Expected number of recaptures		
			Age 1	Age 2	Age 3
1	1	N_1	$N_1 u_1$	$N_1 S_1 u_2$	$N_1 S_1 S_2 u_3$
2	2	N_2		$N_2 u_2$	$N_2 S_2 u_3$
3	3	N_3			$N_3 u_3$

where

N_a = the number of tag releases of age a fish from a specific cohort

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

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

F_i = the instantaneous fishing mortality rate for fish of age i

M_i = the instantaneous natural mortality rate for fish of age i.

Note that S_i represents the annual survival rate of age i fish and u_i represents the annual exploitation rate of age i fish. These equations are the same as those presented in Appendix 5 except they do not include a multiplicative term for reporting rates, since we are assuming reporting rates to be one.

If each tag recapture is assumed to be independent, then the numbers of recaptures at age (including those not recaptured) from releases at a particular age are expected to

be multinomial. The likelihood function for the observed numbers of recaptures from all release ages is then the product of multinomials given by:

$$L_R = \mathrm{K} \times \prod_a \left(\left(\prod_{i \ge a} p_{a,i}^{R_{a,i}} \right) \left(1 - p_{a,\bullet} \right)^{N_a - R_{a,\bullet}} \right) \tag{1}$$

$$\mathrm{K} = \prod_a \frac{N_a!}{\prod_{i \ge a} R_{a,i}! (N_a - R_{a,\bullet})!}$$

where a indexes release age, i indexes recapture age, $R_{a,i}$ is the number of tags recaptured from age i fish released at age a, and $p_{a,i}$ is the probability of a tag being recaptured from an age i fish released at age a. An expression for $p_{a,i}$ can be obtained from the expected number of returns in the above table divided by N_i. Explicitly,

$$p_{a,i} = \begin{cases} u_i & i = a \\ S_a \cdots S_{i-1} u_i & i > a \end{cases} \tag{2}$$

Estimates of the fishing and natural mortality rate parameters (F's and M's) can be obtained by maximizing the likelihood given in equation (1). An estimate of fishing mortality is obtained for each age of recapture, but the information for estimating natural mortality comes from the differential between the expected returns at age $i+1$ of fish released at age i and those released at age $i+1$. Thus, in an experiment with n consecutive release years, estimates can only be obtained for $n-1$ of the natural mortality rate parameters (regardless of the number of recapture years). We address this issue by assuming that $M_i = M_{n-1}$ for $i \ge n$, but other constraints could be used, such as imposing a parametric relationship between natural mortality and age.

Simulation steps

We simulated data from a tagging experiment with 3 consecutive years of releases and recaptures. We assumed that fish of *true* ages 1, 2 and 3 are tagged in each year, and that the number of fish tagged at each age in a given year is proportional to the age

distribution of the population in that year.[1] Therefore, to determine the numbers of
fish tagged at each age in each year we need to know the numbers of fish of ages 1, 2
and 3 in the population in each of years 1, 2 and 3. This involves a total of 5 cohorts,
namely fish born in years $-2, -1$, 0, 1 and 2, which we will refer to as cohorts -2,
-1, 0, 1 and 2. Note that only cohort 0 is tagged at all 3 ages; this is the cohort
assumed to be targeted for the tagging experiment and the one that we will
concentrate on for modelling purposes. However, to keep track of the number of
tagged fish incorrectly classified as belonging to, or not belonging to, this cohort
based on their age estimates, we needed to generate release data for all fish tagged in
each year.

To generate numbers of age 1, 2 and 3 fish in the population for 3 consecutive years
we used the basic population dynamics equation

$$P_{a+1,t+1} = P_{a,t} \exp\left(-F_{a,t} - M_{a,t}\right)$$

where

$P_{a,t}$ = the number of individuals of age i in year t;

$F_{a,t}$ = the instantaneous fishing mortality rate for individuals of age a in year t;

$M_{a,t}$ = the instantaneous natural mortality rate for individuals of age a in year t.

Given the size of each of the 5 cohorts at age 1 (i.e., $P_{1,t}$ for years $t = -1$, 0, 1, 2, 3)
and the age- and year-specific fishing and natural mortality rates for ages 1 to 3 and
years -1 to 2 (i.e., $F_{a,t}$ and $M_{a,t}$ for $a = 1, 2, 3$ and $t = -1$, 0, 1, 2) we can calculate the
population size for the subsequent ages and years of interest; that is, we can calculate
$P_{a,t}$ for ages 1, 2 and 3 in years 1, 2 and 3.

---

[1] Assuming that only fish of true ages 1, 2 and 3 are tagged is a somewhat artificial example because it
assumes that only these age-classes are available for tagging (which would contradict the complete
mixing assumption of most tag-recapture models). A more realistic approach may have been to assume
that fish of all ages are tagged in proportion to abundance. We took the former approach because we
found it more straightforward to simulate the true and estimated data sets. If the latter approach had
been taken, we would expect the biases due to aging errors to be slightly smaller than those presented;
however, we would not expect the general results and conclusions to be greatly affected.

In each year of tagging (t = 1, 2, 3), we assumed that a total of N_T fish belonging to ages 1, 2 and 3 were tagged. For a given year t, the numbers of fish tagged at each age (denoted by $N_{1,t}$, $N_{2,t}$, and $N_{3,t}$) were determined using a random multinomial distribution with probabilities equal to the proportions of each age in the population. That is,

$$(N_{1,t}, N_{2,t}, N_{3,t}) \sim \text{Multinomial}(\,N_T,\ \boldsymbol{p}=(p_1, p_2, p_3))$$

where $p_a = P_{a,t}\,/(\,P_{1,t} + P_{2,t} + P_{3,t})$. Thus, we generated $N_{a,t}$ for ages 1, 2 and 3 in years 1, 2 and 3.

Next we generated age and length data for all fish released in each year because we wanted to keep track of how many were incorrectly classified as belonging to, or not belonging to, the cohort of interest based on their length-based age estimates. To generate a length, l, for a fish of age a, we assumed that fish growth can be described by a von Bertalanffy (VB) equation with asymptotic length parameter L_∞, growth rate parameter k, and age at length 0 of a_0. We allowed for individual variability in growth by allowing the asymptotic length parameter to vary between fish; in particular, we modelled L_∞ as a random normal variate with mean μ_∞ and variance σ_∞^2. We also allowed for additional model error and/or measurement error by including an additive random normal error component with mean 0 and variance σ_ε^2. Specifically, we assumed

$$l(a) = L_\infty \left(1 - \exp\left(-k\left(a - a_0\right)\right)\right) + \varepsilon$$

where $L_\infty \sim \text{N}(\mu_\infty, \sigma_\infty^2)$ and $\varepsilon \sim \text{N}(0, \sigma_\varepsilon^2)$.

Thus, we generated a length for a fish of age a using a normal distribution with mean

$$E(l \mid a) = \mu_\infty \left(1 - \exp\left(-k\left(a - a_0\right)\right)\right)$$

and variance

$$V(l \mid a) = \left(1 - \exp\left(-k\left(a - a_0\right)\right)\right)^2 \sigma_\infty^2 + \sigma_\varepsilon^2.$$

In a tagging experiment, we would not know the true age of a fish at the time of tagging. Instead, we would estimate the fish's age based on its length and a given age-length relationship. If the mean age-length relationship for the animals being tagged is known reasonably accurately (i.e., in our case, if we know fish grow according to a VB curve and we know the true parameter values for μ_∞, k and a_0), then we can use this relationship to calculate length cut-offs for fish of a particular age. For example, we can calculate the mean length at ages $a - 0.5$ and $a + 0.5$, and any fish whose measured length falls within this range will be estimated to be age a.

Using this method, we estimated an age for each fish in our simulations based on its length. We then tallied the number of fish tagged in year t of true age a that were estimated to be age $\hat{a}$; we denote these by $N_{a,\hat{a},t}$. Note that the true number of age a releases in year t is $N_{a,t} = \sum_{\hat{a}} N_{a,\hat{a},t}$, and that the estimated number of age a releases in year t is $\hat{N}_{a,t} = \sum_a N_{a,\hat{a},t}$

The probability of a fish that was released at age a in year t being recaptured in year y (at age $i = a + y - t$) is given by:

$$p_{a,t,y} = \begin{cases} u_{i,t} & y = t \\ S_{a,t}S_{a+1,t+1}\cdots S_{i-1,\,y-1}u_{i,\,y} & y > t \end{cases}$$

where

$$S_{r,s} = \exp\left(-F_{r,s} - M_{r,s}\right)$$

$$u_{r,s} = \frac{F_{r,s}}{F_{r,s} + M_{r,s}}\left(1 - S_{r,s}\right).$$

Consider the releases in year t at true age a, $N_{a,t}$. The numbers of recaptures at age (including those not recaptured) corresponding to these releases are assumed to be multinomial with probabilities given by the expression above. Furthermore, if we

assume that the probability of recapture (i.e. fishing mortality and natural mortality) depends on the fish's true age and not on its estimated age, then the recaptures corresponding to releases in year t at true age a and estimated age $\hat{a}$, $N_{a,\hat{a},t}$, will also be multinomial with the same probabilities. Thus, for our simulations, we generated multinomial recaptures corresponding to each set of releases $N_{a,\hat{a},t}$. Let $R_{a,\hat{a},t,y}$ denote the number of recaptures in year y of fish released at true age a but estimated release age $\hat{a}$ in year t. For our simulations, we are only interested in recaptures from the same years we have releases. Thus, for each set of releases in year 1 we need to generate recapture data for years 1, 2 and 3 (i.e., $R_{a,\hat{a},1,1}, R_{a,\hat{a},1,2}, R_{a,\hat{a},1,3}$); for each set of releases in year 2 we need to generate recapture numbers for years 2 and 3 (i.e., $R_{a,\hat{a},2,2}, R_{a,\hat{a},2,3}$); and for each set of releases in year 3 we only need to generate recapture numbers for year 3 (i.e., $R_{a,\hat{a},3,3}$). For example, for year 1 we generated recaptures corresponding to releases $N_{a,\hat{a},1}$ according to:

$$(R_{a,\hat{a},1,1}, R_{a,\hat{a},1,2}, R_{a,\hat{a},1,3}, N_{a,\hat{a},1} - \sum_{j=0}^{2} R_{a,\hat{a},1,1+j}) \sim \text{Multinomial}(N_{a,\hat{a},1}, \boldsymbol{p})$$

where $\boldsymbol{p} = (p_{a,1,1}, p_{a,1,2}, p_{a,1,3}, 1 - \sum_{j=0}^{2} p_{a,1,1+j})$.

Note that to generate all of the necessary recapture numbers for our simulations we needed to know natural and fishing mortality rates for ages 1 to 3 in year 1, for ages 1 to 4 in year 2, and for ages 1 to 5 in year 3 (i.e., $F_{a,1}$ and $M_{a,1}$ for $a = 1,...,3$; $F_{a,2}$ and $M_{a,2}$ for $a = 1,...,4$; $F_{a,3}$ and $M_{a,3}$ for $a = 1,...,5$).

Using these recapture numbers, we calculated the number of recaptures in year y of fish released at true age a in year t as $R_{a,t,y} = \sum_{\hat{a}} R_{a,\hat{a},t,y}$ (i.e., the "true" recapture data), and we calculated the number of recaptures in year y from fish released at estimated age a in year t is $\hat{R}_{a,t,y} = \sum_{a} R_{a,\hat{a},t,y}$ (i.e., the "estimated" recapture data).

For input into the Brownie model, we are only interested in the "true" and "estimated" data for cohort 0. In particular, we are interested in the releases at true ages 1, 2 and

3 in years 1, 2 and 3, respectively, and the recaptures in years 1, 2 and 3 corresponding to these releases.   These data can be summarized as follows:

Release year	True release age	True number of releases	True number of recaptures		
			Age 1	Age 2	Age 3
1	1	$N_{1,1}$	$R_{1,1,1}$	$R_{1,1,2}$	$R_{1,1,3}$
2	2	$N_{2,2}$		$R_{2,2,2}$	$R_{2,2,3}$
3	3	$N_{3,3}$			$R_{3,3,3}$

Similarly, we are interested in the releases at estimated ages 1, 2 and 3 in years 1, 2 and 3, respectively, and the recaptures in years 1, 2 and 3 corresponding to these releases.   These data can be summarized as follows:

Release Year	Estimated release age	Estimated number of releases	Estimated number of recaptures		
			Age 1	Age 2	Age 3
1	1	$\hat{N}_{1,1}$	$\hat{R}_{1,1,1}$	$\hat{R}_{1,1,2}$	$\hat{R}_{1,1,3}$
2	2	$\hat{N}_{2,2}$		$\hat{R}_{2,2,2}$	$\hat{R}_{2,2,3}$
3	3	$\hat{N}_{3,3}$			$\hat{R}_{3,3,3}$

Note that age and year are redundant for a single cohort, so we can drop the year subscript on $N_{a,t}$ and $R_{a,t,y}$ and refer only to N_a and $R_{a,y}$; similarly, we can drop the year subscript on $\hat{N}_{a,t}$ and $\hat{R}_{a,t,y}$. The data then appears analogous to the way it was presented in the previous section describing the Brownie model.

We applied the Brownie model to each of the above data sets to obtain estimates of fishing and natural mortality for both the true and the estimated data sets. In particular, by maximizing the likelihood given in equation (1), we obtained estimates of $F_{1,1}$, $F_{2,2}$, $F_{3,3}$, $M_{1,1}$ and $M_{2,2}$ for each data set. Again, year and age are redundant

for a single cohort, so we drop the year subscript and simply refer to these as F_1, F_2, F_3, M_1 and M_2.

A number of different scenarios (i.e., sets of parameter values) were used to generate the true and estimated release and recapture data (see next section). For each scenario, we repeated the above steps 100 times and compared the parameter estimates obtained using the true age data with those obtained using the estimated age data.

Scenarios

The effect of aging errors on the mortality rate estimates will depend on the number of fish whose ages are misclassified. This will depend largely on (1) the shape of the age-length relationship, because the steeper the growth curve (i.e., the greater the rate of growth) the less separation between mean lengths of fish of adjacent ages, and (2) the variability in growth (i.e., the greater the variability in growth the greater the overlap in length distributions of fish of adjacent ages). Increasing the variability in growth should have a similar effect on the results as decreasing the steepness of the curve, so we chose to keep the mean age-length relationship the same throughout all our scenarios and just change the amount of variability in the relationship. In particular, throughout all our simulations we used $\mu_\infty = 180$, $k = 0.2$ and $a_0 = -0.5$ for the age-length relationship; these values roughly describe growth for SBT in the 1980s and 1990s (Laslett et al. 2002; Polacheck et al. 2004). There are two components contributing to variability in growth: 1) variability in asymptotic length (L_∞) between fish; 2) additional model and measurement error. The first component, individual variability in L_∞, results in variability in length increasing with age; however, this effect is minimal at the young ages we are considering (ages 1 to 3), so we chose to keep the asymptotic length variance parameter the same throughout our simulations and only vary the parameter for the additional model/measurement error. In particular, we used $\sigma_\infty = 10$ and varied σ_ε.

The number of fish misclassified as belonging to each age class in a year will depend not only on the age-length relationship, but also on the relative number of fish of ages

1, 2 and 3 being tagged in a year. That is, it depends on the relative number of fish of ages 1, 2 and 3 in the population in a year, since we assumed that tagging occurs in proportion to abundance. If a particularly large cohort comes through the population, then there will be a greater number of fish from this cohort in the population in a given year than from other cohorts, and therefore more fish from this cohort will be incorrectly classified as belonging to other cohorts as vice versa. On the contrary, if a particularly small cohort comes through, then there will be fewer fish from this cohort in the population in a given year than from other cohorts, and therefore fewer fish from this cohort will be incorrectly classified as belonging to other cohorts as vice versa. We considered scenarios in which all cohorts being tagged: had the same initial (age 1) size; had the same initial size except for one larger cohort (the cohort being modelled; had the same initial size except for one smaller cohort (again, the cohort being modelled). We also considered a scenario intended to be representative of SBT for which the size of cohorts decreased over time. We kept the absolute number of fish tagged in each year the same throughout our simulations, namely $N_T = 3000$. Although the absolute number of fish tagged will affect the accuracy of the parameter estimates, the effect should be similar for the estimates made using the true release-recapture data and those made using the estimated data.

Finally, the effect of aging errors on the mortality rate estimates will depend on the true mortality rate values – if mortality rates are the same across ages, then age misclassifications should not cause biases in the mortality rate estimates (the uncertainty in the estimates could be affected). We would expect biases in the estimated mortality rates to increase as the difference between the true mortality rates at adjacent ages increases. In order to reduce the number of possible scenarios, we assumed that natural mortality was constant across ages and years, and that fishing mortality varied with age but not year. We considered scenarios in which fishing mortality increased from age 1 to 3, and those in which it decreased from age 1 to 3. We also included a scenario with mortality rates representative of those for SBT, as described below.

Table 1 specifies the parameter values used in the 19 scenarios we considered. Scenarios 1 to 9 all use the same fishing mortality vector, which increases from age 1

to age 3; the initial (age 1) cohort sizes and the variance in the age-length relationship vary. Scenarios 10 to 18 are the same as scenarios 1 to 9 except that fishing mortality decreases from age 1 to age 3. Scenario 19 is intended to be representative of SBT; the initial cohort sizes and age-specific fishing mortality rates are roughly based on estimates from the application to SBT in Appendix 5; the variance parameter for the model/measurement error in the age-length relationship is based on the estimate of the model/measurement error variance parameter for scientist-measured fish presented in Laslett et al. (2002).

Figures 1 to 5 illustrate the amount of overlap in the length-frequency distributions of age 1 to 3 fish in the three years of tagging corresponding to a selection of scenarios (because only fish of ages 1 to 3 are tagged, these are the only age-classes that need to be shown). The greater the overlap, the greater the number of age misclassifications that will be made based on length. The dashed vertical lines delimit the length cut-offs used to classify fish as age 1, 2 or 3 (e.g., any fish whose length falls between the first two lines will be estimated as age 1). Figures 1 to 3 correspond to scenarios 4 to 6. In these three scenarios all cohorts start out the same size at age 1, but the amount of variability in the age-length relationship, and therefore the amount of overlap in lengths of fish of adjacent ages, ranges from low in scenario 4 to high in scenario 6. Figures 4 and 5 correspond to scenarios 2 and 8, respectively, and illustrate the effect on the length-frequency distributions of having a particularly small or large cohort come through. They both have the same amount of variability in the age-length relationship as scenario 5 (Figure 3), but in scenario 2 the cohort of interest for modelling (i.e., the one tagged at age 1 in year 1, age 2 in year 2, and age 3 in year 3) is substantially smaller than the other cohorts, and in scenario 8 the cohort of interest is substantially larger than the other cohorts.

The length-frequency distributions look similar for scenarios 10 to 18 as for scenarios 1 to 9, and therefore are not shown. The two sets of scenarios are analogous except that in scenarios 1 to 9 fishing mortality increases with age, whereas in scenarios 10 to 18 it decreases with age. If we had chosen mortality rates that differed more dramatically in magnitude, then differences between the length-frequency distributions would become more noticeable.

Results

Figure 6 shows the average parameter estimates for scenarios 1 to 9. Generally speaking, when the size of the cohort being modelled is small relative to the other cohorts and/or the variance in the age-length relationship is high, negative biases emerge in the M_1, M_2 and F_3 estimates obtained using the estimated release-recapture data. However, only scenarios 2 and 3 have any biases over 20% (Table 2).

Figure 7 shows the average parameter estimates for scenarios 10 to 18, which are the same as scenarios 1 to 9 except the fishing mortality rates decrease with age instead of increasing. Again, the general pattern is that when the size of the cohort being modelled is small relative to the other cohorts and/or the variance in the age-length relationship is high, biases emerge in the M_1, M_2 and F_3 estimates obtained using the estimated release-recapture data; however, the biases are now positive instead of negative and also tend to be larger in magnitude. For example, scenarios 10, 11, 12, 14 and 15 all have biases of over 20%, with scenarios 11 and 12 having biases of over 80% (Table 2).

The most serious biases occur in scenarios where the size of the cohort being modelled is small compared to other cohorts. This is because cohorts adjacent to the one being modelled have relatively large numbers of fish and therefore the number of fish misclassified from these cohorts as belonging to the cohort of interest can be large and can form a significant fraction of the data for this small cohort (see Figure 4). On the other hand, in all of the scenarios where the size of the cohort being modelled is large compared to other cohorts, the parameter estimates obtained using the estimated release-recapture data are unbiased. This is because relatively few fish from adjacent cohorts will be misclassified as belonging to the cohort of interest even when the variability in length at age is large (see Figure 5). Even though a large number of fish from the cohort of interest may be misclassified as belonging to another cohort, errors in this direction are not serious because they just mean some 'correct' data will be omitted from the model. This will increase the variance of the parameter estimates compared to if all fish had been correctly aged but it will not introduce biases.

There are no large biases in any of the parameter estimates obtained using the true release-recapture data in scenarios 1 to 18 (Figures 6 and 7). This is an important observation because, as shown in Appendix 5 and as will be seen in scenario 19, mortality rates estimated from Brownie models can have inherent biases. We wanted to be sure that the biases observed using the estimated data were in fact due to errors in age classifications and not estimation biases.

The results for scenario 19, which is the scenario intended to be representative of SBT, are presented in Figure 8 and the last row of Table 2. There is evidence of a positive bias in the M_1 estimate and a negative bias in the M_2 estimate using not only the estimated release-recapture data but also the true data (refer to previous paragraph). The biases in the estimates obtained using the estimated data are larger than those using the true data; however, they are not so much larger that aging errors appear to be of great concern (for M_1, the bias is 13.1% using estimated data versus 7.2% using true data; for M_2, the bias is -10.3% using estimated data versus -3.5% using true data). Note that the percent biases for F_1 and F_2 are fairly large (Table 2), but this is a consequence of the true values being so small; the absolute biases are minimal (Figure 8).

Only the bias of the parameter estimates has been discussed so far, not the variance. The variance of the parameter estimates is influenced mainly by the sample size (i.e. the number of releases). We have seen that in scenarios where the cohort being modelled is small, a large number of fish can be misclassified as belonging to this cohort; in such scenarios, the variance can be smaller for the estimates obtained using the estimated release-recapture data than for those obtained using the true data (but the biases of the parameter estimates tend to be large). Generally, however, the variances of the parameter estimates are very similar whether the estimates were obtained using true or estimated data.

Discussion

Our simulation results suggest that in many situations the effects of aging errors on the mortality rate estimates obtained from a Brownie model are minimal. However, this is not always the case – in scenarios where the size of the cohort being modelled

is much smaller than adjacent cohorts and where the variability in lengths at age is large, some of the estimates (namely, fishing mortality at older ages and natural mortality) have large biases. It should be noted that in these scenarios, the differential in the size of adjacent cohorts is much greater than has been estimated for successive cohorts of SBT, and the variability in growth is also greater than estimated for SBT. The true mortality rates also affect the results; we saw larger biases in the parameter estimates when fishing mortality decreased with age than when it increased. In situations where mortality rates are the same at all ages, aging errors have no consequences, as we expected (this was confirmed using simulations but the results were not shown).

In our investigation, we only changed the age-length relationship by changing the variability in lengths at age, but similar results could have been achieved by changing the rate of growth. For example, consider scenario 4, which has growth parameters $k = 0.2$ and $\sigma_\varepsilon = 9$; if we change these to $k = 0.05$ and $\sigma_\varepsilon = 3$ and keep all other parameter values the same, then the length-frequency distributions and degree of overlap between lengths at ages look very similar (compare Figures 3 and 9).

The simulation results presented may underestimate the biases introduced by making errors in the ages of fish at release because the same mean growth curve used to generate lengths for fish was used subsequently to estimate age from length. In reality, the age-length relationship being used to estimate age from length will itself be estimated from data and therefore will not be such an exact description of the true age-length relationship for the fish being tagged. This will especially be true if the data used to estimate the age-length relationship are not representative of the fish being tagged. For example, the data may have been collected in previous years or from a different geographic location than tagging, and growth may vary over time and space. Furthermore, the fish being tagged may represent a particularly fast or slow growing subset of the population. Ideally, a sample of fish should be taken at the same time as tagging and the age-length relationship estimated from these data, either using length-frequency modal analyses if adequate data exists, or else using direct age and length data from hard-parts analyses (e.g. for SBT, otoliths would be removed from the sample and used to age the fish). We are assuming that age cannot be

determined using scales; otherwise a direct age estimate could be obtained for each fish being tagged and an age-length relationship would not be necessary.

For simplicity, we always used integer ages to generate lengths for fish being tagged. Then, for estimating the age-class of a fish, we calculated length cut-offs using the mean length at ages $a - 0.5$ and $a + 0.5$, with integer values for a. In doing so, we are assuming that tagging occurs around the same time as birth/spawning. In a true tagging experiment, fish may be tagged at any time of year and the length cut-offs should be adjusted accordingly based on the relative time of tagging to spawning; specifically, the length cut-offs should be calculated using the approximate decimal age of fish at the time of tagging. For example, if the average time of spawning occurs in January and fish are tagged in July, then fish are already half a year older at the time of tagging and length cut-offs should be calculated using $a = 0.5, 1.5, 2.5, ...$ rather than $a = 0, 1, 2, 3, ...$.

A key assumption in our investigation is that fishing and natural mortality are age-based, so that if fishing and natural mortality at age 1 are F_1 and M_1, respectively, then fish that we estimated to be age 2 but were truly age 1 will have experienced mortality of F_1 and M_1, but we will have assumed that they experienced mortality of F_2 and M_2. Of course, if fishing mortality is truly length-based instead of age-based then the underlying model is incorrect and interpretation of the age-specific mortality rate parameters becomes complicated.

Whether or not aging errors in a Brownie model are of concern depends on the situation and the researcher's purpose. Fortunately, in the situation most representative of SBT, the biases in the parameter estimates that resulted from using estimated release ages were not so large as to cause great concern. This is an important finding because it means we do not expect the results presented throughout this report, and especially in Appendix 12, from applying Brownie models (and variations thereof) to SBT data to be unduly biased due to treating the estimated release ages as precise.

In cases where there is potential for aging errors from cohort slicing to introduce substantial biases, an alternative to relying on cohort slicing for aging of tagged fish should be considered in the experimental design phase. These would include:

1. Collection of a scale from each tagged fish if scales can be used to reliably estimate age for the species being tagged;

2. Collection of otoliths from a sample of fish during tagging operations in order to be able to produce an age-length key that could be directly applied to the tagged fish;

3. Incorporation of aging error directly into the estimation framework.

It should be noted that use of an age-length key would still necessitate the incorporation of aging errors directly into the estimation framework. This is because an age-length key does not provide a unique estimate of the age for each tagged fish but only a probability distribution for its age given its length. Incorporation of aging errors directly into the estimation framework is conceptually straightforward. It would involve specifying a likelihood function for the age of each tagged fish given its length. Then, conditional on its age of release, the other likelihood functions developed in this report can be used to estimate the conditional likelihood of recapturing a tagged fish. These conditional likelihoods would then need to be integrated over the possible ages of releases to provide an overall unconditional likelihood. While conceptually straightforward, it would be computationally complex and time consuming. Thus, in any specific application it is important to determine whether aging errors are likely to be important early in the design phase of the experiment.

References

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.

Laslett, G. M., Eveson, J. P., and Polacheck, T. 2002. A flexible maximum likelihood approach for fitting growth curves to tag-recapture data. Can. J. Fish. Aquat. Sci. **59**: 976-986.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1996. Estimation of natural and fishing mortality rates for juvenile southern bluefin tuna based on multi-year tagging of cohorts. Commission for the Conservation of Southern Bluefin Tuna CCSBT/SC/96/11.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1998. Updated estimates of mortality rates for juvenile southern bluefin tuna from multi-year tagging of cohorts. Council for the Conservation of Southern Bluefin Tuna CCSBT/SC/9807/20.

Polacheck, T., Eveson, J. P., and Laslett, G. M. 2004. Increase in growth rates of southern bluefin tuna (Thunnus maccoyii) over four decades: 1960 to 2000. Can. J. Fish. Aquat. Sci. **61**: 307-322.

Pollock, K. H., Hoenig, J. M., and Jones, C. M. 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.

Table 1. Parameter values used for generating data in each of the scenarios considered. In all scenarios $\mu_\infty = 180$, $k = 0.2$, $a_0 = -0.5$, $\sigma_\infty = 10$, $N_T = 3000$, natural mortality is constant across ages and years, and fishing mortality is constant across years.

Parameter	Scenario																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
σ_ε	3	6	9	3	6	9	3	6	9	3	6	9	3	6	9	3	6	9	3
$P_{1,-1}$ ($\times 10^5$)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	40
$P_{1,0}$ ($\times 10^5$)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	35
$P_{1,1}$ ($\times 10^5$)	0.2	0.2	0.2	1	1	1	5	5	5	0.2	0.2	0.2	1	1	1	5	5	5	30
$P_{1,2}$ ($\times 10^5$)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
$P_{1,3}$ ($\times 10^5$)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	20
$F_{1,t}$	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.01
$F_{2,t}$	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.05
$F_{3,t}$	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2
$F_{4,t}$	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2
$F_{5,t}$	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.2
$M_{a,t}$	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2

Table 2. Percent bias in parameter estimates obtained using estimated release-recapture data (100 simulations per scenario). The true parameter values are given in Table 1.

Scenario	F_1	F_2	F_3	M_1	M_2
1	2.8	5.0	-0.8	10.7	-13.0
2	12.1	2.4	-12.6	-14.5	-29.3
3	25.1	-2.5	-21.9	-52.6	-33.8
4	1.5	1.4	-0.5	5.5	-4.2
5	4.0	0.4	-4.6	-0.5	-10.8
6	7.5	-1.2	-9.1	-13.6	-14.8
7	0.1	0.7	1.1	-0.1	4.3
8	0.9	0.5	-0.3	-0.1	0.1
9	1.8	0.3	-1.6	-2.8	-1.3
10	-3.7	1.0	22.9	-13.0	52.9
11	-3.9	6.3	58.4	23.3	92.0
12	-3.2	12.5	79.7	84.7	82.2
13	-0.4	1.5	8.2	0.5	16.2
14	-0.6	4.0	20.4	8.3	39.0
15	0.1	8.3	36.0	29.0	55.1
16	-0.1	-0.1	1.1	-0.4	1.1
17	0.0	0.3	2.8	2.2	3.3
18	0.3	2.1	7.2	8.6	10.2
19	16.5	10.2	-3.2	13.1	-10.3

Figure 1. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 4 (see Table 1). The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length falls between the two left-most vertical lines will be estimated as age 1). The cohort being modelled is the age 1 fish in year 1 (panel 1), age 2 in year 2 (panel 2), and age 3 fish in year 3 (panel 3).
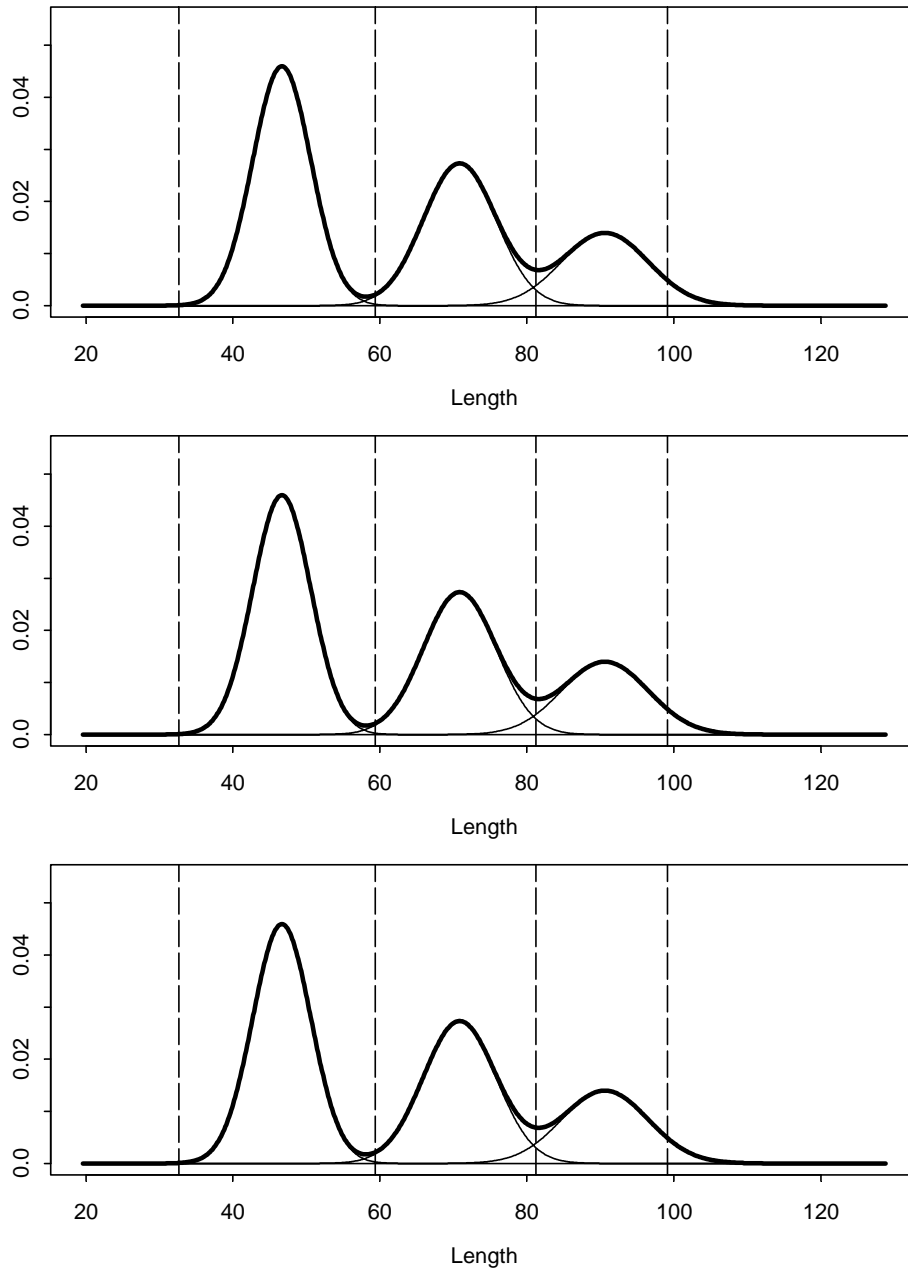
Figure 2. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 5 (see Table 1). The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length falls between the two left-most vertical lines will be estimated as age 1). The cohort being modelled is the age 1 fish in year 1 (panel 1), age 2 in year 2 (panel 2), and age 3 fish in year 3 (panel 3).
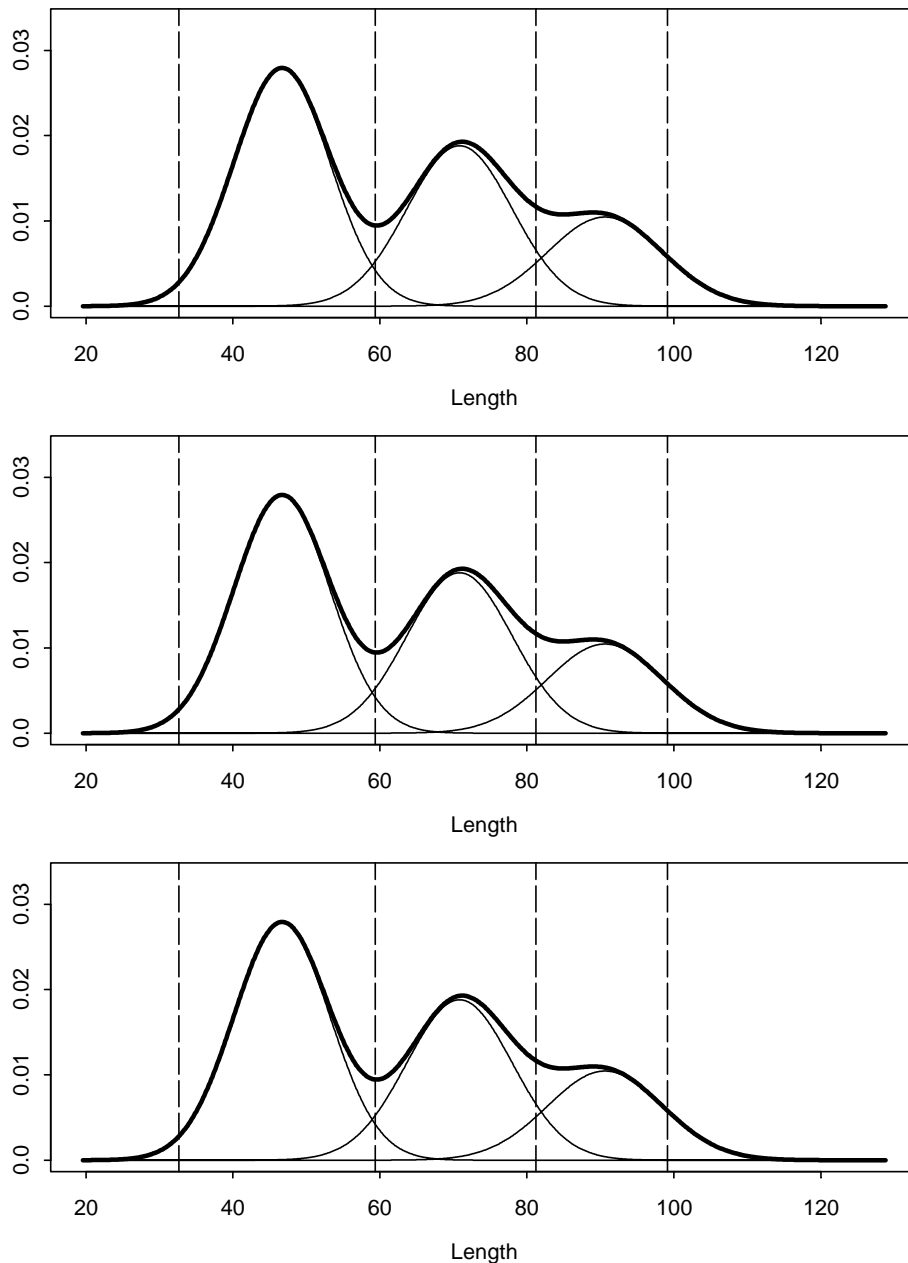
Figure 3. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 6 (see Table 1). The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length falls between the two left-most vertical lines will be estimated as age 1). The cohort being modelled is the age 1 fish in year 1 (panel 1), age 2 in year 2 (panel 2), and age 3 fish in year 3 (panel 3).
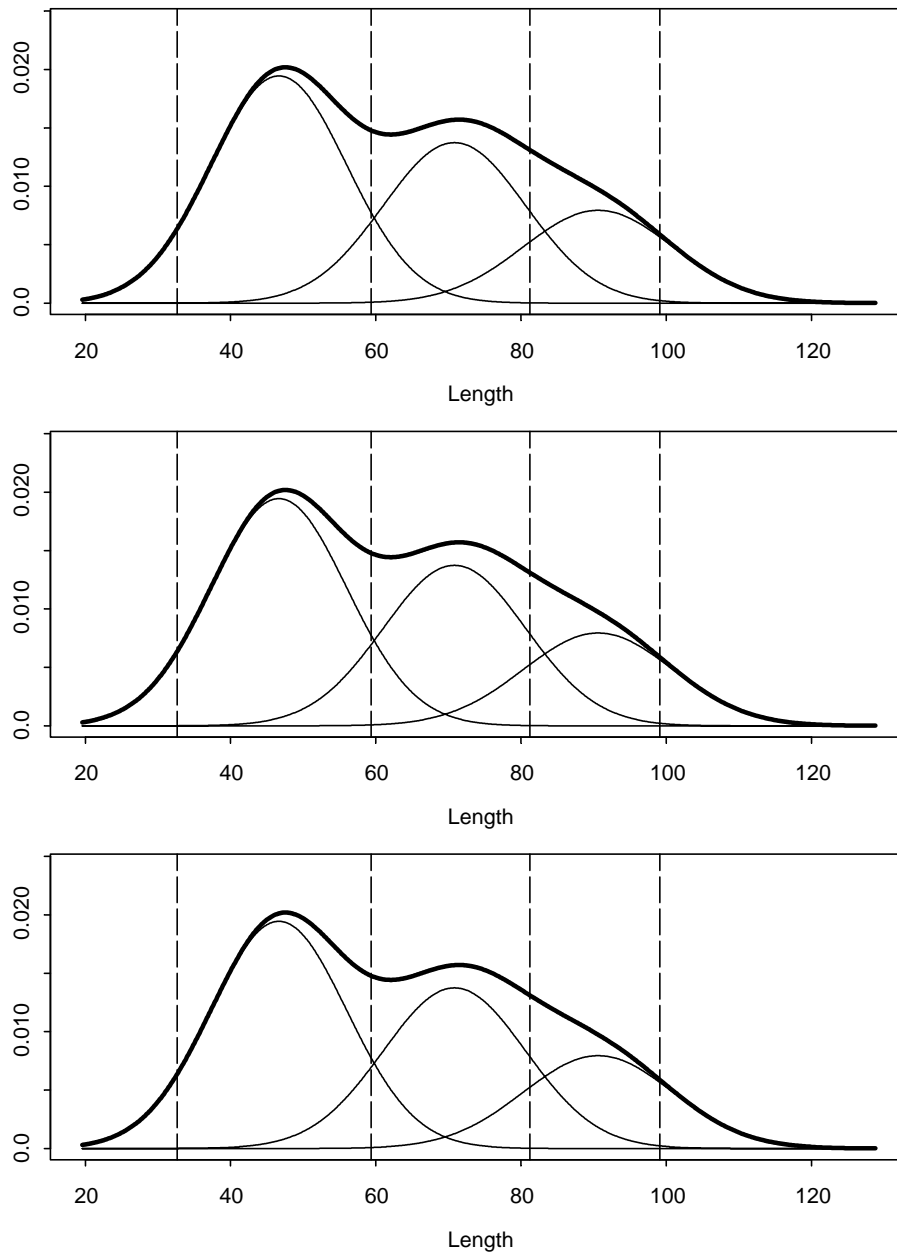
Figure 4. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 2 (see Table 1). The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length falls between the two left-most vertical lines will be estimated as age 1). The cohort being modelled is the age 1 fish in year 1 (panel 1), age 2 in year 2 (panel 2), and age 3 fish in year 3 (panel 3).
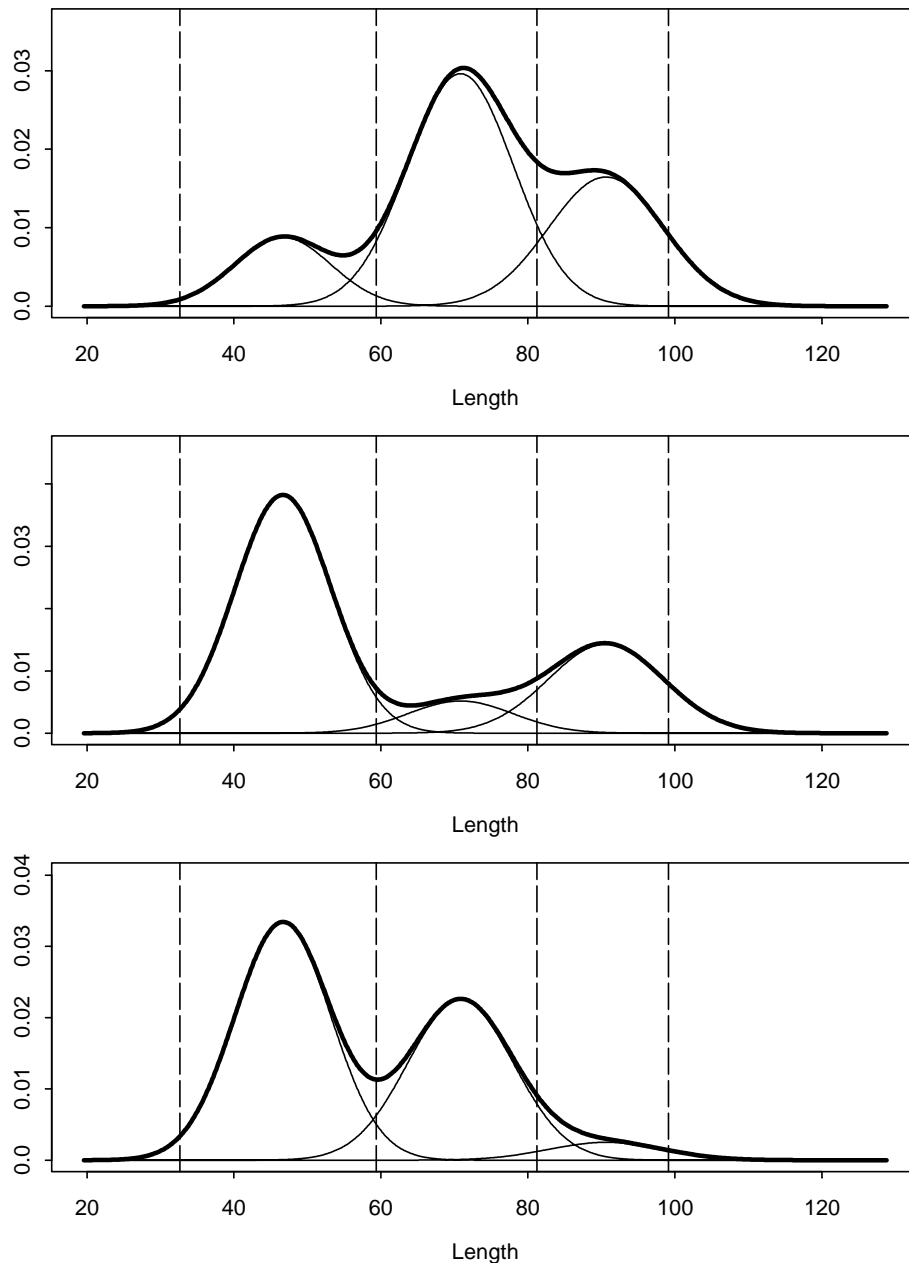
Figure 5. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 8 (see Table 1). The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length fall between the first two vertical lines will be estimated to be age 1).
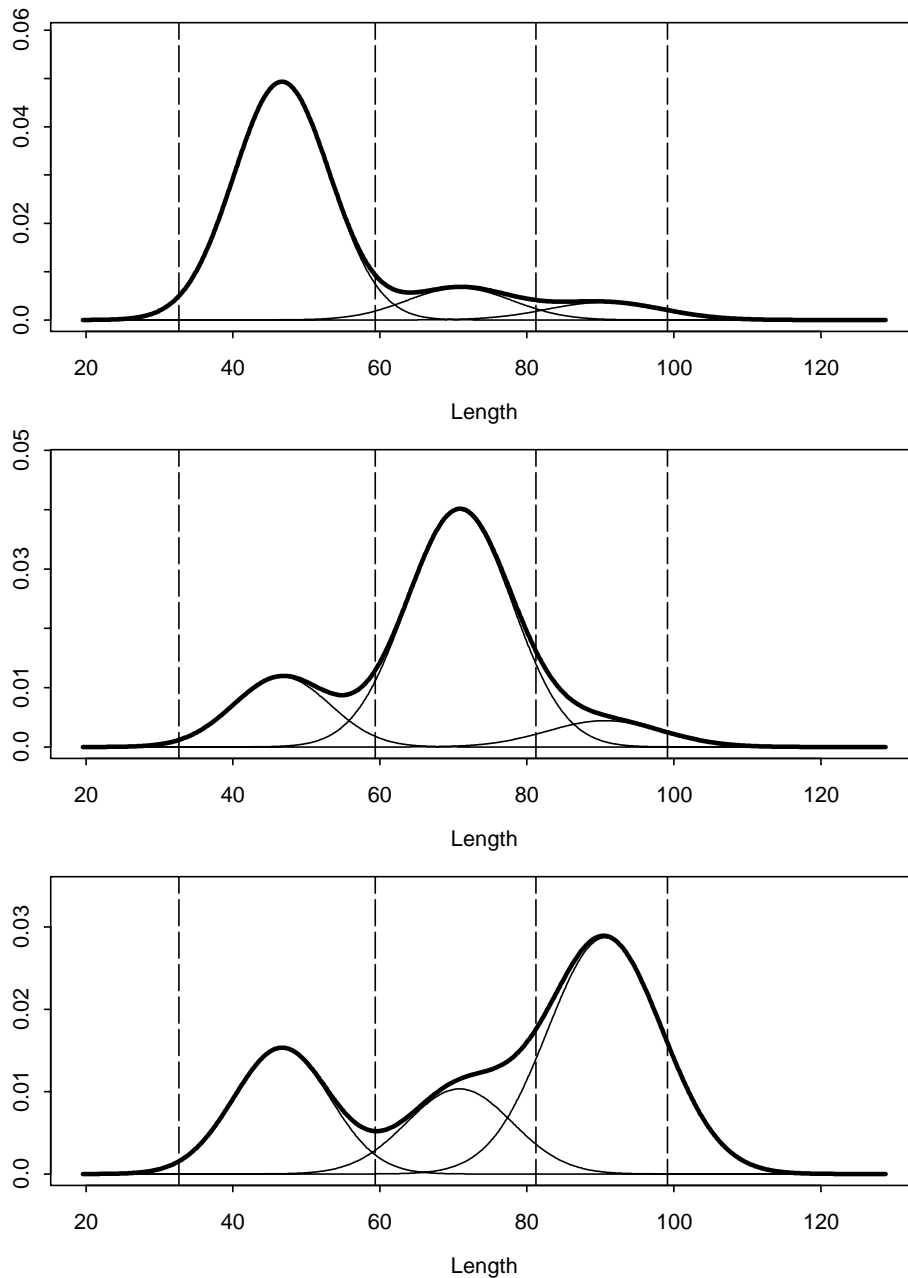
Figure 6. Average parameter estimates from fitting a Brownie model to 100 "true" and "estimated" release-recapture data sets corresponding to scenarios 1 to 9 (upper left panel to lower right panel, by row). Plus = true parameter value; circle = estimated parameter value using "true" data; triangle = estimated parameter value using "estimated" data. (Note: $\sigma = \sigma_\varepsilon$.)
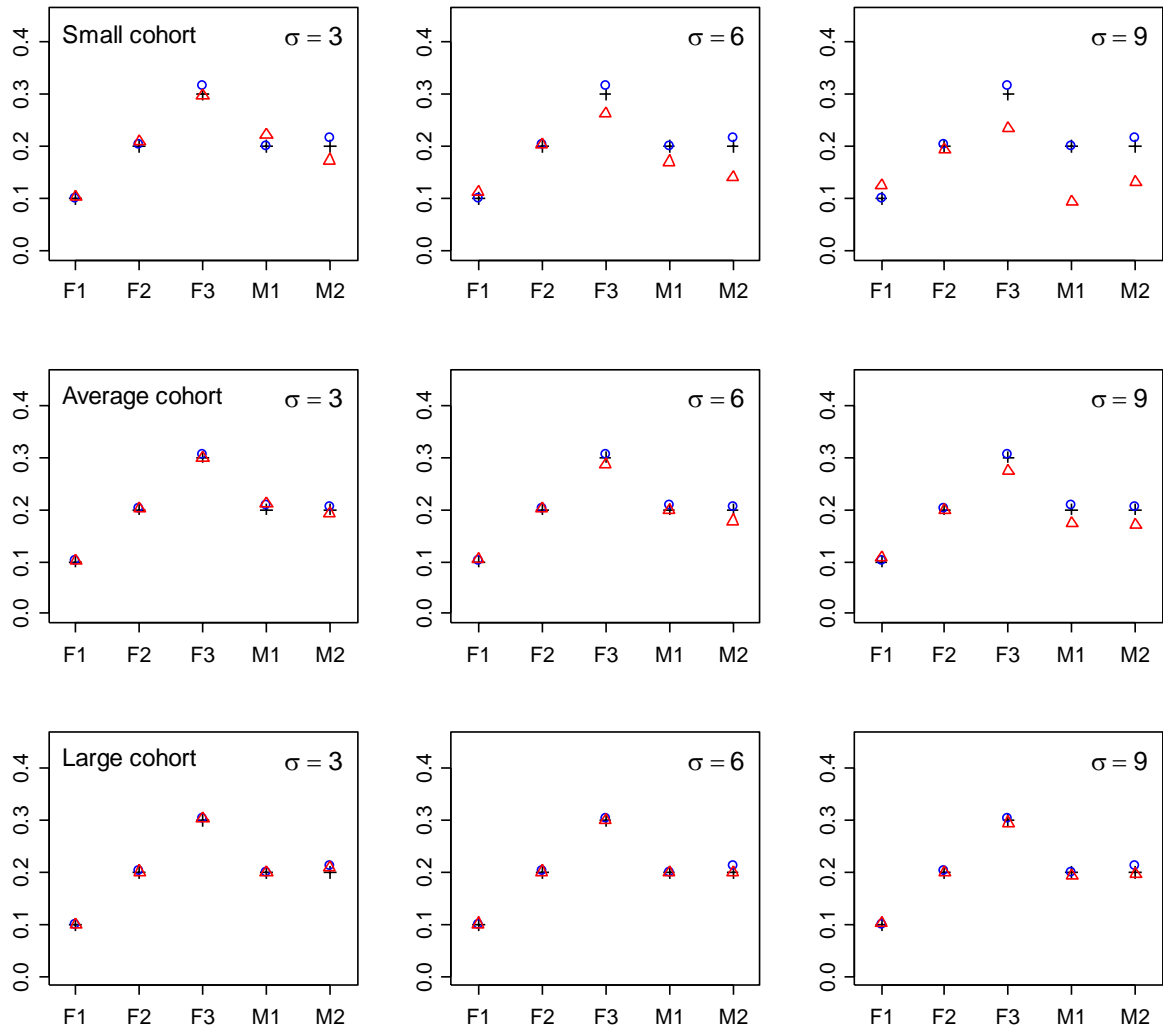
Figure 7. Average parameter estimates from fitting a Brownie model to 100 "true" and "estimated" release-recapture data sets corresponding to scenarios 10 to 18 (upper left panel to lower right panel, by row). Plus = true parameter value; circle = estimated parameter value using "true" data; triangle = estimated parameter value using "estimated" data. (Note: $\sigma = \sigma_\varepsilon$.)
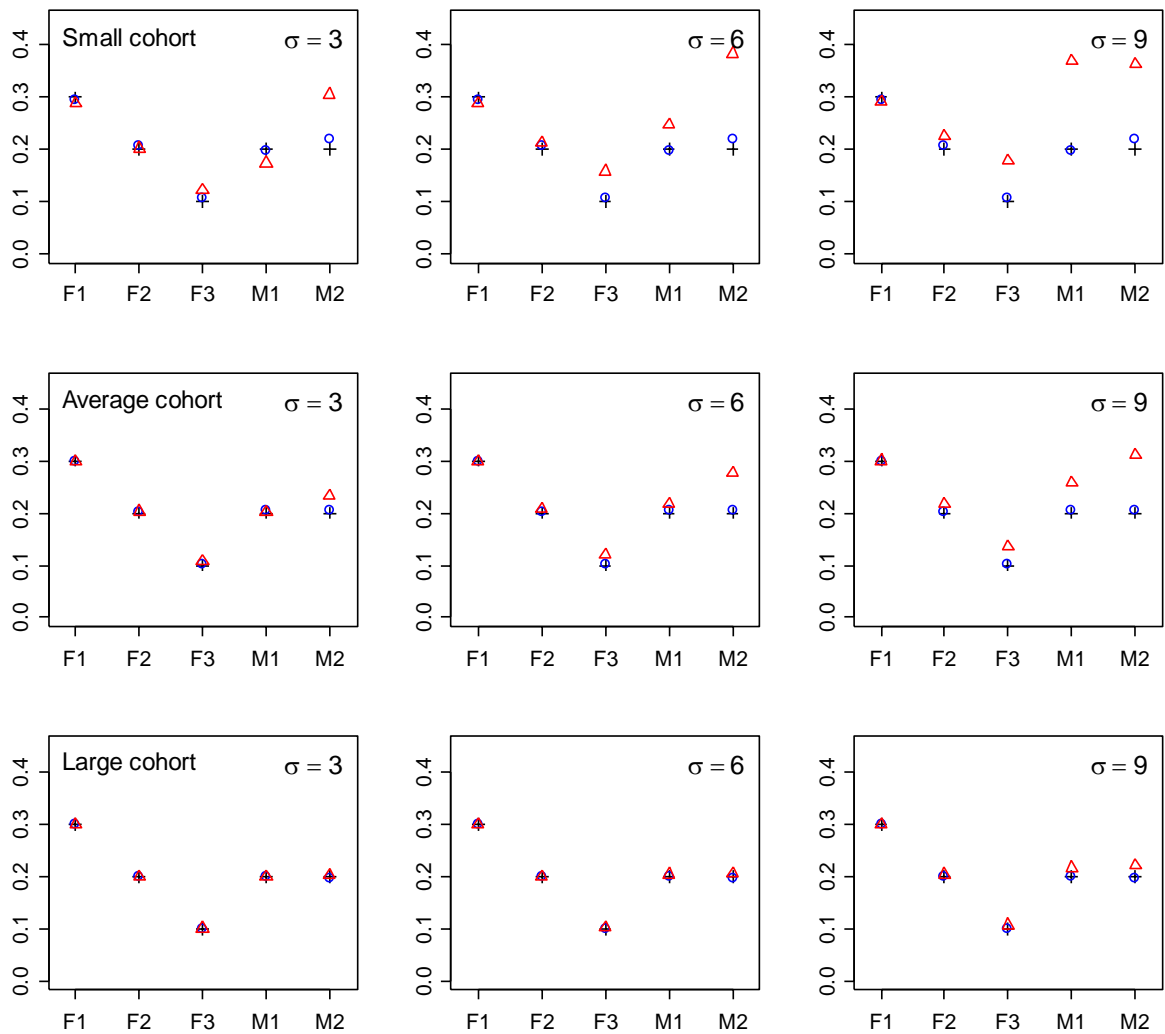
Figure 8. Average parameter estimates from fitting a Brownie model to 100 "true" and "estimated" release-recapture data sets corresponding to scenario 19, intended to be representative of southern bluefin tuna data. Plus = true parameter value; circle = estimated parameter value using "true" data; triangle = estimated parameter value using "estimated" data.
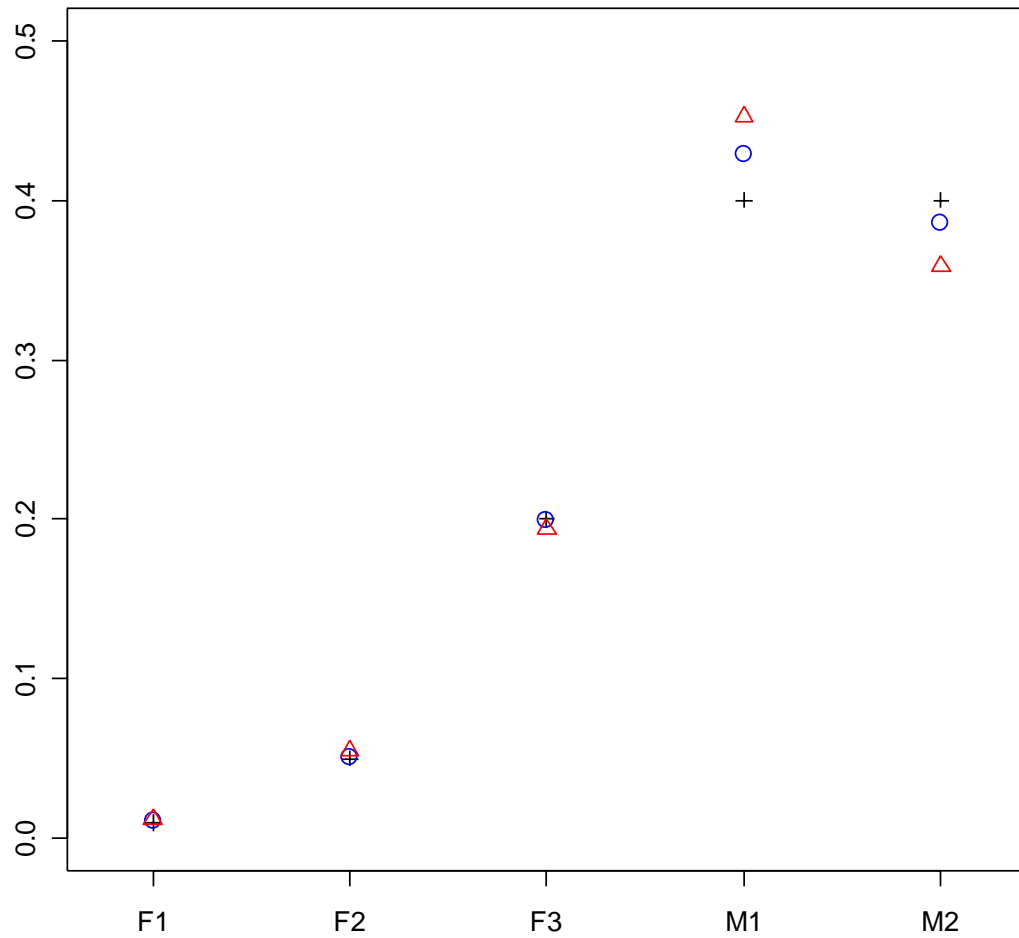
Figure 9. Length-frequency distribution of age 1 to 3 fish in the three years of tagging corresponding to scenario 4 (see Table 1), but with growth rate parameter k equal to 0.05 instead of 0.2. The thick solid line shows the overall distribution; the thin solid lines show the breakdown into age components. The dashed vertical lines delimit the length cut-offs used to estimate the age of a fish based on its length (e.g., any fish whose length falls between the two left-most vertical lines will be estimated as age 1). The cohort being modelled is the age 1 fish in year 1 (panel 1), age 2 in year 2 (panel 2), and age 3 fish in year 3 (panel 3).
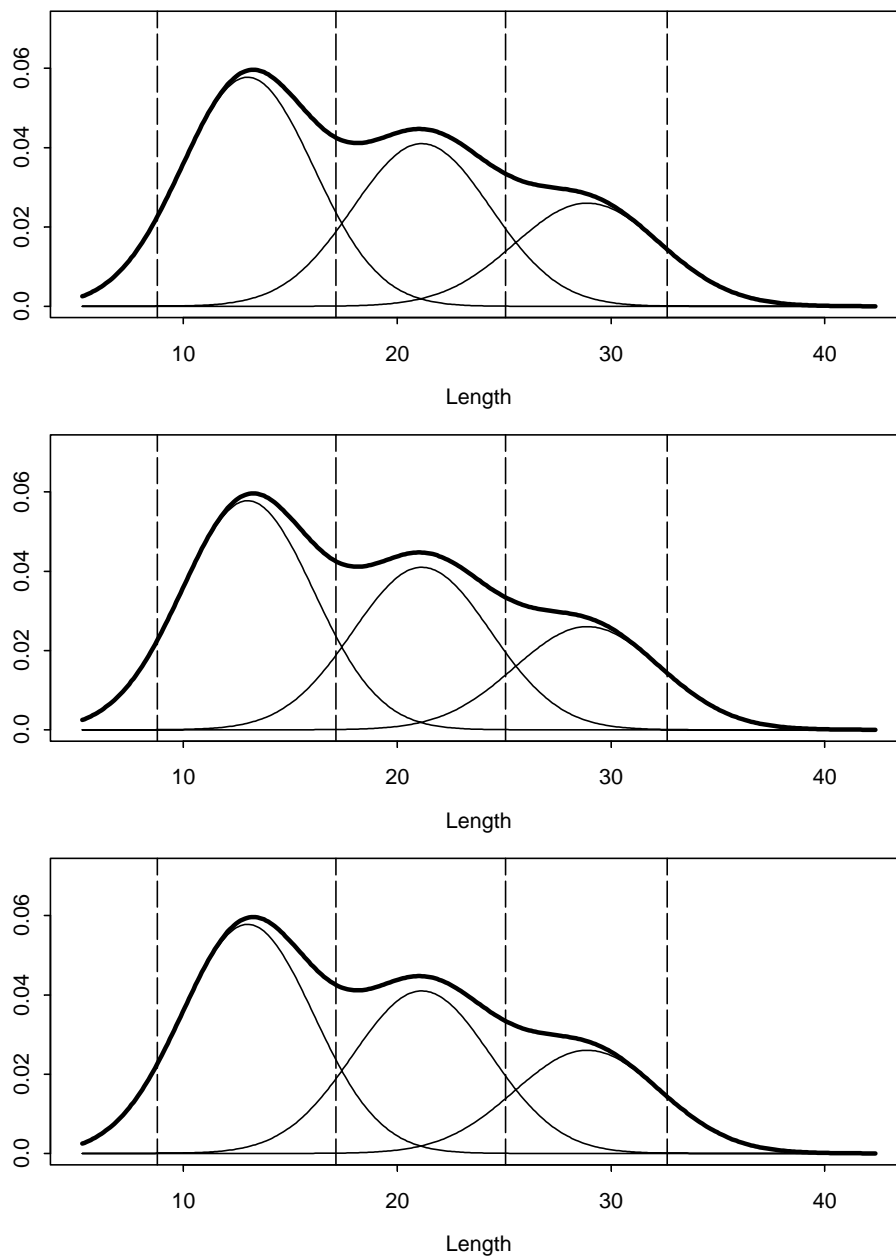
Appendix 9:

Exploring trade-offs in experimental design of a 2-fishery integrated tag-recapture and catch model for estimating mortality rates and abundance

J. Paige Eveson, Tom Polacheck and Geoff M. Laslett

FRDC Project 2002/015

Table of Contents

Abstract

The integrated tag-recapture and catch-at-age model for estimating mortality rates and abundance developed in CCSBT-ESC/0309/22 is extended to a two-fishery situation with a surface purse seine fishery and a longline fishery intended to resemble the southern bluefin tuna (SBT) situation. We also extended the model to allow for overdispersion in the tag return data. Tag reporting rates are assumed to differ between the two fisheries, and they are estimated from tag seeding data in the surface fishery and from observer data in the longline fishery. Simulations are used to investigate design issues for the tagging program currently being conducted on SBT as part of the CCSBT Scientific Research Program (SRP), in particular, to investigate levels of observer coverage and tag releases necessary to achieve reasonable precision in mortality rate and abundance estimates. The results suggest that the number of tags that have been released in recent years as part of the CCSBT SRP tagging program are adequate, but that increasing observer coverage from current levels could potentially lead to significant improvements in the precision of the fishing mortality rate estimates for the longline fishery, as well as smaller improvements in the estimate of population abundance. The results from the model with overdispersion in the tag returns suggest that in order to achieve coefficients of variation of 20% or less for the longline fishing mortality rates at ages 1 to 3, observer coverage must be at least 30% (and at least 20% for the model without overdispersion). Estimates of fishing mortality in the surface fishery are chiefly unaffected by the level of observer coverage in the longline fishery, provided fairly accurate estimates of surface fishery reporting rates and catch-at-age by fishery exist. It is important to note, however, that these results depend on the assumption of complete mixing. If this assumption is violated, then the level of observer coverage in the longline fishery would become more influential on the accuracy and precision of parameter estimates. Without good observer data, and thus good information on differential tag reporting and return rates between fishery components, there would be little power to test the assumption of non-mixing and, if necessary, develop spatially-explicit tag recovery models to account for heterogeneity in recapture probabilities. The results also demonstrate the importance of having reliable and precise estimates of the catch-at-age for each fishery when applying the estimation model presented here. This emphasizes the need to develop appropriate sampling and error models for the catch data; having representative and adequate observer coverage can help to accomplish this in the longline fishery.

Introduction

In CCSBT-ESC/0309/22, we developed an integrated Brownie and Petersen model for estimating abundance and mortality rates (fishing and natural) from multi-year tagging programs and estimates of the catch-at-age data. We explored the situation in which reporting rates are known in the catch-at-age data, and the situation in which reporting rates are estimated from observer data from a portion of the fleet. We showed that combining the catch-at-age data with the multi-year tagging data allows for population abundance to be estimated directly from the model and also provides additional information for estimating mortality rates. We also presented results from simulations on how the relative trade-off between effort put into tagging and observers affects the overall mortality and abundance estimates. The results suggested that observer levels of 20-30% (or even greater) may be required to achieve reasonable levels of precision in the parameter estimates, and raised concerns about whether the current tagging program being undertaken as part of the Council for the Conservation of Southern Bluefin Tuna Scientific Research Program (CCSBT SRP) (Anon. 2001a) will be able to meet its primary objective of being able to estimate mortality rates for southern bluefin tuna (SBT) with sufficient levels of precision to substantially improve the SBT stock assessment.

The results in CCSBT-ESC/0309/22 were based on consideration of a single fishery in which the only source of information for estimating reporting rates was from observers. However, juvenile SBT, which are the target of the SRP tagging program, are harvested both by purse seine and longline vessels. These gears have different age-specific selectivities, different levels of catch, and tag reporting rates are also likely to vary between the purse seine and longline fisheries (and in the case of the latter, there are multiple fleets in which reporting rates are also likely to vary). In addition, observers in the SBT purse seine fishery cannot provide any useful data by which to estimate reporting rates since captured fish are transferred without being removed from the water to cages for farming. Instead, tag seeding is being used to obtain estimates of reporting rates from this fishery component (Stanley and Polacheck 2003; Polacheck and Stanley 2004).

In the current paper, the estimation model in CCSBT-ESC/0309/22 is extended to a two-fishery situation with a purse seine fishery (referred to as the surface fishery) and a single longline fishery. We have reduced the multiple longline fleets to a single fishery with a uniform level of observer coverage and a uniform reporting rate both to simplify the presentation and because exploration of the possible trade-offs in observer coverage amongst different longline fisheries did not seem fruitful given the commitment of the CCSBT to have similar observer target levels in all fisheries. In our model, we allow for fishing mortality rates, as well as reporting rates, to differ between the two fisheries. Reporting rates are estimated from tag seeding data in the surface fishery and from observer data in the longline fishery. We present results on how the amount of effort put into tag releases and observers affects the mortality rate and abundance estimates. We have conditioned the simulations used to generate these results so that they qualitatively resemble the SBT situation.

The motivation for this papers stems from decisions made at the 2003 CCSBT Scientific Committee meeting (Anon. 2003). This meeting concluded that the current levels of observer coverage in the Japanese, Korean and Taiwanese longline fleets are not high enough to provide useful estimates of reporting rates, and thus fishing mortality rates, from these fleets. The overall implication of this conclusion for the ability of the SRP tagging program to meet its primary objectives were not certain because of the differential and much higher reporting

rate in the Australian surface fishery, combined with the fact that the Australian surface fishery catches a substantial portion of the global catch of juvenile SBT. As such the Scientific Committee agreed to convene a Technical Group Meeting in conjunction with its next meeting to deal with this question. Among the terms of reference agreed to for this Technical Group are:

1. To evaluate the level of precision of mortality and abundance estimates that the current tagging program will be likely to provide at current levels of observer coverage and anticipated tag recovery rates (given current efforts directed at increasing them).
2. To evaluate the levels of observer coverage and tag recovery rates that would be required for the tagging program to provide acceptable levels of precision in key mortality and abundance estimates, and how these are influenced by model assumptions.

The results presented here will hopefully assist the Technical Group in its deliberations.

Methods

Estimation model

Underlying the integrated tag and catch model used here are the general population dynamics and catch equations commonly used in fisheries. These equations, presented in CCSBT-ESC/0309/22 and repeated below for fluidity, are expressed in terms of exponential and competing natural and fishing mortality rates. For a cohort of animals of a given age, the number that survive and the number that are caught are given by:

$$P_{i,t+1} = P_{i,t} \exp\{-F_{i,t} - M_{i,t}\} \qquad (1)$$

$$C_{i,t} = \frac{F_{i,t}}{F_{i,t} + M_{i,t}} P_{i,t}\left(1 - \exp\{-F_{i,t} - M_{i,t}\}\right) \qquad (2)$$

where

$P_{i,t}$ = the number of individuals of age i at time t
$C_{i,t}$ = the catch of individuals of age i at time t
$F_{i,t}$ = the instantaneous fishing mortality rate for individuals of age i at time t
$M_{i,t}$ = the instantaneous natural mortality rate for individuals of age i at time t.

In the context of a tagging experiment, the above equations provide the basis for predicting the expected number of returns, assuming that the tagged fish constitute a representative sample of the population. In the current paper, we consider a multi-year tagging experiment involving only a single cohort of fish (tagged at consecutive ages). As such, age and year provide equivalent information, and we can simplify the notation by dropping the reference to year (i.e. the t subscript in the above equations) and expressing everything in terms of age.

In the two-fishery model presented here, we allow for natural mortality to differ between ages, and fishing mortality to differ between ages and fisheries. We also allow fishing mortality to differ between tagged fish in the year of tagging and untagged fish in that year

(following the model presented in Hoenig et al. 1998). This is to allow for the fact that tagged and untagged fish will not be fully mixed directly after tagging, and also because in the SBT situation much of the tagging occurs near the end of the fishing season. This is done in order to prevent a large number of immediate returns, but will obviously mean that fishing mortality in that year will not be the same for tagged and untagged fish. As Hoenig et al. (1998) point out, technically, the model formulation assumes that the relative timing of fishing and natural mortality for tagged fish in the first year after tagging is the same as that for untagged fish and fully mixed tagged fish in subsequent years. However, this is not a critical assumption because the relative timing has only a minor effect on estimation of natural mortality, and furthermore, we are not interested in fishing mortality of newly tagged fish.

Reporting rates are estimated differently for the surface fishery than for the longline fishery. For the longline fishery, we assume observers are on board a percentage of vessels and that the reporting rate is 100% for fish caught on these vessels. Because on average the fishing mortality have been assumed to be the same for all longline vessels, the relative return rate between the observed catches and the unobserved catches provides an estimate of the reporting rate in the unobserved component (i.e. observer coverage is representative of the entire longline fleet). For the surface fishery, we assume reporting rates are estimated from independent data, such as tag seeding data, and that we have estimates of reporting rates, along with standard errors on the estimates, to use in our model.

We divide the tag returns and the corresponding catches into those coming from the surface fishery, the observed component of the longline fishery, and the unobserved component of the longline fishery. However, before proceeding we introduce the following notation:

Table 1. Data (to be inputted into the model).

N_a = number of tag releases of age a fish from a particular cohort

$R_{a,i}^S$ = tag returns of age i fish that were tagged at age a from the surface fishery

$R_{a,i}^{LL,obs}$ = tag returns of age i fish that were tagged at age a from the observed component of the longline fishery

$R_{a,i}^{LL,unobs}$ = tag returns of age i fish that were tagged at age a from the unobserved component of the longline fishery

$\hat{\lambda}_i^S$ = estimated tag reporting rate for fish recaptured at age i in the surface fishery

s_i = standard error of $\hat{\lambda}_i^S$

C_i^S = number of age i fish from the cohort of interest caught in the surface fishery

$C_i^{LL,obs}$ = number of age i fish from the cohort of interest caught in the observed component of the longline fishery

$C_i^{LL,unobs}$ = number of age i fish from cohort of interest caught in the unobserved component of the longline fishery

δ_i = proportion of age i fish in the observed component of the longline fishery

υ_i^S = coefficient of variation of age i catch data from the surface fishery

$\upsilon_i^{LL,obs}$ = coefficient of variation of age i catch data from the observed component of the longline fishery

Table 2. Parameters (to be estimated in the model).

λ_i^S = tag reporting rate for fish captured at age i in the surface fishery

λ_i^{LL} = tag reporting rate for fish captured at age i in the unobserved component of the longline fishery

M_i = instantaneous natural mortality rate for age i fish

F_i^S = instantaneous fishing mortality rate in the surface fishery for age i fish (excluding fish tagged at age i)

F_i^{LL} = instantaneous fishing mortality rate in the longline fishery for age i fish (excluding fish tagged at age i)

F_i^{*S} = instantaneous fishing mortality rate in the surface fishery for age i fish tagged at age i

F_i^{*LL} = instantaneous fishing mortality rate in the longline fishery for age i fish tagged at age i

P_1 = initial population size (at age 1) of tagged cohort

ω_a = Dirichlet variance parameter for the return probabilities of fish released at age a; *only used in model with overdispersion in tag returns*

Now define:

$$F_i^{*Tot} = F_i^{*S} + F_i^{*LL}$$

$$S_i^* = \exp\{-(F_i^{*Tot} + M_i)\}$$

$$f_i^* = \frac{F_i^{*Tot}}{F_i^{*Tot} + M_i}(1 - S_i^*)$$

$$F_i^{Tot} = F_i^S + F_i^{LL}$$

$$S_i = \exp\{-(F_i^{Tot} + M_i)\}$$

$$f_i = \frac{F_i^{Tot}}{F_i^{Tot} + M_i}(1 - S_i)$$

Note that S_i^* represents the survival rate of age i fish tagged at age i; S_i represents the survival rate of age i fish, excluding those tagged at age i; f_i^* represents the exploitation rate of age i fish tagged at age i; and f_i represents the exploitation rate of age i fish, excluding those tagged at age i[1].

First consider the tag-recapture component of the model. The probability that an age i fish that was tagged at age a is returned from the surface fishery is:

---

[1] In the model developed here, the longline and surface fisheries are modelled as taking place throughout the year and their respective fishing mortalities constitute competing risks. In fact, for SBT, the surface and longline fisheries take place almost sequentially. While this detail could be added to the model without much difficulty, it would not be expected to change the general results presented in this paper. However, in an application in which fishing mortality rates are relatively high in one or both fisheries, it could have a small effect on the model predictions of the number of returns. In such situations, the modifications necessary for dealing with sequential fisheries might be worth including in the estimation model.

$$p_{a,i}^{S} = \begin{cases} \lambda_i^S \dfrac{F_i^{*S}}{F_i^{*Tot}} f_i^* & i = a \\[3mm] \lambda_i^S \dfrac{F_i^{S}}{F_i^{Tot}} S_a^* f_i & i = a+1 \\[3mm] \lambda_i^S \dfrac{F_i^{S}}{F_i^{Tot}} S_a^* \left(\displaystyle\prod_{k=a+1}^{i-1} S_k \right) f_i & i > a+1 \end{cases} \qquad (3)$$

The probability that an age i fish that was tagged at age a is returned from the observed component of the longline fishery is:

$$p_{a,i}^{LL,obs} = \begin{cases} \delta_i \dfrac{F_i^{*LL}}{F_i^{*Tot}} f_i^* & i = a \\[3mm] \delta_i \dfrac{F_i^{LL}}{F_i^{Tot}} S_a^* f_i & i = a+1 \\[3mm] \delta_i \dfrac{F_i^{LL}}{F_i^{Tot}} S_a^* \left(\displaystyle\prod_{k=a+1}^{i-1} S_k \right) f_i & i > a+1 \end{cases} \qquad (4)$$

The probability that an age i fish that was tagged at age a is returned from the unobserved component of the longline fishery is:

$$p_{a,i}^{LL,unobs} = \begin{cases} (1-\delta_i)\lambda_i^{LL} \dfrac{F_i^{*LL}}{F_i^{*Tot}} f_i^* & i = a \\[3mm] (1-\delta_i)\lambda_i^{LL} \dfrac{F_i^{LL}}{F_i^{Tot}} S_a^* f_i & i = a+1 \\[3mm] (1-\delta_i)\lambda_i^{LL} \dfrac{F_i^{LL}}{F_i^{Tot}} S_a^* \left(\displaystyle\prod_{k=a+1}^{i-1} S_k \right) f_i & i > a+1 \end{cases} \qquad (5)$$

These probability statements assume no mortality due to tagging and no tag shedding. If these assumptions are not met, additional parameters and potentially additional data will need to be introduced to account for these factors.

For tags released at a particular age, the numbers of returns by age from all sources (i.e. the surface fishery, the observed component of the longline fishery, and the unobserved component of the longline fishery), as well as those tags not returned, are expected to be multinomial with probabilities given in equations (3), (4) and (5). Thus, the likelihood equation for the tag return data corresponding to all release ages is:

$$L_R = \gamma \times \prod_a \left(\left(1 - p_{a,\bullet}^{Tot}\right)^{N_a - R_{a,\bullet}^{Tot}} \prod_{i \geq a} \left(\left(p_{a,i}^{S}\right)^{R_{a,i}^{S}} \left(p_{a,i}^{LL,obs}\right)^{R_{a,i}^{LL,obs}} \left(p_{a,i}^{LL,unobs}\right)^{R_{a,i}^{LL,unobs}} \right) \right) \qquad (6)$$

where

$$\gamma = \prod_a \frac{N_a}{\left(N_a - R_{a,\bullet}^{Tot}\right)! \prod_{i \geq a}\left(R_{a,i}^S! R_{a,i}^{LL,obs}! R_{a,i}^{LL,unobs}!\right)},$$

$$R_{a,\bullet}^{Tot} = \sum_i \left(R_{a,i}^S + R_{a,i}^{LL,obs} + R_{a,i}^{LL,unobs}\right),$$

and

$$p_{a,\bullet}^{Tot} = \sum_i \left(p_{a,i}^S + p_{a,i}^{LL,obs} + p_{a,i}^{LL,unobs}\right).$$

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

Estimates of the fishing and natural mortality rates (F^*'s, F's and M's) can be obtained from the multi-year tagging data by maximizing the above likelihood. Note, however, that the 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 $i+1$; thus, in an experiment with n consecutive release years, estimates can only be obtained for $n-1$ of the natural mortality rate parameters (regardless of the number of recapture years). Estimates of the reporting rates for the longline fishery (λ_i^{LL}'s) can also be obtained from the above likelihood using the differential between the return rates from the observed and unobserved catches, provided the ratios of observed to unobserved catches (δ_i's) are known.

There is not enough information in likelihood (6) to be able to estimate the reporting rates from the surface fishery. We assume instead that an estimate of the reporting rate at each age ($\hat{\lambda}_i^S$) and an associated standard error (s_i) has been obtained from independent tag seeding data, and that the estimate follows a standard beta distribution with mean λ_i^S and variance approximated by s_i^2. We chose a standard beta distribution because it gave a reasonably bell-shaped distribution that was constrained to lie between 0 and 1 (as desired for reporting rates). Thus, the likelihood component for the surface fishery reporting rate data is:

$$L_{\lambda^S} = \prod_i \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)}\left(\hat{\lambda}_i^S\right)^{\alpha_i-1}\left(1 - \hat{\lambda}_i^S\right)^{\beta_i-1} \tag{7}$$

where

$$\alpha_i = \left(\frac{\lambda_i^S}{s_i}\right)^2\left(1 - \lambda_i^S\right) - \lambda_i^S$$

and

$$\beta_i = \frac{\alpha_i\left(1 - \lambda_i^S\right)}{\lambda_i^S}.$$

Note that $\Gamma(\cdot)$ denotes the gamma function.

Now consider the catch component of the model. The probability that an age i fish from the cohort being studied is caught in the surface fishery is:

$$
\pi_i^S = \begin{cases} \dfrac{F_i^S}{F_i^{Tot}} f_i & i = 1 \\[4mm] \dfrac{F_i^S}{F_i^{Tot}} \left(\displaystyle\prod_{k=1}^{i-1} S_k \right) f_i & i > 1 \end{cases} \tag{8}
$$

The probability that an age i fish from the cohort of interest is caught in the observed component of the longline fishery is:

$$
\pi_i^{LL,obs} = \begin{cases} \delta_i \dfrac{F_i^{LL}}{F_i^{Tot}} f_i & i = 1 \\[4mm] \delta_i \dfrac{F_i^{LL}}{F_i^{Tot}} \left(\displaystyle\prod_{k=1}^{i-1} S_k \right) f_i & i > 1 \end{cases} \tag{9}
$$

The probability that an age i fish from the cohort of interest is caught in the unobserved component of the longline fishery is:

$$
\pi_i^{LL,unobs} = \begin{cases} (1-\delta_i) \dfrac{F_i^{LL}}{F_i^{Tot}} f_i & i = 1 \\[4mm] (1-\delta_i) \dfrac{F_i^{LL}}{F_i^{Tot}} \left(\displaystyle\prod_{k=1}^{i-1} S_k \right) f_i & i > 1 \end{cases} \tag{10}
$$

If we assume the numbers of fish caught at each age are known accurately (and that each fish has an equal probability of being caught), then the catch-at-age data, including those fish from the cohort not caught, are random multinomial, where each fish has a probability of being captured at age i in one of the fishery components (given by the expressions in (8), (9) and (10)) or else not captured. Usually, however, the catch-at-age data are not known accurately. In the case of SBT, the age distribution of the catch is determined by taking a sample, estimating the ages of fish in the sample (either from lengths or from direct aging of hard parts), and using the estimated age frequencies of the sample to represent the total catch. We have chosen to model the error in the catch-at-age data that results from this sampling procedure as Gaussian, with a coefficient of variation (CV) that depends on the level of sampling. The CV is intended to capture variability in the catch-at-age data due to non-homogeneous spatial and temporal distribution of fish, as well as different size/age selectivities among vessels (i.e. if these factors are significant, then the CV of the catch-at-age data would be large because the age distributions derived from different samples could vary a lot).

To fit a model with both multinomial "process" error and Gaussian sampling error would require a relatively sophisticated approach, such as a Kalman filter. However, in most fishery situations, the number of fish in the cohort from which catches are being taken will be very large such that the multinomial error will be negligible compared to the Gaussian sampling error, and only the latter source of error needs to be considered. This is the approach taken in the current paper.

For the surface fishery, we assume that catches are routinely sampled and that there is an appropriate sampling design and estimation model that allows for the variance in the catch data to be well estimated. We have assumed that the CV of the catch data for each year (υ_i^S) is known and independent of the tag data. For the longline fishery, we assume that all fish caught in the observed component are sampled, but that no fish from the unobserved component are sampled. Thus, there is no age information for the unobserved catches, and only catch-at-age data from the observed component is included in the model. The CV for the longline catch data in a given year (υ_i^{LL}) will be determined by the level of observer coverage (since this determines the level of sampling).

The likelihood for the surface and observed longline catch data is:

$$L_C = \frac{1}{2\pi}\prod_i\left\{\frac{1}{\upsilon_i^S P_1\pi_i^S}\exp\left(-\frac{1}{2}\left(\frac{C_i^S - P_1\pi_i^S}{\upsilon_i^S P_1\pi_i^S}\right)^2\right)\right\}\times$$

$$\prod_i\left\{\frac{1}{\upsilon_i^{LL,obs} P_1\pi_i^{LL,obs}}\exp\left(-\frac{1}{2}\left(\frac{C_i^{LL,obs} - P_1\pi_i^{LL,obs}}{\upsilon_i^{LL,obs} P_1\pi_i^{LL,obs}}\right)^2\right)\right\}$$

$$(11)$$

The overall likelihood for the combined recapture and catch data can be obtained by multiplying likelihoods for the tag-recapture data, the tag-seeding data and the catch data together:

$$L_{Tot} = L_R \times L_{\lambda^S} \times L_C \tag{12}$$

The inclusion of the catch component in the overall likelihood allows for the initial cohort size P_1 to be estimated and also provides more information on the mortality rate estimates. Additionally, for our specific model which allows for non-mixing of tagged fish in the first year after release, the tag-recapture data does not provide an estimate of the fishing mortality at age 1 of untagged fish (F_1^S and F_1^{LL}); inclusion of the catch data provides these estimates. Thus, by maximizing (12), estimates of all parameters given in Table 2 can be estimated, with the exception that we can only estimate $n-1$ natural mortality rate parameters, where n is the number of consecutive release years.

In the current model formulation, there is not enough information to estimate the proportion of observer coverage in each year (the δ's). To do so, we would need to know the total observer catch in each year as well as the total overall catch in each year. Currently, the model only requires catch data from a single cohort (i.e. from a single age class in each year). Rather than bringing the total catch data into the model, we assume that the total catch numbers are known well enough that the δ's are estimated accurately, and we treat the δ's as being known without error in our model.

The model allows for the catch CV in each fishery to vary with year, and we assume that these CV's are known (there is not enough information with which to estimate them). If we were to assume a constant CV in all years for a given fishery, then, in theory, the CV should

be estimable from the likelihood; however, we found in practice that its estimation is very poor (often converging to zero).

Overdispersion in tag return data

The tag-recapture component of the model presented above assumes a multinomial distribution for the tag returns; this is only valid if all fish of a particular age have the same probability of being caught. If there is unsystematic incomplete mixing of tagged and untagged fish[2] (e.g. if fish tagged in the same school and/or in close proximity on the same day have positively correlated recapture probabilities), then the numbers of returns at age will have more variability than a multinomial distribution would predict. Differential age/size selectivities among fishing vessels will also contribute to overdispersion if tagged fish are not homogeneously mixed within the untagged population. One way of incorporating this overdispersion is to model the tag return data as Dirichlet-multinomial. Essentially, the probabilities of return corresponding to releases at age a are modelled as Dirichlet random variables with variance parameter ω_a (see Appendix A). Then the numbers of returns conditional on the probabilities of return follow a multinomial distribution, and the unconditional numbers of returns follow the compound distribution referred to as the Dirichlet-multinomial (see Appendix A).

The likelihood for the tag return data when these data are modelled as Dirichlet-multinomial is:

$$
L_R^{\omega} = \gamma^{\omega} \times \prod_a \left\{ \frac{\Gamma(\omega_a)}{\Gamma(N_a + \omega_a)} \prod_i \frac{\Gamma\left(\left(N_a - R_{a,\bullet}^{Tot}\right) + \omega_a\left(1 - p_{a,\bullet}^{Tot}\right)\right)}{\Gamma\left(\omega_a\left(1 - p_{a,\bullet}^{Tot}\right)\right)} \times \right.
$$
$$
\left. \prod_i \frac{\Gamma\left(R_{a,i}^S + \omega_a p_{a,i}^S\right) \Gamma\left(R_{a,i}^{LL,obs} + \omega_a p_{a,i}^{LL,obs}\right) \Gamma\left(R_{a,i}^{LL,unobs} + \omega_a p_{a,i}^{LL,unobs}\right)}{\Gamma\left(\omega_a p_{a,i}^S\right) \Gamma\left(\omega_a p_{a,i}^{LL,obs}\right) \Gamma\left(\omega_a p_{a,i}^{LL,unobs}\right)} \right\} \tag{13}
$$

where

$$
\gamma^{\omega} = \prod_a \frac{N_a!}{\left(N_a - R_{a,\bullet}^{Tot}\right)! \prod_i \left(R_{a,i}^S! R_{a,i}^{LL,obs}! R_{a,i}^{LL,unobs}!\right)} \,,
$$

and, as in equation (6),

$$
R_{a,\bullet}^{Tot} = \sum_i \left(R_{a,i}^S + R_{a,i}^{LL,obs} + R_{a,i}^{LL,unobs}\right),
$$

and

$$
p_{a,\bullet}^{Tot} = \sum_i \left(p_{a,i}^S + p_{a,i}^{LL,obs} + p_{a,i}^{LL,unobs}\right).
$$

---

[2] Unsystematic incomplete mixing is meant to refer to situations where there is still large amounts of mixing among tagged and untagged fish and the pattern of mixing has a large "random" component such that on average the probability of recapture of tagged and untagged fish are the same. This should be distinguished from the situation where there is a systematic and repeatable pattern of non-mixing between tagged and untagged fish -- for example, if all tagging was done late in the season in one location and fish in that location and time period only remain in one part of the stock's overall range. Such systematic non-mixing will induce biases into the population and mortality estimates if it is not accounted for within the estimation model. A basic assumption of the estimation model used here is that the tagged fish constitute a representative sample of the population.

Note that γ^ω is a constant that can be left out of the likelihood.

The overall likelihood is now analagous to (12) except L_R^ω replaces L_R:

$$L_{Tot}^\omega = L_R^\omega \times L_{\lambda S} \times L_C \qquad (14)$$

The parameters that we estimate by maximizing the likelihood in (14) are the same as before, except now we also have overdispersion parameters (ω_a's) to estimate. Rather than estimating an overdispersion parameter for each release event (which would not likely be possible with the current model formulation), we constrain the ω_a's so that they lead to an increase in the variance of the returns at age of x times over that of multinomial returns. This can be accomplished by setting $\omega_a = (N_a - x)/(x - 1)$ (refer to Appendix A). Now, instead of having several additional overdispersion parameters to estimate, we have just one, x.

Data and parameters used to condition the simulations

In generating data for the simulations, our aim was to choose input values that emulate the most recent years of SBT tag-recapture and catch data as closely as possible. SBT are generally tagged at ages 1 to 3, therefore in our simulations we assume that we tag a single cohort of fish in 3 consecutive years at ages 1, 2 and 3. Most SBT tags are returned within the first 5 years after release, so we generate 5 years of recapture data, along with 5 years of corresponding catch data.

The input values used to generate the tag-recapture and catch-at age data sets for our simulations are given in Table 3. The number of releases were determined by averaging the number of tags released at ages 1 to 3 as part of the CCSBT tagging program in years 2002, 2003 and 2004. We also looked at the effect of halving and doubling the number of releases.

The mortality rates were assumed to follow a negative linear trend with age; the slope and intercept were chosen to give values that closely resemble the mortality rate vector commonly used in past stock assessments. The reason for assuming a linear trend is that with only 3 release years, we can only estimate 2 mortality rate parameters. By constraining the mortality rates to be linear with age, we reduced the number of mortality rate parameters to 2 as required. Other constraints could have been imposed but a linear trend is consistent with previous assumptions about natural mortality rates for SBT.

The total fishing mortality rates (across fisheries) were based on total SBT catches from years 1998 to 2000. The average total catches in numbers of ages 1 to 5 fish over these 3 years were calculated to be 1959, 58208, 225015, 69982 and 26817 respectively. Thus, using the mortality rates discussed above and assuming an initial (age 1) population size of 2 million fish, we could calculate the age-specific fishing mortality rates required to give these catch numbers using equations (1) and (2). These are the F_i^{Tot} values reported in Table 3. We chose 2 million for P_1 because it is within the plausible range of values for SBT based on recent stock assessments (e.g. Hirmatsu and Tsuji 2001; Kolody and Polacheck 2001; Polacheck and Preece 2001), and it also resulted in reasonable fishing mortality rates. However, varying P_1 over the range of 1 to 4 million had a negligible effect on the results.

We expect the total fishing mortality rates for tagged fish in their first year after release (i.e. the F_i^{*Tot}'s) to be quite low for SBT because tagging generally occurs near the end of the surface fishery season. The values we chose were rather arbitrary (0.05 for all ages), but they do not have much influence on the results, with an exception being if one of the values is so close to zero that the simulated tagging data has no tag returns at that age. In such a case, not all parameters are estimable unless constraints/assumptions are imposed regarding the mortality rates and reporting rate for that age. Zero tag returns, especially at ages 1 and 2, may be an issue with real SBT tag-recapture data and, if so, would have to be dealt with appropriately.

To apportion the total fishing mortality between the surface fishery and longline fishery, we need to know the proportion of surface versus longline catches at each age. Using SBT catch data for years 1998 to 2000, we calculated the proportion of surface catches at each age in each year and then took the average of the 3 years; these values are reported as θ_i in Table 3. Then the fishery-specific F_i's and F_i^*'s were calculated by simply multiplying the total fishing mortalities by θ_i for the surface fishery and $(1 - \theta_i)$ for the longline fishery.

The reporting rates used for the surface fishery were based on a preliminary analysis of data from a pilot tag seeding experiment conducted on SBT farm cages in 2002/2003 (Polacheck and Stanley 2004). This analysis suggested an average reporting rate of 0.65 with a standard error of 0.10. These values were assumed to apply in all years in the results presented here (i.e. for all ages in our single-cohort formulation). The reporting rates used for the unobserved component of longline fishery (0.10 for all ages) were based on longline reporting rate estimates from previous analyses of the 1990s SBT tagging data (which ranged between 0 and around 0.40 depending on the fleet) combined with concerns that promotional activities (particularly direct personal contact) encouraging fisherman to return tags has been less during the SRP than during the 1990s. However, the effect of increasing the reporting rates for the longline fishery was explored.

The CV for the catch-at-age data from the surface fishery was chosen to be 0.2 in all years. This figure is rather arbitrary but currently there are no estimates, or developed statistical models for obtaining estimates, of the error in the age composition of the surface catches. In addition, the actual CV is likely to vary among years. The CV of the catch-at-age data for the observer component of the longline fishery is assumed to be related to the level of observer coverage, because more observers means more catch sampling. A hypothetical relationship between the level of observer coverage and the CV of the catch data, which we believe to be reasonable for our purposes, is shown in Figure 1. The formula used to generate this curve is $\upsilon_i^{LL} = 0.75 * (0.05)^{\sqrt{\delta_i}}$. Note that even with 100% observer coverage, the CV does not go to zero because there is still variability in the catch process (referred to previously as multinomial process error) and aging error in going from measured length distributions to estimated age distributions. This relationship is rather arbitrary; however, sufficient data and information are not available on the actual sampling protocols to develop a more realistic model.

We kept the level of observer coverage in the longline fishery the same in all years (i.e. at all ages) and, initially, set the level of be 0.1. This value was chosen because 10% observer coverage has been the goal set by CCSBT members in past years (Anon. 2001b). The CV of the observer catches that corresponds to this observer level is 0.29 (calculated using the

relationship given in the previous paragraph). One of the primary goals of this paper is to investigate how the level of observer coverage affects our ability to estimate mortality rate and abundance parameters; thus, we also considered observer levels of 0.05, 0.2, 0.3 and 0.5, with corresponding catch CV's of 0.38, 0.20, 0.15 and 0.09 respectively.

In the model that incorporates overdispersion in the tag return data, we also need to specify how much extra variability we want in the tag returns compared to that of multinomial returns. We chose a factor of 3 (i.e., in the notation used in the Methods section, $x = 3$). Note that overdispersion in the tag return data is likely to be associated with higher variability in the longline catch data, especially at low levels of observer coverage (e.g. a large source of the variability in the catch data would come from large inter-vessel variability in the size/age composition of their catches, especially if catches from only a few vessels were sampled; this would be the case when observer coverage is low because observers would likely be constrained to a limited number of relatively long cruises). However, in the absence of information on this, the same CV/observer coverage relationship for the longline catch data was used in both the model with and without overdispersion in the tag returns.

Table 3. Parameter values for reference case simulation run.

	Age/year, i				
	1	2	3	4	5
P_i	2×10^6	–	–	–	–
N_i	2718	5807	1223	–	–
M_i	0.4	0.35	0.3	0.25	0.2
F_i^{*Tot}	0.05	0.05	0.05	–	–
F_i^{Tot}	0.001	0.053	0.340	0.183	0.103
θ_i	0.882	0.825	0.828	0.407	0.120
F_i^{*S}	0.044	0.041	0.041	–	–
F_i^{*LL}	0.006	0.009	0.009	–	–
F_i^{S}	0.001	0.044	0.282	0.075	0.012
F_i^{LL}	0.000	0.009	0.058	0.108	0.091
λ_i^{S}	0.65	0.65	0.65	0.65	0.65
s_i	0.10	0.10	0.10	0.10	0.10
υ_i^{S}	0.20	0.20	0.20	0.20	0.20
λ_i^{LL}	0.10	0.10	0.10	0.10	0.10
δ_i	0.10	0.10	0.10	0.10	0.10
υ_i^{LL}	0.29	0.29	0.29	0.29	0.29

Results

Multinomial tag returns

Using the values in Table 3, we simulated 100 multinomial tag-recapture and Gaussian catch-at-age datasets. We then obtained parameter estimates corresponding to each of the 100

datasets by maximizing the likelihood in (12). We refer to the simulations carried out using the values in Table 3 as the 'reference case' simulations.

Effect of observer coverage

In addition to the reference case simulations, we also ran 100 simulations using each of the alternative levels of observer coverage being considered, namely, $\delta = 0.05, 0.2, 0.3$ and 0.5 (with corresponding longline catch CV's of $\upsilon^{LL} = 0.38, 0.20, 0.15$ and 0.09). The means and standard deviations of the 100 maximum likelihood estimates for the parameters of key interest are given in Tables 4 and 5. Results for the remaining parameters can be found in Appendix B, Tables B1 and B2. The mean estimates of almost all parameters are within two standard errors of the true value, suggesting they are estimated without bias (standard error equals standard deviation divided by square root of sample size, where the sample size is 100 in our case). An exception is the estimate of the age 1 population size, P_1, which has a slight negative bias; however, the bias is small (<7%) and disappears as the level of observer coverage increases. There are also significant biases (statistically speaking) in some of the reporting rate estimates, but these biases are small in practical terms and they diminish as the level of observer coverage increases. The reporting rate estimates are not of primary interest, and since small biases in these estimates do not appear to induce biases in the mortality rate estimates, they are not of concern. Although insignificant, there is some suggestion of a small negative bias for both M_1 and M_5 (recall that natural mortality is constrained to be linear with age so it can be fully described by 2 parameters; we have chosen to parameterize the line in terms of M_1 and M_5). Interestingly, the bias for M_1 decreases as observer coverage increases, but the bias for M_5 increases.

Our ability to estimate almost all of the parameters improves as the level of observer coverage in the longline fishery increases (as seen by a decrease in standard deviation as observer coverage increases; Table 5). The degree of improvement differs between parameters and can be better evaluated by looking at the coefficient of variation (CV = standard deviation/mean) of the estimates as opposed to the standard deviation (Figure 1). As we would expect, the CV's of the fishing mortality rate estimates for the longline fishery are most improved by increases in longline observer coverage, with improvements in CV ranging from 8% to 27% when observer coverage goes from 5% to 50%. We note that the CV of M_5 is large in all situations (~90%), whereas the CV of the initial cohort size P_1 is always small (~10%); we discuss these findings in the Discussion.

Table 4. Mean of key reference case parameter estimates (from 100 simulations) for various levels of observer coverage (δ). True parameter values are given below parameter names. The values for P_1 are expressed in millions.

	P_1	M_1	M_5	F_1^S	F_2^S	F_3^S	F_4^S	F_5^S	F_1^{LL}	F_2^{LL}	F_3^{LL}	F_4^{LL}	F_5^{LL}
δ	2.0	0.4	0.2	0.001	0.044	0.282	0.074	0.012	0.000	0.009	0.058	0.109	0.091
0.05	1.86	0.390	0.197	0.001	0.044	0.293	0.081	0.014	0.000	0.009	0.062	0.109	0.100
0.10	1.88	0.392	0.198	0.001	0.044	0.293	0.082	0.014	0.000	0.009	0.061	0.111	0.100
0.20	1.92	0.392	0.189	0.001	0.043	0.287	0.079	0.013	0.000	0.009	0.059	0.111	0.097
0.30	1.94	0.393	0.180	0.001	0.044	0.285	0.076	0.013	0.000	0.009	0.059	0.108	0.094
0.50	1.97	0.396	0.178	0.001	0.044	0.283	0.076	0.013	0.000	0.009	0.059	0.109	0.093

Table 5. Standard deviation of key reference case parameter estimates (from 100 simulations) for various levels of observer coverage (δ). The values for P_1 are expressed in millions.

δ	P_1	M_1	M_5	F_1^S	F_2^S	F_3^S	F_4^S	F_5^S	F_1^{LL}	F_2^{LL}	F_3^{LL}	F_4^{LL}	F_5^{LL}
0.05	0.23	0.057	0.176	0.000	0.007	0.050	0.023	0.006	0.000	0.003	0.016	0.032	0.051
0.10	0.22	0.056	0.177	0.000	0.007	0.046	0.023	0.007	0.000	0.003	0.013	0.030	0.042
0.20	0.20	0.056	0.174	0.000	0.007	0.045	0.020	0.007	0.000	0.002	0.010	0.027	0.043
0.30	0.18	0.052	0.165	0.000	0.008	0.044	0.019	0.006	0.000	0.001	0.008	0.023	0.037
0.50	0.16	0.054	0.168	0.000	0.008	0.040	0.018	0.005	0.000	0.001	0.006	0.023	0.034

Effect of number of releases

While the level of observer coverage is one factor of the experimental design that can be controlled, the number of releases is another. We repeated the reference case simulations except we first halved, and then doubled, the number of releases at each age. Again, the mean parameter estimates were unbiased for the most part, and any biases were small and not of concern; as such, we do not present the mean estimates. The CV's are of more interest (Figure 2). The general direction of the results is as expected — halving the number of releases degrades the estimates and doubling the number of releases improves the estimates, at least for parameters of interest (those shown in Figure 2). However, the response appears to be asymmetric; the loss in precision from halving the number of releases appears to be greater than the gain in precision from doubling the number of releases. It is also worth noting that changing the number of releases had a larger effect on the precision of the natural mortality rate estimates, M_1 and M_5, than changing the level of observer coverage.

These results were obtained using a 10% level of observer coverage since this is the reference level; however, the general relative effect of halving and doubling the number of releases on the precision of the parameter estimates remained the same at other levels of observer coverage.

Effect of other factors

We increased the longline reporting rate from the reference case value of 0.1 to 0.5 in all years, then reran the simulations. There was almost no improvement in the parameter estimates (Figure 4). This is expected because in the likelihood, the tag returns from the unobserved component of the longline fishery are scaled up by the estimated reporting rate to give an estimate of the actual number of tag recaptures. The reporting rates are determined by the return rate in the observer component, so that the age distribution of the returns always ends up the same for the unobserved component as the observed component. As such, it does not matter whether the reporting rate is 0.1 or 0.5; it is the accuracy of the observer tag return data that matters (as we saw in our previous simulations).

Preliminary analyses of data from recent tag seeding experiments suggested a value of 65% for the surface fishery reporting rates, so we used this value in our reference case simulations. However, in previous analyses of SBT tagging data , the reporting rate in the surface fishery has generally been assumed to be 100% (Polacheck et al. 1996, 1998). Thus, we ran simulations using 100% surface reporting rates (and assumed they were known without error) and found only a minimal improvement in the fishing mortality rate estimates for the surface fishery, and no improvement in the natural mortality rate and abundance estimates (Figure 5a). Furthermore, in the case of 65% reporting rates, we looked at the effect of changing the precision with which these rates are estimated; in particular we increased that standard error

of the estimates from 0.10 to 0.30. This made virtually no difference to the results (Figure 5b). These results suggest that the return data from the surface fishery are already sufficiently informative that neither an increase in the magnitude of the reporting rates, nor an increase in the precision of the reporting rate estimates, has much effect.

Lastly, we considered the effect of changing the CV of the catch-at-age data in the surface fishery from 0.2 in all years to 0.05, and also 0.30, in all years. The estimates of fishing mortality at ages 1 and 2 in the surface fishery were most affected, with the CV of the age 1 estimates decreasing by over 20% when the catch CV was improved from 0.30 to 0.05 (Figure 6).

Dirichlet-multinomial tag returns

We repeated all of the simulations done in the case of multinomial tag returns using the model with Dirichlet-multinomial tag returns.

We first present the results from the simulations looking at the effect of changing the level of observer coverage. The means and standard deviations of the estimates for the parameters of key interest are summarized in Tables 6 and 7; those for the remaining parameters can be found in Tables B3 and B4 of Appendix B. The CV's of the key parameter estimates are shown in Figure 7. Comparing these results with the analogous results for the case of multinomial tag returns (i.e., Tables 4, 5, B1 and B2, and Figure 2), we see that:

- Again, the mean estimates are all within one standard deviation of the true value, with the exception of the increased-variance factor, x, for the Dirichlet distribution, which is consistently underestimated (see Tables B1 and B2 of Appendix B).
- The slight biases seen in the abundance and natural mortality rate estimates in the case of multinomial returns no longer appear to exist.
- The standard deviations (and hence CV's) of the estimates are larger for all parameters (and significantly so for some parameters, in particular for the fishing mortalities at older ages in both fisheries).
- Again, the standard deviations (and hence CV's) of almost all parameter estimates decline as the level of observer coverage increases, and for a given parameter, the amount that the CV declines is roughly the same. For example, the declines in the CV's are still largest for the fishing mortality rates in the longline fishery and they are in the range of 10 to 30% when the observer level increases from 5% to 50%.

Qualitatively, the results from varying any of the factors were very similar in the model with Dirichlet-multinomial returns as the model with multinomial tag returns. The parameters were almost always estimated with less precision (i.e. their CV's were larger) with Dirichlet-multinomial returns, but the relative changes in CV's and general observations made did not change significantly between the models.

Table 6. Mean of key reference case parameter estimates (from 100 simulations) for various levels of observer coverage (δ) when overdispersion is incorporated in tag return data. True parameter values are given below parameter names. The values for P_1 are expressed in millions.

	P_1	M_1	M_5	F_1^S	F_2^S	F_3^S	F_4^S	F_5^S	F_1^{LL}	F_2^{LL}	F_3^{LL}	F_4^{LL}	F_5^{LL}
δ	2.0	0.4	0.2	0.001	0.044	0.282	0.074	0.012	0.000	0.009	0.058	0.109	0.091
0.05	1.94	0.398	0.193	0.001	0.045	0.297	0.084	0.015	0.000	0.009	0.059	0.111	0.106
0.10	1.95	0.397	0.198	0.001	0.044	0.288	0.084	0.015	0.000	0.009	0.059	0.111	0.104
0.20	2.00	0.400	0.219	0.001	0.044	0.299	0.084	0.016	0.000	0.009	0.060	0.117	0.116
0.30	2.02	0.411	0.205	0.001	0.044	0.298	0.082	0.015	0.000	0.009	0.060	0.114	0.114
0.50	1.99	0.399	0.190	0.001	0.045	0.289	0.078	0.013	0.000	0.009	0.058	0.112	0.103

Table 7. Standard deviation of key reference case parameter estimates (from 100 simulations) for various levels of observer coverage (δ) when overdispersion is incorporated in tag return data. The values for P_1 are expressed in millions.

δ	P_1	M_1	M_5	F_1^S	F_2^S	F_3^S	F_4^S	F_5^S	F_1^{LL}	F_2^{LL}	F_3^{LL}	F_4^{LL}	F_5^{LL}
0.05	0.30	0.097	0.211	0.000	0.010	0.066	0.033	0.010	0.000	0.004	0.023	0.050	0.074
0.10	0.30	0.104	0.215	0.000	0.009	0.065	0.032	0.011	0.000	0.003	0.018	0.039	0.065
0.20	0.28	0.087	0.223	0.000	0.009	0.063	0.029	0.010	0.000	0.002	0.015	0.043	0.075
0.30	0.24	0.086	0.218	0.000	0.008	0.060	0.027	0.009	0.000	0.002	0.012	0.038	0.065
0.50	0.25	0.086	0.204	0.000	0.009	0.059	0.024	0.008	0.000	0.001	0.010	0.035	0.056

Discussion

The estimation framework and simulation results presented in this paper provide insights into design issues for the tagging program currently being conducted as part of the CCSBT SRP, in particular into appropriate levels of observer coverage and tag releases. Observer coverage to date has generally been minimal (<5%) (Anon. 2003). The results suggest that increasing observer coverage can lead to significant improvements in the precision of the fishing mortality rate estimates for the longline fishery, as well as smaller improvements in the estimate of population abundance. The number of tags that have been released in recent years as part of the CCSBT SRP tagging program appear to be adequate. Doubling the number of releases at each age led to only marginal improvements in any of the parameter estimates. On the contrary, halving the number of releases noticeably degraded some of the parameter estimates; thus, we would caution against reducing the number of releases without further investigation.

An advantage to having a multi-component fishery is that, if the catches by component are known well, then reporting rates only need to be estimated well in one component in order to get reasonable estimates of reporting rates (and hence other parameters) in other components (Hearn et al. 2003). This relies on the assumption that recapture rates of tagged fish (i.e. number of tags per unit of catch) are the same in all components (i.e. complete mixing). Then, knowing the reporting rate in one component means the recapture rate is known in that component, so that the number of tags that should have been returned in another component to achieve the same recapture rate can be calculated. We see evidence of this in our simulations because the reporting rates in the surface fishery are estimated very well so that even when the level of observer coverage is only 5% in the longline fishery, the reporting rates (and hence fishing mortality rates) in the longline fishery can still be estimated reasonably well. In essence, the observer data are not contributing substantially to the

estimation of reporting rates; instead, the reporting rates for the longline fishery are being derived in most part from extrapolation from the surface fishery return rates. On the contrary, if we were to assume that there is no information about reporting rates in the surface fishery (e.g. no tag seeding data), then our ability to estimate the reporting rates in the longline fishery, as determined by the level of observer coverage, would have a larger influence on the reporting rate and fishing mortality rate estimates in the surface fishery (see Figure 8). The degree of influence will be greater when the surface catches are known with high precision (Figure 8a) versus when they are known with less precision (Figure 8b). It is interesting to note that when there is no information on reporting rates in the surface fishery, the level of observer coverage becomes more influential not only on the estimation of the surface fishery parameters, but also on the estimation of the initial population size parameter.

There are several disadvantages of basing reporting rates for the longline fishery on extrapolation from the surface fishery instead of obtaining independent estimates from observer data. Firstly, extrapolation from the surface fishery precludes the ability to test for significant non-mixing. Low return rates of tags in one fishery component could be the result of either low reporting rates or the fact that tagged fish did not mix with the portion of the stock being fished by this fishery component. These two possibilities are unresolvable without direct information on the reporting rates in the different fishery components. This issue is particularly of concern for SBT longline fisheries given the large spatial/temporal scales on which these fisheries operate and the spatially-restricted nature of the current tagging operations. For example, if low tag return rates are found for longline vessels fishing off South Africa, this could be due to low reporting rates or the fact that low numbers of tagged fish actually mixed with fish off South Africa. The implications of these two alternatives could be large in terms of estimates of mortality rates and population size; simply assuming complete mixing when it does not exist will bias these estimates. Furthermore, if non-mixing exists, then the extrapolated reporting rates will be biased, which will compound the biases already introduced into the mortality rate and population size estimates due to non-mixing. Moreover, the use of extrapolated reporting rates prevents the application of more spatially-explicit tag recovery models to account for heterogeneity in recapture probabilities as a result of non-mixing.

The model with multinomial tag returns assumes complete mixing of tagged and untagged fish, and that the fate of each tagged fish is independent of the fate of other tagged fish. The first of these assumptions may be violated in the case of SBT because their distribution is often patchy and juvenile fish tend to form schools. The second assumption is also likely to be violated for SBT because tagging generally occurs over a limited geographic area and a limited time period, and multiple fish from the same school are often tagged. If fish tagged from the same school or within close time and proximity of each other have a tendency to behave similarly, then their recapture probabilities would be positively correlated. Either non-mixing or dependence between tagged fish would mean that the return data are overdispersed. We attempted to incorporate overdispersion into our model by modelling the tag return data as Dirichlet-multinomial, which allows for extra variability compared to that of a multinomial distribution. We parameterized the Dirichlet-multinomial distribution so that the amount of extra variability was a constant factor, regardless of the number of releases. It may be argued that if the overdispersion stems mainly from non-independence among tagged fish, then tagging more fish will reduce this source of variance (assuming more releases would mean fish from a larger number of schools and a larger geographical and temporal range would be tagged). In such a case, the overdispersion should be modelled as a function of the number of releases. Determining the sources of overdispersion, their relative

magnitudes, and the most appropriate way to model them is an issue requiring further investigation.

In modelling the catch-at-age data as Gaussian, we argued that the multinomial process error should be negligible compared to the sampling error (assumed to Gaussian), so that only the latter source of error needed to be considered. However, if fish are not distributed homogeneously in space or time, or if there is large variability in the size/age selectivities of vessels, then the process error would be overdispersed relative to a multinomial distribution. Furthermore, the sampling error would be larger (i.e. the CV of the Gaussian distribution would be larger). In this case, it is not clear if the process error would still be negligible compared to the sampling error, nor is it clear how the relationship between the CV of the sampling error and the level of sampling (i.e. the level of observer coverage for the longline fishery) should be modelled. More observer data should still mean a reduction in the sampling error (i.e. a smaller CV), but the amount of reduction will depend on the nature of the increased observer coverage. If all of the additional observer data comes from only a few vessels/cruises, then the gain will be much less than if it comes from a large number of vessels/cruises operating over a wide geographic range. Developing an appropriate error model for the catch-at-age data is an important area for future work because it is critical for understanding the statistical properties of the parameter estimates obtained from the tagging and catch model.

In all of our results, the natural mortality rate at age 1, M_1, is estimated with reasonable precision, even in the model with overdispersion (CV around 20-25%). On the other hand, the natural mortality rate at age 5, M_5, is estimated with very low precision (CV over 100% in the case of overdispersion), and a histogram of the estimates for any set of simulations shows that the estimate of M_5 usually equals either the lower bound (0.01) or upper bound (0.4) set for this parameter. While this causes some concern, it is important to recall that natural mortality has been constrained to be a linear function of age, so that the natural mortality rate estimates for ages 2 to 4 will have CV's intermediate to those at ages 1 and 5. For example, we calculated the natural mortality estimates at all ages for the reference case simulations and found their CV's to be 0.14, 0.14, 0.28, 0.52 and 0.89 for ages 1 to 5 respectively.

The initial population size (P_1) was estimated well in all cases (CV less than 20%), even when many of the fishing and natural mortality rate parameters were not. At first this seems counter-intuitive. However, on further consideration, it can be explained by the presence of high positive correlations between natural mortality and fishing mortality (see discussion below). If natural mortality is overestimated for a particular set of data, then fishing mortality is also likely to be overestimated since fishing and natural mortality are positively correlated. An overestimation of natural mortality would mean the probability of catching a fish is underestimated, whereas an overestimation of fishing mortality would mean the probability of catching a fish is overestimated; thus, the two counteract each other such that the probability of catching a fish may be estimated without any bias. A similar argument holds if natural mortality was underestimated. For estimating population size, it is the estimate of the probability of catching a fish that matters, not the actual estimates of natural and fishing mortalities (since, in simplistic terms, catch equals population size times probability of catching a fish, so if we know the catch and the probability of catch well, then we know the population size well).

High correlations exist between many of the parameter estimates (see Appendix B, Table B5). As already mentioned, there are some high positive correlations between the natural mortality and fishing mortality estimates, especially at older ages. This is expected because an increase in natural mortality means that less fish are still alive in the population; thus, in order to achieve a particular level of catch, fishing mortality must increase (i.e. the percentage of the population caught must increase) as natural mortality increases. For the same reason, the fishing mortality rates between ages and fisheries are often highly positively correlated. For example, if fishing mortality at age i increases, then there are less fish of age $i+1$ alive in the population the next year; thus, in order to achieve a particular level of catch at age $i+1$, the fishing morality at age $i+1$ would have to increase if fishing mortality at age i increased. Finally, there are high negative correlations between the initial population size and the fishing mortality estimates. These can be explained in a similar fashion, because to have achieved a particular level of catch, the population size must have been larger if the fishing mortality had been low than if it had been high.

The results presented here are for a tagging experiment involving a single cohort. In practice, it would likely be feasible and cost efficient to tag two or more cohorts in any given year, and this is done in the case of SBT. This could improve the information available for estimating reporting rates since we assume that reporting rates differ only by year, and not age. Perhaps more importantly, if mortality rates are assumed to vary only with age and not year, then having data from more cohorts could potentially improve our ability to estimate natural mortality rates, which we have seen is quite poor. In order to evaluate the potential benefit of including more cohorts, we ran some simulations using data for two consecutive cohorts, both with 3 consecutive release years and 5 recapture years (i.e. cohort 1 was tagged in years 1, 2 and 3 at ages 1, 2 and 3 and recaptured in years 1 to 5; cohort 2 was tagged in years 2, 3 and 4 at ages 1, 2 and 3 and recaptured in years 2 to 6). We allowed fishing mortality rates to vary with age, year and fishery; natural mortality rates to vary with age; and reporting rates to vary with year and fishery. A small improvement was seen in the estimate of M_1 (3-4% decrease in CV) and a slightly larger improvement in the estimate of M_5 (5-10% decrease in CV). Further improvements would be expected with the inclusion of even more cohorts, and data from multiple cohorts should be available from the current SBT tagging program.

Literature Cited

Anon. 2001a. Report of the tagging program workshop. Commission for the Conservation of Southern Bluefin Tuna. 2-4 October 2001. Canberra, Australia.

Anon. 2001b. Report of the eighth annual meeting. Commission for the Conservation of Southern Bluefin Tuna. 15-19 October 2001. Miyako, Japan.

Anon. 2003. Report of the eighth meeting of the scientific committee. Commission for the Conservation of Southern Bluefin Tuna. 1-4 September 2003. Christchurch, New Zealand.

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 multicomponent fishery. N. Am. J. Fish. Manage. **23**:66-77.

Hirmatsu, K. and S. Tsuji. 2001. Stock assessment and future projection of the southern bluefin tuna based on the ADAPT VPA.CCSBT-SC/0108/31.

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.

Kolody, D. and T. Polacheck. 2001. Application of a statistical catch-at-age and -length integrated analysis model for the assessment of southern bluefin tuna stock dynamics 1951-2000. CCSBT-SC/0108/13.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1996. Estimation of natural and fishing mortality for juvenile southern bluefin tuna based on multi-year tagging of cohorts. CCSBT-SC/96/11.

Polacheck, T., W. Hearn, C. Millar and C. Stanley. 1998. Updated estimates of mortality rates for juvenile SBT from multi-year tagging cohorts. CCSBT-SC/9808/20.

Polacheck, T., J. P. Eveson, and G. M. Laslett. 2001. Exploring the trade-off between tag releases and observer coverage in the estimation of mortality rates through an integrated Brownie and Petersen mark-recapture estimation approach. CCSBT-ESC/0309/22.

Polacheck, T. and A. Preece. 2001. An integrated statistical time series assessment of the southern bluefin tuna stock based on catch at age data. CCSBT-SC/0108/19.

Polacheck, T. and C. Stanley. 2004. Update on tag seeding activities and preliminary estimates of reporting rates from the Australian surface fishery based on tag seeding experiments. CCSBT-ESC/0409/15.

Stanley, C. and T. Polacheck. 2003. Report from a pilot tag seeding program for estimating tag reporting rates from the Australian surface fishery. CCSBT-ESC/0309/25.

Acknowledgements

Figure 1. The assumed relationship between level of observer coverage and accuracy (i.e. the coefficient of variation) of the catch-at-age data for the longline fishery.
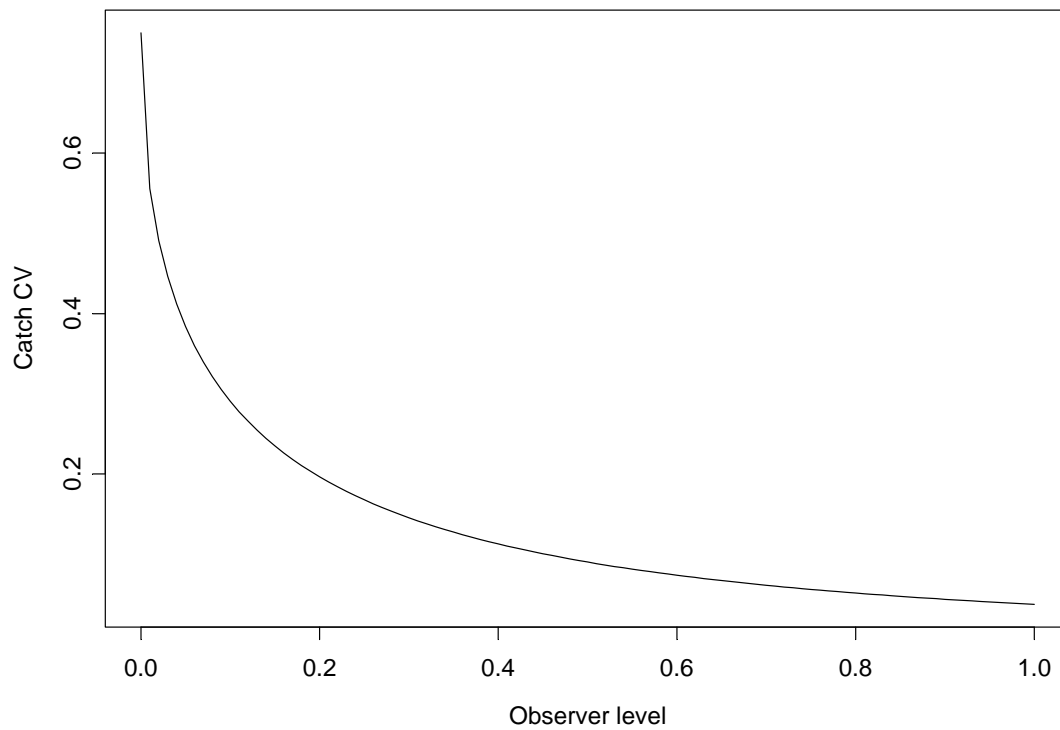
Figure 2. Effect of varying the level of observer coverage on the coefficient of variation (CV) of the key parameter estimates.
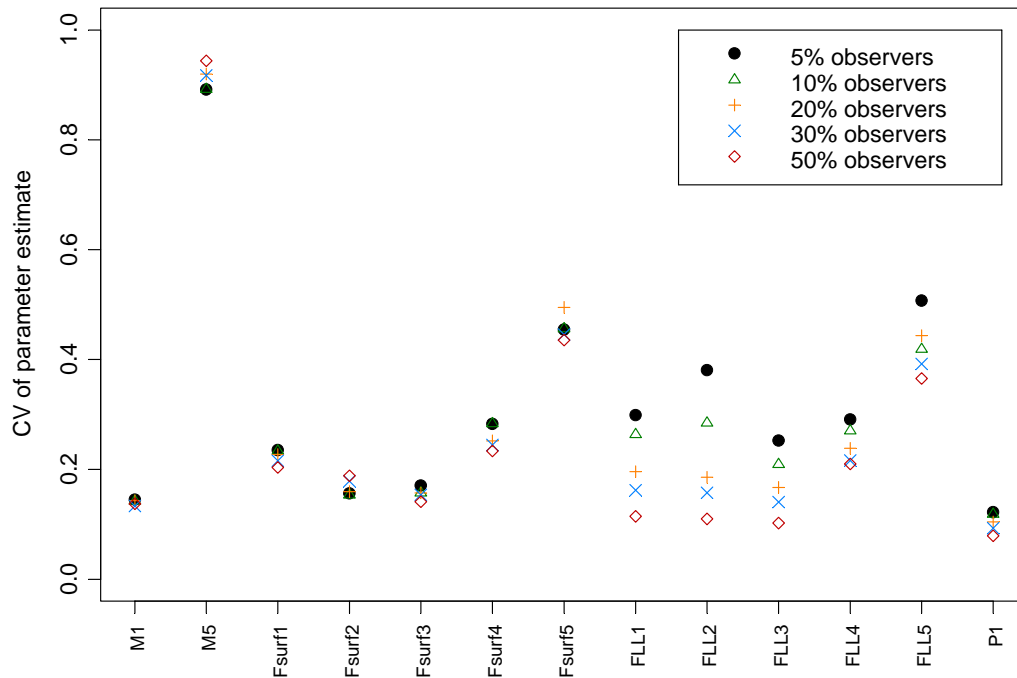
Figure 3. Effect of varying the number of releases (N) on the coefficient of variation (CV) of the key parameter estimates when the level of observer coverage is 0.10. N refers to the reference case number of releases.
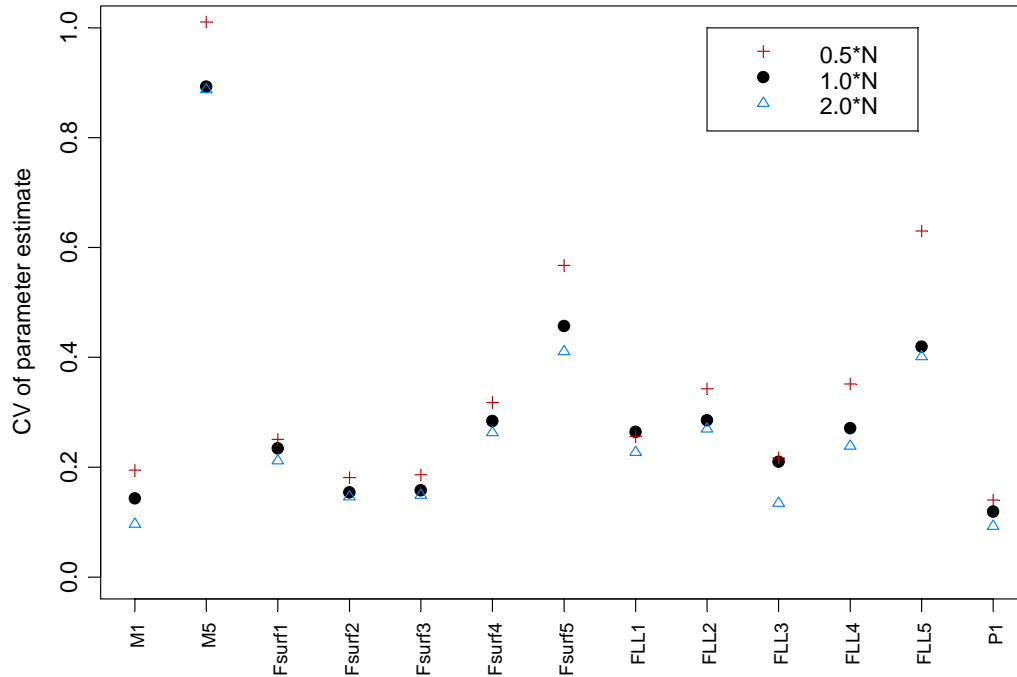
Figure 4. Effect of varying the reporting rate in the unobserved component of the longline fishery (LL rep rate) on the coefficient of variation (CV) of the key parameter estimates.
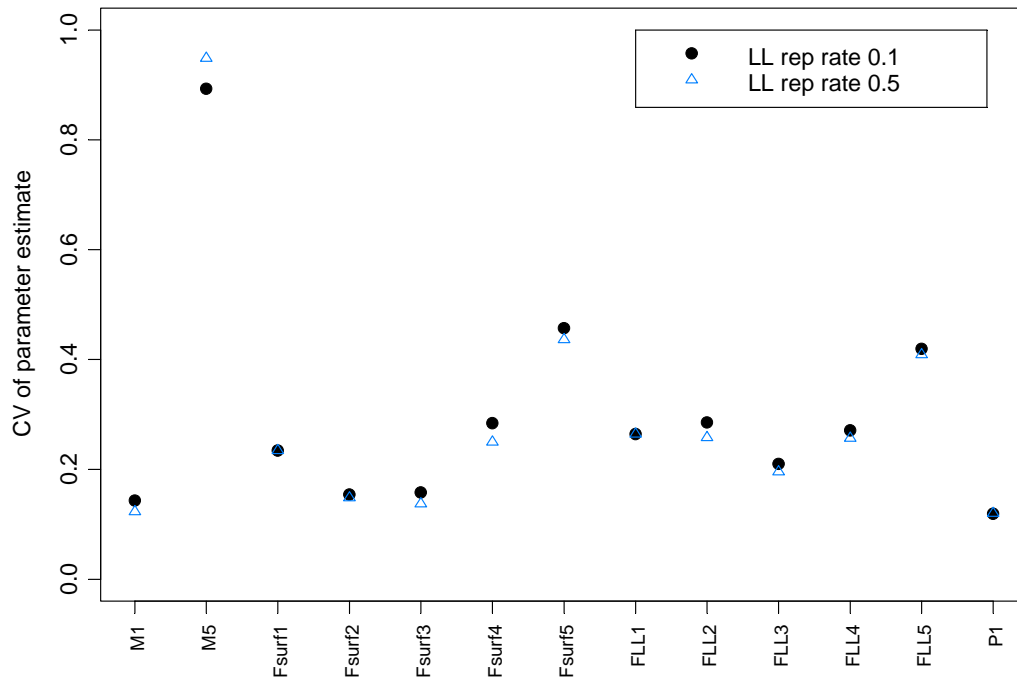
Figure 5. Effect of varying a) the magnitude of the reporting rates in the surface fishery (surf rep rate); and b) the standard error (SE) of the reporting rate estimates for the surface fishery, on the coefficient of variation (CV) of the key parameter estimates.
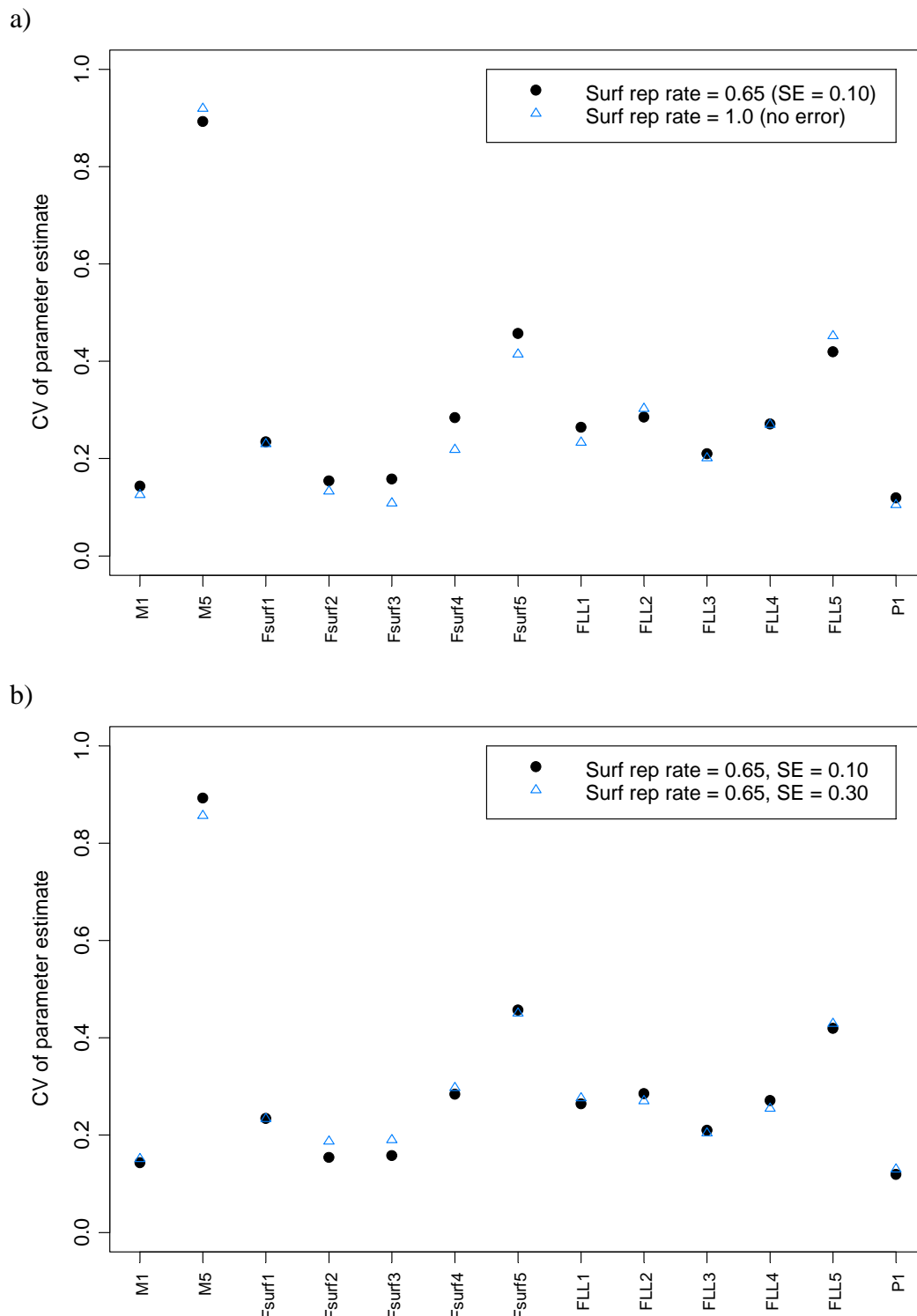
a)

b)

Figure 6. Effect of varying the coefficient of variation of the catch-at-age data in the surface fishery (CV_surf) on the coefficient of variation (CV) of the key parameter estimates.
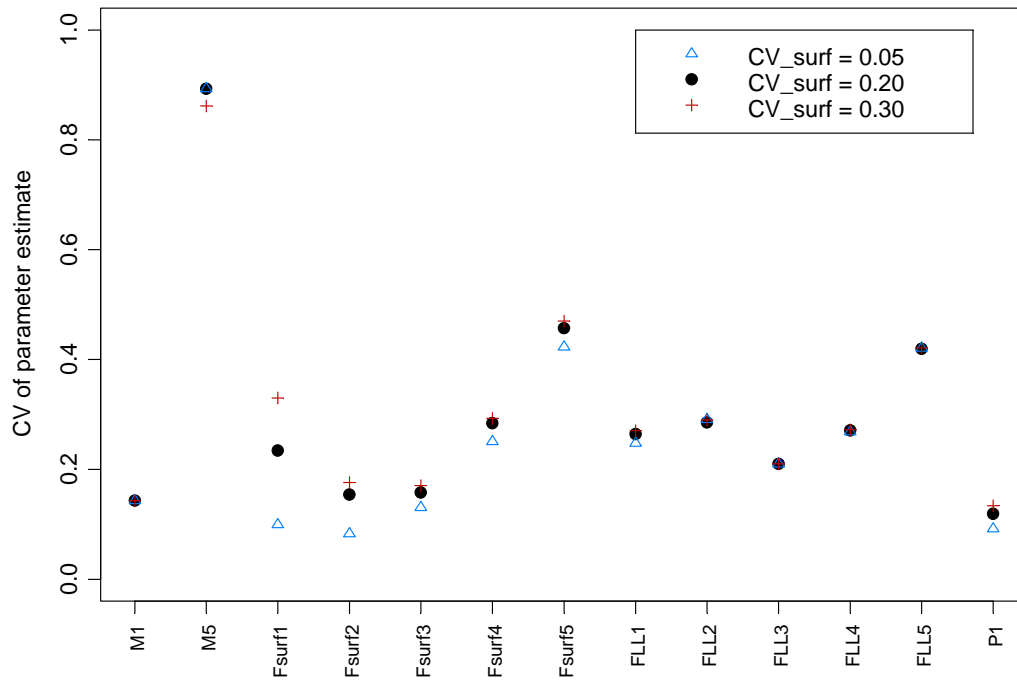
Figure 7. Effect of varying the level of observer coverage on the coefficient of variation (CV) of the key parameter estimates for the model with overdispersion in the tag return data.
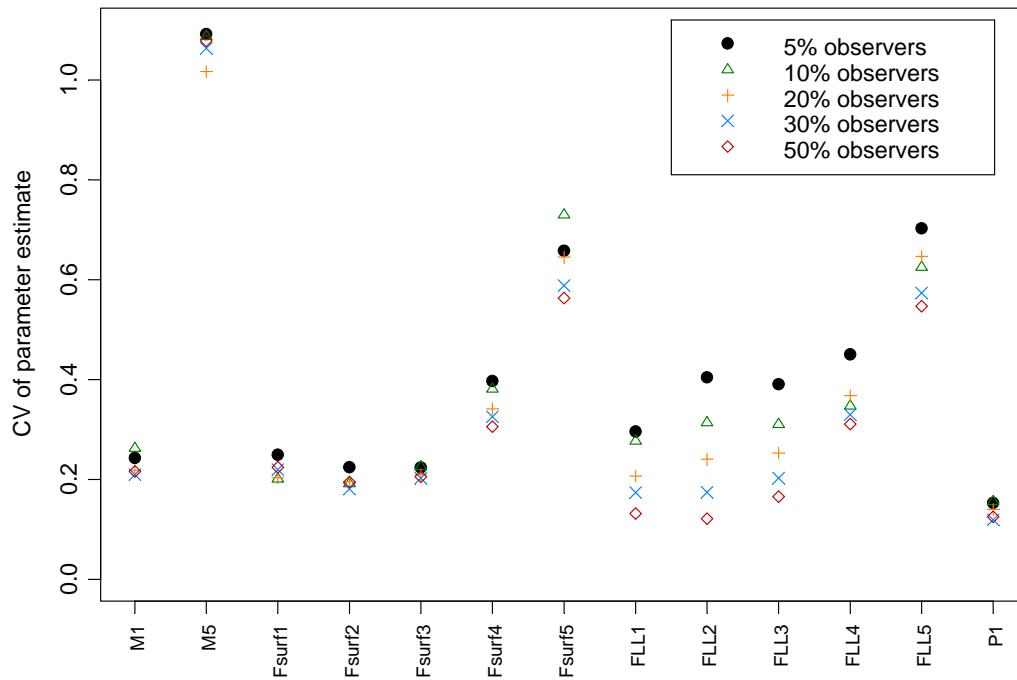
Figure 8. Effect of varying the level of observer coverage on the coefficient of variation (CV) of the key parameter estimates when there is no information about reporting rates for the surface fishery and a) the surface catch data are known with high precision (CV=0.05); b) the surface catch data are know with less precision (CV=0.20). Results are shown for model with multinomial tag returns.

a)
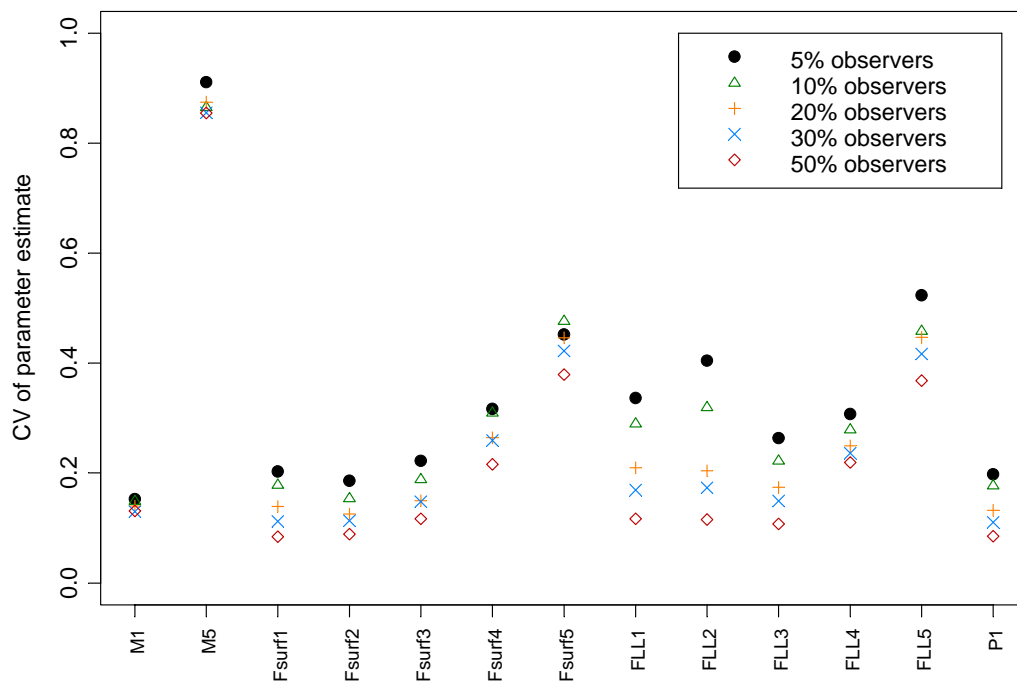
b)
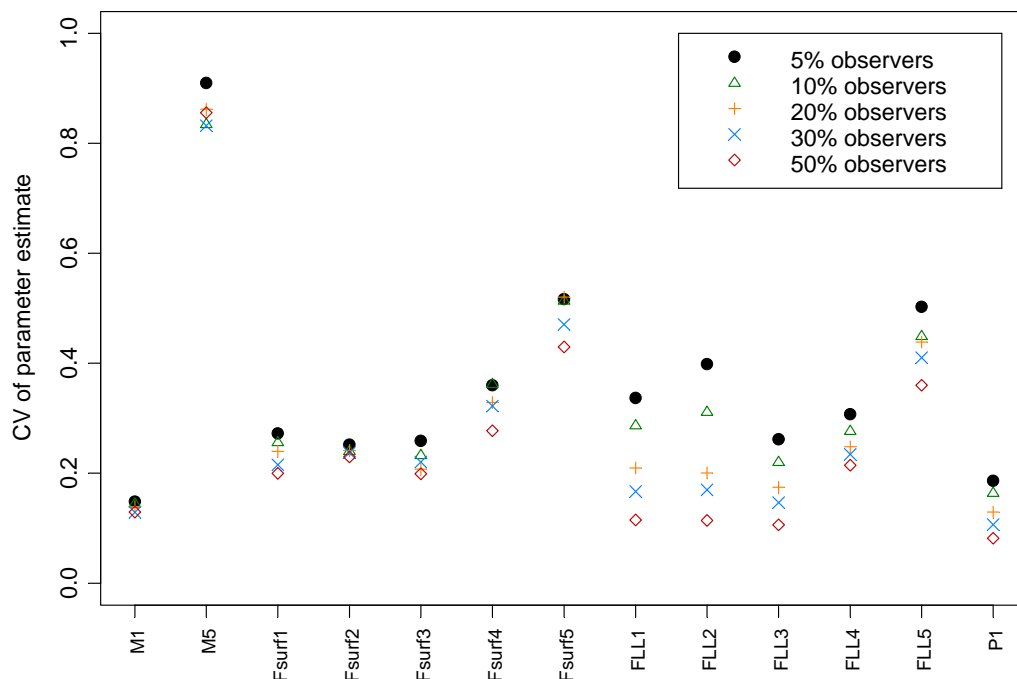
Appendix A.

The notation and parameterizations used in this Appendix were chosen to be representative of the estimation model with overdispersion presented in the main body of the paper. Thus, in the presentation below, N represents the number of tag releases at a particular age, $R_1, \ldots, R_{k-1}$ represent the number of tag returns at ages 1 to $k-1$, and R_k represents the number of tags that were not returned by age k. The π's are the random Dirichlet probabilities of return at age and the p's are their expected values (in our estimation model with overdispersion, the p's are analogous to the return probabilities given in equations (3)-(5)).

The Dirichlet distribution

The Dirichlet distribution is used to describe the variation in a set of proportions that sum to 1. The probability density of a set of proportions $\underline{\pi} = \{\pi_1, \ldots, \pi_k\}$ with parameter set $\underline{p} = \{\omega, p_1, \ldots, p_k\}$ is given by:

$$\Pr(\underline{\pi}) = \frac{\Gamma(\omega)}{\prod_{i=1}^{k}(\omega p_i)} \prod_{i=1}^{k} \pi_i^{\omega p_i - 1}$$

where $0 < p_i < 1$ for all i and $\sum_{i=1}^{k} p_i = 1$.

The mean and variance of the proportions are:

$$E[\pi_i] = p_i$$

and

$$V[\pi_i] = \frac{p_i(1-p_i)}{\omega+1}.$$

Note that the Dirichlet distribution with $k = 2$ reduces to the beta distribution.

The Dirichlet-multinomial distribution

The multinomial distribution describes a situation in which N independent random trials are conducted and the outcome of each trial can fall into one of k categories; the probability of falling into category i is π_i ($\sum_{i=1}^{k} \pi_i = 1$). The final category counts $\underline{R} = \{R_1, \ldots, R_k\}$, where $\sum_{i=1}^{k} R_i = N$, have a multinomial distribution with probability density:

$$\Pr(\underline{R}) = \frac{N!}{\prod_{i=1}^{k} R_i!} \prod_{i=1}^{k} \pi_i^{R_i}$$

When the category probabilities are themselves viewed as random variables following a Dirichlet distribution, then the multinomial probability density given above describes the

conditional distribution of the category counts *given* the probabilities, which we denote by $\Pr\left(\underline{R}\mid\underline{\pi}\right)$. Then the *unconditional* distribution of the category counts is given by the compound distribution called the Dirichlet-multinomial with probability density:

$$\Pr\left(\underline{R}\right) = \int_{\underline{\pi}} \Pr\left(\underline{R}\mid\underline{\pi}\right)\Pr\left(\underline{\pi}\right)\,d\underline{\pi}$$

The integral is *k*-dimensional over all values of $\underline{\pi}$ such $0 \le \pi_i \le 1$ and $\sum_{i=1}^{k}\pi_i = 1$. It is easy to show that the resulting distribution is:

$$\Pr\left(\underline{R}\right) = \frac{N!}{\prod_{i=1}^{k}R_i!}\prod_{i=1}^{k}\frac{\Gamma\left(R_i+\omega p_i\right)}{\Gamma\left(\omega p_i\right)}$$

The mean and variance of the category counts are:

$$E\left[R_i\right] = N\,p_i$$

and

$$V\left[R_i\right] = \left(\frac{N+\omega}{\omega+1}\right)N\,p_i\left(1-p_i\right).$$

Recall that the variance of the category counts for the multinomial distribution is $N\,p_i\left(1-p_i\right)$ so that the variance for the Dirichlet-multinomial is a factor of $\left(N+\omega\right)/\left(\omega+1\right)$ times larger.

Appendix B. Additional results

Table B1. Mean of remaining reference case parameter estimates (from 100 simulations) for various levels of observer coverage (δ). True parameter values are given below parameter names.

	F_1^{*S}	F_2^{*S}	F_3^{*S}	F_1^{*LL}	F_2^{*LL}	F_3^{*LL}
δ	0.044	0.041	0.041	0.006	0.009	0.009
0.05	0.045	0.044	0.043	0.005	0.008	0.010
0.10	0.045	0.044	0.043	0.006	0.009	0.007
0.20	0.045	0.043	0.042	0.006	0.009	0.008
0.30	0.045	0.042	0.043	0.006	0.009	0.008
0.50	0.045	0.041	0.042	0.006	0.009	0.008

	λ_1^S	λ_2^S	λ_3^S	λ_4^S	λ_5^S	λ_1^{LL}	λ_2^{LL}	λ_3^{LL}	λ_4^{LL}	λ_5^{LL}
δ	0.65	0.65	0.65	0.65	0.65	0.10	0.10	0.10	0.10	0.10
0.05	0.648	0.631	0.632	0.622	0.623	0.189	0.141	0.103	0.110	0.107
0.10	0.644	0.633	0.633	0.624	0.626	0.146	0.117	0.101	0.101	0.101
0.20	0.644	0.638	0.641	0.631	0.621	0.134	0.107	0.097	0.098	0.100
0.30	0.644	0.644	0.641	0.627	0.623	0.143	0.101	0.096	0.101	0.093
0.50	0.643	0.650	0.645	0.630	0.624	0.142	0.112	0.097	0.097	0.100

Table B2. Standard deviation of remaining reference case parameter estimates (100 simulations) for various levels of observer coverage (δ).

δ	F_1^{*S}	F_2^{*S}	F_3^{*S}	F_1^{*LL}	F_2^{*LL}	F_3^{*LL}
0.05	0.008	0.009	0.043	0.005	0.005	0.008
0.10	0.008	0.009	0.043	0.005	0.004	0.007
0.20	0.008	0.010	0.042	0.004	0.003	0.005
0.30	0.007	0.009	0.043	0.003	0.002	0.004
0.50	0.006	0.009	0.042	0.002	0.002	0.004

δ	λ_1^S	λ_2^S	λ_3^S	λ_4^S	λ_5^S	λ_1^{LL}	λ_2^{LL}	λ_3^{LL}	λ_4^{LL}	λ_5^{LL}
0.05	0.084	0.084	0.074	0.084	0.091	0.165	0.091	0.041	0.033	0.038
0.10	0.082	0.083	0.074	0.083	0.089	0.137	0.067	0.026	0.032	0.032
0.20	0.082	0.084	0.076	0.083	0.090	0.116	0.051	0.028	0.023	0.029
0.30	0.082	0.078	0.072	0.082	0.087	0.119	0.042	0.024	0.023	0.029
0.50	0.082	0.076	0.073	0.084	0.086	0.116	0.057	0.028	0.027	0.038

Table B3. Mean of remaining reference case parameter estimates not in Table 5 (from 100 simulations) for various levels of observer coverage (δ) when overdispersion is incorporated in tag return data. True parameter values are given below parameter names.

	F_1^{*S}	F_2^{*S}	F_3^{*S}	F_1^{*LL}	F_2^{*LL}	F_3^{*LL}	x
δ	0.044	0.041	0.041	0.006	0.009	0.009	3.0
0.05	0.045	0.043	0.045	0.007	0.010	0.011	1.86
0.10	0.045	0.044	0.044	0.006	0.009	0.008	1.89
0.20	0.046	0.044	0.045	0.006	0.009	0.009	1.81
0.30	0.047	0.043	0.044	0.007	0.009	0.008	1.86
0.50	0.046	0.042	0.043	0.007	0.009	0.009	1.82

	λ_1^S	λ_2^S	λ_3^S	λ_4^S	λ_5^S	λ_1^{LL}	λ_2^{LL}	λ_3^{LL}	λ_4^{LL}	λ_5^{LL}
δ	0.65	0.65	0.65	0.65	0.65	0.10	0.10	0.10	0.10	0.10
0.05	0.643	0.640	0.641	0.631	0.613	0.185	0.134	0.122	0.120	0.110
0.10	0.643	0.636	0.653	0.624	0.622	0.175	0.126	0.111	0.109	0.106
0.20	0.643	0.636	0.642	0.627	0.621	0.119	0.125	0.104	0.104	0.097
0.30	0.643	0.646	0.648	0.623	0.611	0.133	0.121	0.108	0.109	0.096
0.50	0.649	0.642	0.651	0.632	0.617	0.126	0.113	0.104	0.103	0.099

Table B4. Standard deviation of remaining reference case parameter estimates (100 simulations) for various levels of observer coverage (δ) when overdispersion is incorporated in tag return data.

δ	F_1^{*S}	F_2^{*S}	F_3^{*S}	F_1^{*LL}	F_2^{*LL}	F_3^{*LL}	x
0.05	0.012	0.011	0.018	0.012	0.008	0.016	0.35
0.10	0.012	0.011	0.016	0.008	0.006	0.010	0.42
0.20	0.013	0.012	0.016	0.005	0.005	0.010	0.31
0.30	0.015	0.011	0.017	0.005	0.004	0.008	0.42
0.50	0.013	0.008	0.015	0.004	0.003	0.007	0.34

δ	λ_1^S	λ_2^S	λ_3^S	λ_4^S	λ_5^S	λ_1^{LL}	λ_2^{LL}	λ_3^{LL}	λ_4^{LL}	λ_5^{LL}
0.05	0.083	0.089	0.084	0.084	0.097	0.170	0.087	0.070	0.060	0.065
0.10	0.083	0.088	0.084	0.080	0.094	0.162	0.090	0.055	0.046	0.053
0.20	0.082	0.083	0.085	0.083	0.092	0.119	0.092	0.043	0.039	0.046
0.30	0.081	0.091	0.079	0.091	0.104	0.126	0.094	0.046	0.043	0.042
0.50	0.085	0.084	0.095	0.082	0.099	0.127	0.083	0.056	0.047	0.049

Table B5. Correlations between key parameter estimates for the reference case simulations (those with magnitude $\geq$ 0.5 are shaded).

	M_1	M_5	F_1^S	F_2^S	F_3^S	F_4^S	F_5^S	F_1^{LL}	F_2^{LL}	F_3^{LL}	F_4^{LL}	F_5^{LL}	P_1
M_1	1.0	-0.27	-0.07	0.15	0.24	0.14	0.05	-0.04	0.08	0.20	0.17	0.07	0.13
M_5		1.00	-0.04	-0.01	0.16	0.50	0.71	-0.01	-0.01	0.26	0.54	0.71	0.02
F_1^S			1.00	0.50	0.51	0.42	0.31	0.35	0.36	0.44	0.34	0.23	-0.69
F_2^S				1.00	0.64	0.52	0.4	0.36	0.49	0.56	0.49	0.38	-0.69
F_3^S					1.00	0.76	0.61	0.38	0.45	0.68	0.66	0.57	-0.69
F_4^S						1.00	0.82	0.28	0.37	0.70	0.80	0.77	-0.57
F_5^S							1.00	0.23	0.27	0.61	0.81	0.86	-0.40
F_1^{LL}								1.00	0.32	0.35	0.27	0.20	-0.48
F_2^{LL}									1.00	0.34	0.32	0.27	-0.51
F_3^{LL}										1.00	0.68	0.56	-0.63
F_4^{LL}											1.00	0.78	-0.52
F_5^{LL}												1.00	-0.36
P_1													1.00

Appendix 10:

Incorporating spatial structure into an integrated tag-recapture and catch model using a generic movement model

Geoff M. Laslett, J. Paige Eveson and Tom Polacheck

FRDC Project 2002/015

1 Introduction

In the design and analysis of tagging experiments, spatial heterogeneity and incomplete mixing of tagged and untagged fish throughout the population of interest are critical issues to consider. If non-trivial levels of spatial heterogeneity and/or incomplete mixing exist, then they must be accounted for in the design and analysis stages in order to get reliable estimates of mortality rates and abundance as well as the uncertainty in these estimates. In the current appendix, we expand the integrated multi-year tagging and catch-at-age model developed in Appendix 5 to allow for abundance and fishing mortality to differ among defined regions and for fish to move among these regions. This model assumes 100% (or known) reporting rates. We then expand the spatial model to incorporate the estimation of reporting rates using observer data, in a similar manner to how the non-spatial model was expanded in Appendix 7.

The spatial framework considered here is generic, without any restrictions on the movement patterns (i.e., transition probabilities) between regions. As long as the underlying spatial structure and time periods in the model are appropriate and the transition probabilities are allowed to vary with age and year/time-period, this framework should be able to represent most movement dynamics reasonably well.

For a given cohort of tagged fish, the spatial integrated multi-year tagging and catch-at-age models can provide age-specific estimates of natural mortality, age- and region-specific estimates of fishing mortality, region-specific estimates of abundance at the time of initial tagging, as well as age-specific transition rates between regions (issues of parameter identifiability are discussed in Appendix 13). In the model using observer data, age- and region-specific reporting rate estimates are also obtained. Often, however, it is not the regional parameter estimates for fishing mortality and abundance that are of primary interest, but rather the population-wide estimates (i.e., over all regions). Such population-wide estimates can be obtained using a non-spatial model, however we would expect them to be biased if spatial heterogeneity exists. To investigate the consequences (biases) of not allowing for

spatial heterogeneity when it exists, we compare estimates of total fishing mortality and abundance when regional data are analysed using the spatial model versus when the data are pooled across all regions and analysed using the equivalent non-spatial model (e.g., results from the spatial model with observers are compared with results from the non-spatial model with observers). Results for a range of scenarios are presented in order to explore in which situations spatial tag designs and models are necessary.

The above spatial models only consider tagging experiments using conventional tags, for which the time and location of release and recapture are known but no information between release and recapture is available. For a spatial model, the location of a fish between release and recapture is potentially very useful information; thus, we consider the benefit of supplementing the conventional tag releases with some archival tag releases, for which the entire spatial history of a fish between release and recapture is known.

Finally, the above models make the common assumption that the tag-recapture data are multinomially distributed. In most cases, the tag-recapture data will exhibit more variation than a multinomial model predicts (i.e., will be overdispersed multinomial) due to spatial heterogeneity and incomplete mixing. Although the spatial model is intended to account for much of this, some overdispersion is still likely to exist due to the patchy distribution (e.g., schooling nature) of many fish species. Thus, we show how the spatial model can be expanded to incorporate overdispersion in the tag-recapture data.

In summary, details and results for the following models will be presented in this appendix (in all cases, multi-year tag-recapture data and catch-at-age data are assumed to be available):

1. Spatial model applied to regional data, 100% reporting rates.

2. Non-spatial model applied to data pooled over regions, 100% reporting rates.

3. Spatial model applied to regional data, reporting rates estimated using observer data.

4. Non-spatial model applied to data pooled over regions, reporting rates estimated using observer data.

5. Model (1) including a proportion of archival tag releases.

6. Model (1) allowing for more variability in the tag-recapture data than a multinomial distribution predicts.

2 Model 1: the generic spatial Brownie model

Model (1) will, for brevity, be referred to as the generic spatial Brownie model (or often just the generic base model), noting that historically Brownie models do not include catch-at-age data and it would more accurately be referred to as the spatial integrated Brownie and Petersen model .

In order to make the problem tractable, we divide time into $t = 1, 2, \ldots, T$ time periods, and space into $r = 1, 2, \ldots, R$ regions. Here time period t refers to the interval $(t - 1, t]$. For convenience, we will consider time periods equal to years. For computational convenience, we shall impose a slightly artificial structure on the fishery. During year t (that is, during the period $(t - 1, t)$), the fish in a region stay within the region, but they may be caught or die naturally. Exactly at the end of each year (that is, at time t exactly), the fish will move between regions, according to a Markov chain model. The states of the Markov chain are the R regions, and the movement is governed by transition probabilities that depend on the age of the fish. This model conveniently separates the fishing and natural mortality processes from the movement processes.

Data of two types are available in each region and time period: tag-recapture and catch-at-age data. In our presentation of the data and the model, we shall focus on a single cohort of fish. Because of this focus, and because we are considering time periods of years, the age of the fish, denoted by a, equals the time period t. In short, $t = a$ in the model statements below.

We now describe the model in more detail. We assume that natural mortality depends only on age, denoted by M_a, and that fishing mortality depends on age and region, donted by F_{ar} (so that for T years and R regions, there are TR fishing mortality parameters). The conditional probability that a fish, alive at the beginning of time t (age a), survives is

$$S_{ar} = \exp(-M_a - F_{ar}).$$

Given that the fish does not survive, the probability that it is caught rather than dies naturally is

$$f_{ar} = \frac{F_{ar}}{M_a + F_{ar}}.$$

Thus, the probability that a fish, alive at the beginning of year t, is caught during that year is

$$(1 - S_{ar}) f_{ar}.$$

If we let R_t be the region that the fish is in at the start of (and during) time period t and $L_t = 1$ if the fish is alive at the start of time period t, then the probability that the fish is alive and in region r at the start of time period $t+1$ is given by:

$$\Pr\{R_{t+1} = r, L_{t+1} = 1\} \;\; = \;\; \sum_{r'=1}^{R} \Pr\{R_t = r', L_t = 1\} S_{tr'} \pi_{r'r|t}. \tag{1}$$

Here $\pi_{rs|t}$ is the probability of moving from region r to s at time t, that is, immediately after the interval $(t-1, t)$. If the fish is initially in region r_0 at time $t = 0$, it is assumed to remain in that region for period 1, so that $\Pr\{R_1 = r_0, L_1 = 1\} = 1$ and $\Pr\{R_1 = r, L_1 = 1\} = 0$ for $r \neq r_0$.

Set $p_{tr} = \Pr\{R_t = r, L_t = 1\}$. Given S_{tr} and $\pi_{rs|t}$, we can write the updating equations (1) for p_{tr} in matrix form. For example, in the case of two regions

$$\begin{pmatrix} p_{t+1,1} \\ p_{t+1,2} \end{pmatrix} = \begin{pmatrix} \pi_{11|t} & \pi_{21|t} \\ \pi_{12|t} & \pi_{22|t} \end{pmatrix} \begin{pmatrix} S_{t1} & 0 \\ 0 & S_{t2} \end{pmatrix} \begin{pmatrix} p_{t1} \\ p_{t2} \end{pmatrix} \tag{2}$$

where $p_{11} = 1, p_{12} = 0$ or $p_{11} = 0, p_{12} = 1$. Reading the right-hand side of (2) from right to left, the terms refer to the probability that a fish is alive and in a particular

region at the start of time period t, survives the time period, and then stays in its current region or else migrates to the other region. It is conventional to write the transition matrix as $\Pi_t = \{\pi_{rs|t}\}$. The rows then sum to 1. Note that the matrix

$$
\begin{pmatrix}
\pi_{11|t} & \pi_{21|t} \\
\pi_{12|t} & \pi_{22|t}
\end{pmatrix}
$$

that appears in (2) is the transpose of Π_t. Thus, (2) may be written in matrix form as

$$
p_{t+1} = \Pi_t' S_t p_t , \tag{3}
$$

where S_t is the $R \times R$ diagonal matrix with rth diagonal element $S_{tr} = e^{-M_t - F_{tr}}$.

2.1   Tagging data

The cohort of fish is tagged in each region at the beginning of each of years $t' = 1, 2, \ldots, T$. The number of fish tagged in region r' at the beginning of time t' is $N_{t'r'}$; we shall refer to these fish as tagging group $G_{t'r'}$. Recaptures of tagged fish occur over the same regions and years as releases.

For any given tagging group, the vectors p_t from (3) can be generated by following the principles outlined in the previous section. After this, we can calculate the probability that the tag is returned from region r during year t. In the base model we assume 100% reporting of tagged fish, so the relevant probability is

$$
q_{tr} = f_{tr}(1 - S_{tr})p_{tr} . \tag{4}
$$

The probability that the tag is not returned is

$$
\bar{q} = 1 - \sum_{t=1}^{T} \sum_{r=1}^{R} q_{tr} .
$$

We therefore obtain a set of probabilities.

$$
q = (\bar{q}, q_{11}, q_{12}, \ldots, q_{1R}, q_{21}, q_{21}, \ldots, q_{2R}, \ldots, q_{T1}, q_{T2}, \ldots, q_{TR}) \tag{5}
$$

For the tagging group $G_{t'r'}$, we will denote its corresponding vector q by $q_{|t'r'}$, with elements $\bar{q}_{|t'r'}$ and $q_{tr|t'r'}$. If, for example, there are $R = 3$ regions and $T = 3$ years, there will be $RT = 9$ tagging groups and, thus, $RT = 9$ separate vectors q. For the R groups tagged at the beginning of year 1, the only substantive difference is that the initial vectors p_1 are different. For the R groups tagged at the beginning of year 2, the elements q_{1r} of q are all 0. Similarly, the groups tagged at the beginning of year $t + 1$ have all elements $q_{jr} = 0$ for $j \leq t$ (i.e., the probability of a tag being returned from a fish before it has been tagged is 0).

2.2 Catch-at-age data

We now look at the model for catch-at-age data. Again we concentrate on a single cohort, the same as for the tag-recapture data. A complication with catch-at-age data is that, since the fish do not have tags, we do not know their origins. However, we do know their ages, hence we can work out in which year they were one-year-olds. Suppose the population size of age 1 fish at $t = 0$ (i.e., at the beginning of year 1) in region r_0 is P_{r_0}. Then the probability of a fish originating in region r_0 being recaptured in region r in year t is $q_{tr|1r_0}$. Thus, the probability of being recaptured in region r during time period t regardless of its source is

$$q_{tr|1.} = \frac{\sum_{r_0=1}^{R} q_{tr|1r_0} P_{r_0}}{\sum_{r_0=1}^{R} P_{r_0}}.$$

Denote the expected catch of this cohort in region r during time period t as μ_{tr}. Then

$$\mu_{tr} = \sum_{r_0=1}^{R} q_{tr|1r_0} P_{r_0}.$$

2.3 Inference

Inference for these data is fairly trivial. Suppose we put the tag-recapture data into a matrix n of dimension $I \times J$, in which the data in row i are the numbers of recaptures by region and time period from tagging group i (thus, $I = RT$). The first entry of each row is the number of unrecaptured tags, so that $J = 1 + RT$. The

probabilities of recapture are placed into a corresponding matrix Q. Row i of Q is the vector q from (5) for the relevant tagging group. Each row of data is treated independently as multinomial data — this implies that we regard sampling variation as the only significant source of variation in the data.

The log-likelihood is (apart from an additive constant)

$$l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \log Q_{ij} \, .$$

If there are structural zeroes in the data, this method of computing the log-likelihood will fail, because $Q_{ij} = 0$ for structural zeroes, and $\log Q_{ij} = -\infty$ then. The likelihood for structural zeroes is $0^0 = 1$, and the log-likelihood is 0. We can achieve this result by the following simple device. Let

$$Q_{ij}^* = \begin{cases} Q_{ij} & \text{if } n_{ij} > 0; \\ 0.01 & \text{otherwise.} \end{cases}$$

Then

$$l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \log Q_{ij}^* \, .$$

For the catch-at-age data, we adopt a Gaussian model with known coefficient of variation c. The reasons for choosing this model are explained in Appendices 5 and 7; briefly, the Gaussian model is intended to capture the variability that occurs due to the catch-at-age being estimated from a small length or age sample and due to catches from different vessels having different size/age compositions. The catches in each region and time are assumed to be statistically independent. If the catch in region r and time t is C_{tr}, the log-likelihood is (apart from an additive constant)

$$l_{catch} = \sum_{t=1}^{T} \sum_{r=1}^{R} \left[-\log \mu_{tr} - \frac{1}{2} \frac{(C_{tr} - \mu_{tr})^2}{c^2 \mu_{tr}^2} \right] \, .$$

The parameters are estimated by maximising the total log-likelihood

$$l = l_{tag} + l_{catch} \, .$$

Note that when a cohort is tagged in T consecutive years, only $T-1$ natural mortality rate parameters can be estimated (this is shown in Appendix 13, and is a well-known feature in non-spatial Brownie models). Thus, the parameters esimated are F_{tr} ($t = 1, \ldots, T$ and $r = 1, \ldots, R$), M_t ($t = 1, \ldots, T - 1$), P_{r_0} ($r_0 = 1, \ldots, R$), and $\pi_{rs|t}$ ($t = 1, \ldots, T$, $r = 1, \ldots, R$ and $s = 1, \ldots, R$), with the constraint that $\sum_{s=1}^{R} \pi_{rs|t} = 1$.

2.4 Overall (population-wide) fishing mortality

It is often of interest to compute population-wide fishing mortality parameters, averaged over all regions. Let P_{tr} be the number of fish in the cohort in region r at time $t - 1$, that is, at the beginning of time period t. The number of fish that survive over all R regions is

$$\sum_{r=1}^{R} P_{tr} \exp(-F_{tr} - M_t).$$

The overall fishing mortality parameter may be defined as

$$P_t \exp(-F_t - M_t)$$

where $P_t = \sum_{r=1}^{R} P_{tr}$ is the total (population-wide) abundance. This equation suggests that the overall fishing mortality operates in the same way as the regional fishing mortalities, but on the total population. Solving for F_t, we obtain

$$F_t = -\log\left(\frac{\sum_{r=1}^{R} P_{tr} \exp(-F_{tr})}{\sum_{r=1}^{R} P_{tr}}\right).$$

In practice, the expected population sizes are substituted for P_{tr}. The true overall fishing mortalities are computed using the true expected P_{tr} values and the true fishing mortalities, and the estimated overall mortalities use the estimated expected P_{tr} values and the estimated fishing mortalities F_{tr}. This type of reasoning is called a *counting argument* in statistical circles.

Note that a slightly different definition for the overall fishing mortality would result if we used a counting argument on the catch, or the number of natural deaths,

in time period t. This reflects the fact that the population dynamics have been modelled in terms of exponential and competing natural and fishing mortality rates (i.e., both occur simultaneously and continuously throughout the time periods).

2.5 Numerical example

In order to crystallise the nature of the data, we simulate some data from the base model using the true parameter values listed in Table 3. There are $R = 3$ regions and $T = 3$ years.

Table 1: Tag-recapture data from the generic base model

t'	r'	$N_{t'r'}$	unreturned tags	returned tags								
				$t=1$			$t=2$			$t=3$		
				$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$
1	1	1415	1012	114	0	0	136	5	4	110	16	18
1	2	1415	1008	0	102	0	12	140	10	19	105	19
1	3	1415	979	0	0	123	12	4	145	18	27	107
2	1	949	640	0	0	0	162	0	0	112	18	17
2	2	949	640	0	0	0	0	148	0	12	137	12
2	3	949	644	0	0	0	0	0	161	12	22	110
3	1	636	476	0	0	0	0	0	0	160	0	0
3	2	636	495	0	0	0	0	0	0	0	141	0
3	3	636	487	0	0	0	0	0	0	0	0	149

In Table 1, t' and r' are the time and region of the tagging group, and $N_{t'r'}$ is the number of fish tagged in region r' at the beginning of time period t'. There are 1415 fish of age 1 released in each of the 3 regions, 949 fish of age 2 and 636 of age 3. The declining number of tags reflects the declining population size with age. Of the 1415 tags released in region 1 just prior to time period 1, 114 are returned from region 1 during time period 1. None of these can be returned from regions 2 and 3 in time period 1, because, according to the model, these fish do not migrate to the other regions until time period 2. In each row of Table 1 the row sums are $N_{t'r'}$ — for example, $1415 = 1012 + 114 + 0 + 0 + 136 + 5 + 4 + 110 + 16 + 18$ in row 1. For age 2 fish tagged just prior to time period 2, there are of course no returns in time period

1. However, it is convenient, for programming purposes, to keep all tag-recapture data in the same format. Similar remarks apply to age 3 fish tagged just prior to period 3. Overall, 6381/9000=70.9% of tags are not returned.

The corresponding catch-at-age data are presented in Table 2. Here we have used a coefficient of variation of $c = 0.2$.

Table 2: Catch-at-age data from the generic base model

	$t = 1$			$t = 2$			$t = 3$	
$r = 1$	$r = 2$	$r = 3$	$r = 1$	$r = 2$	$r = 3$	$r = 1$	$r = 2$	$r = 3$
6161.02	7912.8	9300.83	8375.81	12591.2	15421.3	10956.3	14385.3	5665.28

Of course, theoretically, the numbers in Table 2 should be integers. In our simulated example they are not because they have been generated using a continuous Gaussian distribution. However, even in reality, it is possible (and common) for catch-at-age data to be non-integers since these numbers are often scaled up from samples. Note that we simply have the catch in each region for each year. The total catch is 90770 fish.

We simulated 100 such data sets and estimated the parameters by maximum likelihood, assuming that the coefficient of variation in the catch-at-age data, $c = 0.2$, is known. In Table 3 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 3: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of data generated from a model in which there is tagging of a single cohort of fish in three regions for three years

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.223	0.293	0.354	$\pi_{32|1}$	0.050	0.014	0.049	0.089
M_2	0.200	0.091	0.193	0.297	$\pi_{33|1}$	0.900	0.853	0.901	0.943
F_{11}	0.100	0.083	0.099	0.116	$\pi_{11|2}$	0.800	0.752	0.798	0.868
F_{21}	0.200	0.175	0.200	0.228	$\pi_{12|2}$	0.100	0.057	0.103	0.140
F_{31}	0.300	0.247	0.297	0.352	$\pi_{13|2}$	0.100	0.057	0.099	0.136

Table 3: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
F_{12}	0.100	0.081	0.100	0.113	$\pi_{21\|2}$	0.100	0.073	0.106	0.143
F_{22}	0.200	0.174	0.198	0.226	$\pi_{22\|2}$	0.800	0.738	0.790	0.852
F_{32}	0.300	0.258	0.299	0.359	$\pi_{23\|2}$	0.100	0.065	0.104	0.146
F_{13}	0.100	0.080	0.099	0.118	$\pi_{31\|2}$	0.100	0.073	0.100	0.143
F_{23}	0.200	0.175	0.199	0.222	$\pi_{32\|2}$	0.100	0.055	0.098	0.141
F_{33}	0.300	0.245	0.295	0.339	$\pi_{33\|2}$	0.800	0.754	0.802	0.848
$\pi_{11\|1}$	0.900	0.861	0.901	0.946	$\log_{10} P_1$	5.000	4.866	4.982	5.076
$\pi_{12\|1}$	0.050	0.018	0.049	0.079	$\log_{10} P_2$	5.000	4.846	4.998	5.092
$\pi_{13\|1}$	0.050	0.022	0.050	0.087	$\log_{10} P_3$	5.000	4.838	4.990	5.113
$\pi_{21\|1}$	0.050	0.017	0.047	0.072	F_1	0.100	0.088	0.099	0.110
$\pi_{22\|1}$	0.900	0.856	0.904	0.950	F_2	0.200	0.182	0.199	0.216
$\pi_{23\|1}$	0.050	0.018	0.049	0.087	F_3	0.300	0.258	0.296	0.338
$\pi_{31\|1}$	0.050	0.016	0.050	0.084					

All of the parameter estimates appear to be reasonably unbiased.

3   Model 2: pooled data from the generic spatial Brownie model

In this case the data are generated from the generic spatial Brownie model, and then pooled over regions. The data are analysed ignoring regions. The idea is to see whether any substantial biases are incurred by treating the data as arising from a single region.

3.1   Tagging data

Focus on a cohort of fish. Suppose that N_i age i fish are tagged in each of years $t = 1, \ldots, T$, and we have the number of returns for each year. Let $S_i = \exp(-M_i - F_i)$ denote the probability of survival to the end of age i, and $f_i = F_i/(M_i + F_i)$ be the conditional probability that the fish, if it does not survive, is caught rather than dies naturally. Here natural mortality (M_i) and fishing mortality (F_i) depend only on

age. For $T = 3$ years of releases and returns, we end up with the usual (non-spatial) Brownie model probabilities, as set out in Table 4.

Table 4: Probabilities of tag return by year after release.

t'	$N_{t'.}$	$t = 1$	$t = 2$	$t = 3$
1	N_1	$(1 - S_1)f_1$	$S_1(1 - S_2)f_2$	$S_1 S_2(1 - S_3)f_3$
2	N_2		$(1 - S_2)f_2$	$S_2(1 - S_3)f_3$
3	N_3			$(1 - S_3)f_3$

3.2 Catch-at-age data

For the catch-at-age data, the probabilities of catching a fish at each age are simply taken from the probabilities in row 1 of Table 4. Since we are ignoring the spatial structure in the data, there is no need to model the data as a mixture over R regions.

3.3 Inference

Each row of the tag-recapture data is treated as multinomial with the probabilities in the relevant table, similar to Table 4. The catch-at-age data are treated as Gaussian with known coefficient of variation. Note that there are essential inconsistencies in this model, in that mixtures of multinomial are not multinomial and mixtures of Gaussians with known coefficient of variation are not Gaussian with the same coefficient of variation. However, the point of this exercise is to see if ignoring the underlying spatial structure matters in practice. The parameters are estimated by maximising the joint likelihood of both data sets. Note that this is the model described in Appendix 5.

3.4 Numerical example

The data in Table 1, when pooled over regions, results in Table 5.

Table 5: Pooled tag-recapture data from the generic base model

t'	$N_{t'.}$	unreturned tags	returned tags $t=1$	$t=2$	$t=3$
1	4245	2999	339	468	439
2	2847	1924	0	471	452
3	1908	1458	0	0	450

Similarly, the corresponding catch-at-age data from Table 2, when pooled over regions, results in Table 6.

Table 6: Pooled catch-at-age data from the generic base model

$t=1$	$t=2$	$t=3$
23374.65	36388.29	31006.82

We repeated this pooling of data sets over regions for the same 100 data sets simulated in numerical example section for the generic base model, then estimated the parameters by maximum likelihood using the pooled model. In Table 7 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 7: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of *pooled* data, generated from a model in which there is tagging of a single cohort of fish in three regions for three years

Par	True	2.5%	Mean	97.5%
M_1	0.300	0.227	0.294	0.355
M_2	0.200	0.091	0.194	0.300
F_1	0.100	0.089	0.099	0.111
F_2	0.200	0.183	0.199	0.216
F_3	0.300	0.261	0.297	0.338
$\log_{10} P$	5.477	5.400	5.464	5.520

The parameter estimates appear to be unbiased in this situation, despite ignoring the underlying spatial structure.

3.5   Comparison of estimates from spatial and pooled data

We have investigated 18 scenarios to see how well the estimation procedure works under varying conditions. The 18 scenarios cover releases proportional to cohort abundance in each region versus equal releases in all regions, balanced versus unbalanced transitions between regions, equal versus unequal cohort abundance in each region, and equal and unequal fishing mortalities by region.

Balanced transitions mean that there is no change to the relative abundances of a cohort among regions; e.g., if a cohort starts out with equal proportions of age 1 fish in all regions in year 1, then the cohort will still be found in equal proportions in all regions at age 2 after transitions have occurred at the end of year 1. We consider scenarios with either relatively high or relatively low balanced transition rates between all regions. Unbalanced transitions mean that the proportional distribution of a cohort among regions will change with time/age. Also note that equal or unequal fishing mortalities by region refer to within a specific year/age and does not reflect how fishing mortality varies with age (in fact, in all 18 scenarios fishing mortality decreases with age).

Table 8 summarizes the general features of the 18 scenarios, and Tables 9 and 10 provide the specific parameter values used in each of the scenarios.

Table 8: Summary of the general features of the 18 scenarios considered (refer to text for full explanation of column headings; note that $\propto$ means 'proportional to')

	Releases		Transition rates			Initial abundance		Fishing mortality	
Scenario	$\propto$ abund	Equal	Bal,low	Bal,high	Unbal	Equal	Unequal	Equal	Unequal
1	x		x			x		x	
2	x		x			x			x
3	x		x				x	x	

Table 8: continued

Scenario	Releases		Transition rates			Initial abundance		Fishing mortality	
	∝ abund	Equal	Bal,low	Bal,high	Unbal.	Equal	Unequal	Equal	Unequal
4	x		x				x		x
5	x			x		x		x	
6	x			x		x			x
7	x			x			x	x	
8	x			x			x		x
9	x				x	x		x	
10	x				x	x			x
11	x				x		x	x	
12	x				x		x		x
13		x	x				x	x	
14		x	x				x		x
15		x		x			x	x	
16		x		x			x		x
17		x			x		x	x	
18		x			x		x		x

Table 9: True parameter values in the first 9 of the 18 scenarios

	releases proportional to abundance								
	balanced								unbal.
	low				high				→ area 3
	$P \stackrel{}{=}$		$P \neq$		$P \stackrel{}{=}$		$P \neq$		$P \stackrel{}{=}$
	$F \stackrel{}{=}$	$F \neq$	$F \stackrel{}{=}$	$F \neq$	$F \stackrel{}{=}$	$F \neq$	$F \stackrel{}{=}$	$F \neq$	$F \stackrel{}{=}$
	1	2	3	4	5	6	7	8	9
M_1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
M_2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
F_{11}	0.1	0.15	0.1	0.15	0.1	0.15	0.1	0.15	0.1
F_{21}	0.2	0.3	0.2	0.3	0.2	0.3	0.2	0.3	0.2
F_{31}	0.3	0.45	0.3	0.45	0.3	0.45	0.3	0.45	0.3
F_{12}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
F_{22}	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
F_{32}	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
F_{13}	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1
F_{23}	0.2	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
F_{33}	0.3	0.15	0.3	0.15	0.3	0.15	0.3	0.15	0.3
$\pi_{12|1}$	0.05	0.05	0.05	0.05	0.30	0.30	0.15	0.15	0.25
$\pi_{13|1}$	0.05	0.05	0.05	0.05	0.30	0.30	0.15	0.15	0.00
$\pi_{21|1}$	0.05	0.05	0.08	0.08	0.30	0.30	0.22	0.22	0.05
$\pi_{23|1}$	0.05	0.05	0.08	0.08	0.30	0.30	0.22	0.22	0.25
$\pi_{31|1}$	0.05	0.05	0.15	0.15	0.30	0.30	0.45	0.45	0.00
$\pi_{32|1}$	0.05	0.05	0.15	0.15	0.30	0.30	0.45	0.45	0.05

Table 9: scenarios 1 to 9: continued

	1	2	3	4	5	6	7	8	9
$\pi_{12\|2}$	0.10	0.10	0.05	0.05	0.40	0.40	0.15	0.15	0.25
$\pi_{13\|2}$	0.10	0.10	0.05	0.05	0.40	0.40	0.15	0.15	0.00
$\pi_{21\|2}$	0.10	0.10	0.08	0.08	0.40	0.40	0.22	0.22	0.05
$\pi_{23\|2}$	0.10	0.10	0.08	0.08	0.40	0.40	0.22	0.22	0.25
$\pi_{31\|2}$	0.10	0.10	0.15	0.15	0.40	0.40	0.45	0.45	0.00
$\pi_{32\|2}$	0.10	0.10	0.15	0.15	0.40	0.40	0.45	0.45	0.05
P_1	100000	100000	150000	150000	100000	100000	150000	150000	100000
P_2	100000	100000	100000	100000	100000	100000	100000	100000	100000
P_3	100000	100000	50000	50000	100000	100000	50000	50000	100000
N_{11}	1415	1413	2123	2154	1415	1414	2123	2154	1415
N_{12}	1415	1413	1415	1436	1415	1414	1415	1436	1415
N_{13}	1415	1413	708	718	1415	1414	708	718	1415
N_{21}	949	908	1423	1384	949	944	1423	1406	759
N_{22}	949	947	949	963	949	948	949	963	949
N_{23}	949	988	474	495	949	953	474	473	1138
N_{31}	636	577	954	859	636	651	954	900	413
N_{32}	636	636	636	645	636	639	636	641	611
N_{33}	636	704	318	347	636	624	318	308	884

Table 10: True parameter values in the last 9 of the 18 scenarios

	releases $\propto$ to abundance				releases equal				
	unbalanced				balanced			unbalanced	
	$\to$ area 3			low		high		$\to$ area 3	
	$\frac{P}{=}$ $\frac{F}{\neq}$	$\frac{P}{\neq}$ $\frac{F}{=}$	$\frac{P}{\neq}$ $\frac{F}{\neq}$	$\frac{P}{\neq}$ $\frac{F}{=}$	$\frac{P}{\neq}$ $\frac{F}{\neq}$	$\frac{P}{\neq}$ $\frac{F}{=}$	$\frac{P}{\neq}$ $\frac{F}{\neq}$	$\frac{P}{\neq}$ $\frac{F}{=}$	$\frac{P}{\neq}$ $\frac{F}{\neq}$
	10	11	12	13	14	15	16	17	18
M_1	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
M_2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
F_{11}	0.15	0.1	0.15	0.1	0.15	0.1	0.15	0.1	0.15
F_{21}	0.3	0.2	0.3	0.2	0.3	0.2	0.3	0.2	0.3
F_{31}	0.45	0.3	0.45	0.3	0.45	0.3	0.45	0.3	0.45
F_{12}	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
F_{22}	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
F_{32}	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
F_{13}	0.05	0.1	0.05	0.1	0.05	0.1	0.05	0.1	0.05
F_{23}	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2	0.1
F_{33}	0.15	0.3	0.15	0.3	0.15	0.3	0.15	0.3	0.15
$\pi_{12\|1}$	0.25	0.25	0.25	0.05	0.05	0.15	0.15	0.25	0.25
$\pi_{13\|1}$	0.00	0.00	0.00	0.05	0.05	0.15	0.15	0.00	0.00
$\pi_{21\|1}$	0.05	0.05	0.05	0.08	0.08	0.22	0.22	0.05	0.05
$\pi_{23\|1}$	0.25	0.25	0.25	0.08	0.08	0.22	0.22	0.25	0.25
$\pi_{31\|1}$	0.00	0.00	0.00	0.15	0.15	0.45	0.45	0.00	0.00

Table 10: scenarios 10 to 18: continued

	10	11	12	13	14	15	16	17	18
$\pi_{32\mid1}$	0.05	0.05	0.05	0.15	0.15	0.45	0.45	0.05	0.05
$\pi_{12\mid2}$	0.25	0.25	0.25	0.05	0.05	0.15	0.15	0.25	0.25
$\pi_{13\mid2}$	0.00	0.00	0.00	0.05	0.05	0.15	0.15	0.00	0.00
$\pi_{21\mid2}$	0.05	0.05	0.05	0.08	0.08	0.22	0.22	0.05	0.05
$\pi_{23\mid2}$	0.25	0.25	0.25	0.08	0.08	0.22	0.22	0.25	0.25
$\pi_{31\mid2}$	0.00	0.00	0.00	0.15	0.15	0.45	0.45	0.00	0.00
$\pi_{32\mid2}$	0.05	0.05	0.05	0.15	0.15	0.45	0.45	0.05	0.05
P_1	100000	150000	150000	150000	150000	150000	150000	150000	150000
P_2	100000	100000	100000	100000	100000	100000	100000	100000	100000
P_3	100000	50000	50000	50000	50000	50000	50000	50000	50000
N_{11}	1409	2123	2145	1000	1000	1000	1000	1000	1000
N_{12}	1409	1415	1430	1000	1000	1000	1000	1000	1000
N_{13}	1409	708	715	1000	1000	1000	1000	1000	1000
N_{21}	721	1115	1074	1000	1000	1000	1000	1000	1000
N_{22}	935	1044	1038	1000	1000	1000	1000	1000	1000
N_{23}	1179	688	718	1000	1000	1000	1000	1000	1000
N_{31}	359	595	523	1000	1000	1000	1000	1000	1000
N_{32}	592	700	677	1000	1000	1000	1000	1000	1000
N_{33}	987	613	680	1000	1000	1000	1000	1000	1000

For each of 18 scenarios, we ran 500 simulations of the data and estimated the parameters by maximum likelihood. For each simulation, the parameters were estimated from applying the generic spatial model to the regional data, and then population-wide fishing mortality and abundance estimates were calculated from the regional estimates. Population-wide parameter estimates were also estimated from applying the non-spatial model to the pooled data for comparison. The coefficient of variation of the catch-at-age data was assumed to be 0.2 in both the simulations and estimation phases. The mean $\bar{\hat{\theta}}$ of the 500 maximum likelihood estimates was the calculated for each parameter θ. We then computed the relative bias:

$$100 \times \frac{(\bar{\hat{\theta}} - \theta)}{\theta}.$$

The biases are reported in Table 11. Where the true parameter value was 0, we decided to report the absolute bias instead. However, the absolute biases were in fact tiny, and we have indicated such cases by $\sim$0.0.

Table 11: Relative biases of the parameter estimates in the 18 scenarios

	releases proportional to abundance												releases equal					
	balanced								unbalanced				balanced				unbal.	
	low				high				$\rightarrow$ area 3				low		high		$\rightarrow$ 3	
	$P \stackrel{=}{}$		$P \neq$		$P \stackrel{=}{}$		$P \neq$		$P \stackrel{=}{}$		$P \neq$		$P \neq$		$P \neq$		$P \neq$	
	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$	$F=$	$F\neq$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	-0.8	-0.9	0.4	-1.2	-1.6	-0.7	-0.9	-0.9	0.1	-0.7	-0.5	0.0	-0.1	-0.3	-1.5	0.3	-1.1	-0.3
M_2	-1.2	-0.7	-1.0	-0.6	-0.9	0.1	0.3	-1.5	-1.1	0.6	-0.2	-1.5	-0.7	-0.9	1.2	-1.7	-0.2	-0.5
F_{11}	-0.4	0.0	0.4	-0.5	-0.6	-0.2	-0.3	0.2	0.2	-0.5	0.2	0.3	0.4	0.4	0.1	0.6	-0.6	-0.1
F_{21}	-0.1	-0.2	-0.2	-0.2	0.2	0.1	0.1	-0.4	-0.8	0.5	-0.1	-0.1	0.0	0.4	-0.1	0.5	0.3	-0.1
F_{31}	-0.5	-0.3	0.2	0.4	0.5	0.7	0.3	0.1	0.3	0.4	0.6	0.1	-0.5	0.1	0.3	0.0	0.1	0.0
F_{12}	0.4	-0.1	-0.3	-0.3	-0.7	-0.1	-0.1	-0.4	0.6	-0.3	-0.7	-0.1	-0.5	-0.1	-0.4	0.4	-0.4	-1.0
F_{22}	-0.3	0.1	0.0	-0.3	-0.7	0.5	-0.3	0.0	0.1	0.2	0.1	0.0	-0.3	0.1	0.2	0.2	0.3	0.3
F_{32}	0.3	-0.2	-0.1	0.1	0.2	0.4	-0.2	0.1	0.4	0.3	0.2	0.1	0.3	-0.1	0.7	0.1	0.1	0.5
F_{13}	-0.4	-0.3	0.4	1.1	0.5	-0.5	-0.1	1.2	0.2	-0.6	0.4	-1.4	0.7	0.3	0.1	0.7	-0.6	0.9
F_{23}	-0.3	-0.1	0.1	-0.7	0.2	0.5	0.6	-0.5	-0.2	0.1	-0.2	-0.1	0.9	-0.4	0.5	0.0	-0.4	0.0
F_{33}	-0.1	-0.3	0.0	-0.6	-0.9	1.1	0.6	0.3	0.1	0.4	0.5	-0.4	0.0	0.3	-0.2	-0.6	-0.5	0.1
$\pi_{11\|1}$	0.0	0.1	0.0	-0.1	-0.8	1.2	-0.3	-0.2	0.1	0.1	0.1	0.2	0.0	0.1	0.0	-0.2	0.2	0.2
$\pi_{12\|1}$	-1.7	-1.0	-1.3	0.7	0.8	-0.8	0.9	-0.4	-0.4	-0.5	-0.3	-0.5	-0.2	-1.9	0.0	-1.0	-0.5	-0.7
$\pi_{13\|1}$	0.8	-0.3	1.3	0.7	0.3	-0.9	0.2	1.6	$\sim$0.0	$\sim$0.0	$\sim$0.0	$\sim$0.0	0.7	0.3	0.2	1.8	$\sim$0.0	$\sim$0.0
$\pi_{21\|1}$	-1.7	-0.3	0.8	0.2	-0.8	0.8	-0.1	-0.1	0.3	1.7	-0.6	1.1	-1.2	0.1	0.3	-1.3	0.0	3.0
$\pi_{22\|1}$	0.1	-0.2	0.2	-0.1	0.5	0.4	-0.1	-0.2	-0.1	-0.4	0.3	0.0	0.0	0.1	-0.3	-0.1	-0.1	-0.2
$\pi_{23\|1}$	0.1	4.0	-2.6	0.7	0.1	-1.4	0.2	0.7	0.2	0.7	-0.8	-0.2	1.7	-1.4	0.4	1.6	0.2	-0.1
$\pi_{31\|1}$	0.7	0.1	0.4	-2.7	0.2	-0.3	-0.3	-0.6	$\sim$0.0	$\sim$0.0	$\sim$0.0	$\sim$0.0	0.5	-1.0	-0.3	-0.8	$\sim$0.0	$\sim$0.0
$\pi_{32\|1}$	-2.2	2.3	0.1	-2.2	0.6	0.5	0.2	0.1	-1.6	0.2	-2.0	-1.1	1.3	-0.1	0.3	0.6	-4.3	-2.7
$\pi_{33\|1}$	0.1	-0.1	-0.1	1.0	-0.6	-0.1	0.3	2.4	0.1	0.0	0.0	0.0	-0.4	0.2	0.0	1.3	0.2	0.1
$\pi_{11\|2}$	-0.1	0.1	-0.1	-0.2	0.1	-0.1	-0.2	-0.7	0.1	0.0	-0.4	0.1	-0.1	0.0	0.0	-0.2	-0.3	0.0
$\pi_{12\|2}$	0.5	-0.1	-0.5	-0.4	-0.4	0.3	0.6	0.8	-0.3	-0.1	1.1	-0.2	-0.6	1.1	0.3	0.3	0.9	0.1
$\pi_{13\|2}$	0.0	-0.4	2.0	4.4	0.3	-0.3	0.2	2.5	$\sim$0.0	$\sim$0.0	$\sim$0.0	$\sim$0.0	1.7	-0.3	-0.2	0.6	$\sim$0.0	$\sim$0.0
$\pi_{21\|2}$	1.1	0.0	0.6	-0.1	-0.1	0.0	-0.2	0.5	0.5	-2.6	-3.1	0.1	0.0	-0.3	-0.3	0.0	-0.2	-1.3
$\pi_{22\|2}$	-0.2	0.1	-0.1	0.1	-0.6	0.5	0.1	-0.4	0.0	0.1	-0.1	-0.5	-0.1	-0.1	0.0	-0.3	-0.1	0.0
$\pi_{23\|2}$	0.9	-0.5	0.0	-0.6	0.3	-0.3	-0.2	0.5	0.0	0.3	0.8	1.3	0.8	1.8	0.2	0.7	0.4	0.2
$\pi_{31\|2}$	0.9	1.9	-0.3	0.4	-0.1	-0.7	0.1	-0.7	$\sim$0.0	$\sim$0.0	$\sim$0.0	$\sim$0.0	0.7	-0.3	0.3	0.1	$\sim$0.0	$\sim$0.0
$\pi_{32\|2}$	-0.5	-0.7	0.9	-1.2	-0.4	-0.2	-0.2	0.8	-0.4	0.9	0.9	0.5	0.7	0.0	-0.6	-0.6	1.8	1.3
$\pi_{33\|2}$	0.0	-0.1	-0.1	0.2	1.0	1.9	0.4	-0.5	0.0	-0.1	-0.1	-0.1	-0.3	0.1	1.3	2.3	-0.1	-0.1
P_1	-1.4	-1.4	-2.0	-0.8	0.7	-3.1	-1.0	-0.9	-1.2	-0.4	-1.1	-1.9	-1.0	-2.0	-0.7	-1.0	-1.3	-1.2
P_2	-2.2	-1.8	-0.6	-1.3	-1.2	-1.6	-0.6	-2.1	-2.1	-1.0	-1.2	-0.9	-1.1	-0.5	-2.1	-1.3	-0.2	-0.8
P_3	-1.3	-1.7	-0.3	-2.3	-2.4	-1.4	-3.2	-0.9	-1.9	-2.3	-1.7	-1.4	-1.9	-2.1	-2.6	-0.4	-2.2	-1.8
relative biases of population-wide estimates from the spatial data																		
M_1	-0.8	-0.9	0.4	-1.2	-1.6	-0.7	-0.9	-0.9	0.1	-0.7	-0.5	0.0	-0.1	-0.3	-1.5	0.3	-1.1	-0.3
M_2	-1.2	-0.7	-1.0	-0.6	-0.9	0.1	0.3	-1.5	-1.1	0.6	-0.2	-1.5	-0.7	-0.9	1.2	-1.7	-0.2	-0.5
F_1	-0.3	-0.3	-0.1	-0.4	-1.0	-1.1	-0.7	-0.4	0.1	-0.4	-0.4	-0.4	-0.2	-0.2	-0.7	-0.1	-1.0	-0.6
F_2	-0.4	-0.2	-0.2	-0.4	-0.5	0.1	-0.2	-0.7	-0.4	0.5	-0.2	-0.3	-0.1	0.0	-0.2	-0.2	-0.1	-0.1
F_3	-0.3	-0.3	-0.2	-0.1	-0.5	0.1	-0.2	-0.4	0.1	0.4	0.2	-0.4	-0.3	-0.2	0.1	-0.4	-0.3	0.0

Table 11: relative biases: continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
P	-1.6	-1.6	-1.2	-1.2	-1.0	-2.0	-1.3	-1.3	-1.7	-1.3	-1.3	-1.5	-1.2	-1.5	-1.5	-1.0	-1.1	-1.2

relative biases of population-wide estimates from the pooled data

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	-0.7	-0.8	0.5	-0.8	-1.0	-0.3	-0.5	-0.4	0.3	-0.5	-0.2	0.1	0.3	-7.6	-0.9	-14.0	-0.8	20.9
M_2	-0.7	-0.6	-0.4	0.0	0.1	1.8	1.3	0.4	-0.8	1.2	0.2	-0.9	-0.3	-9.2	1.3	-45.2	0.2	37.8
F_1	-0.1	0.0	0.2	-0.2	-0.1	0.0	0.0	0.3	0.4	-0.3	0.1	0.0	0.4	-14.6	0.2	-14.8	-0.5	-11.7
F_2	-0.2	-0.1	0.0	-0.2	-0.2	0.4	0.1	-0.2	-0.3	0.3	-0.1	0.0	0.2	-14.0	0.1	-14.9	0.0	-2.8
F_3	-0.1	-0.2	0.1	0.3	0.0	0.9	0.2	0.4	0.3	0.5	0.4	0.1	-0.1	-13.3	0.3	-18.3	-0.1	8.9
P	-3.1	-2.9	-2.6	-2.2	-2.5	-3.5	-2.7	-2.7	-3.1	-2.4	-2.7	-2.8	-2.6	6.6	-2.7	3.2	-2.4	9.6

In all 18 scenarios, all of the parameter estimates obtained using the spatial model appear to be unbiased. Although most estimates obtained from the pooled data and non-spatial model are unbiased, there are some exceptions. First, in all scenarios, the population-wide initial abundance estimate has a larger bias when estimated from the pooled data than when estimated from the spatial data. Second, for scenarios 14, 16 and 18, large biases occur in almost all parameter estimates obtained from the pooled data; these are the scenarios for which, within a given year, tagging is not in proportion to abundance and fishing mortalities are unequal across regions.

The coefficients of variation are reported in Table 12. The very large values occur when the true value of a parameter is 0.

Table 12: Coefficients of variation of the parameter estimates in the 18 scenarios

	releases proportional to abundance												releases equal					
	balanced								unbalanced				balanced				unbal.	
	low				high				$\rightarrow$ area 3				low		high		$\rightarrow$ 3	
	$\underset{=}{P}$		$\underset{\neq}{P}$		$\underset{=}{P}$		$\underset{\neq}{P}$		$\underset{=}{P}$		$\underset{\neq}{P}$		$\underset{\neq}{P}$		$\underset{\neq}{P}$		$\underset{\neq}{P}$	
	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$	$\underset{=}{F}$	$\underset{\neq}{F}$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	13	13	13	12	12	13	13	12	12	14	12	13	14	13	14	13	13	15
M_2	26	28	25	25	26	26	26	26	27	27	26	27	22	21	24	24	21	24
F_{11}	9	7	7	6	9	8	7	6	8	7	7	6	10	9	10	9	10	9
F_{21}	7	6	5	5	8	7	5	5	7	7	6	6	7	6	7	6	7	6
F_{31}	9	8	7	8	10	8	8	7	10	10	9	9	7	6	8	7	7	7
F_{12}	9	9	9	9	9	9	9	9	9	9	8	9	11	10	11	10	10	11
F_{22}	6	7	6	6	8	8	7	7	7	7	7	6	7	7	7	7	7	7
F_{32}	9	9	9	9	10	10	10	9	9	9	9	9	7	7	8	7	7	8
F_{13}	9	12	12	16	10	12	13	17	9	12	12	16	10	14	11	16	11	15

Table 12: coefficients of variation: continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
F_{23}	7	9	9	11	8	10	11	14	6	8	7	9	7	9	8	11	7	8
F_{33}	9	10	11	12	10	12	13	17	8	9	9	9	7	9	7	9	7	9
$\pi_{11\mid1}$	3	3	2	2	11	10	4	5	4	4	4	3	3	3	6	7	5	5
$\pi_{12\mid1}$	34	34	26	26	13	14	15	16	12	13	11	11	38	40	22	22	14	15
$\pi_{13\mid1}$	34	45	27	35	14	16	17	23	2236	1119	4	22	39	52	21	28	1714	545
$\pi_{21\mid1}$	34	28	25	21	14	12	14	13	31	26	32	28	29	27	18	16	37	32
$\pi_{22\mid1}$	3	3	3	3	12	11	8	9	5	6	5	6	4	4	8	9	6	7
$\pi_{23\mid1}$	34	46	25	36	14	16	17	21	13	16	13	15	32	41	18	21	16	19
$\pi_{31\mid1}$	34	28	24	21	13	12	13	11	479	1623	312	503	21	18	11	10	318	519
$\pi_{32\mid1}$	36	33	26	25	12	14	12	11	31	32	48	45	20	21	11	10	38	39
$\pi_{33\mid1}$	3	2	7	6	10	12	35	48	2	2	2	2	6	6	31	39	2	2
$\pi_{11\mid2}$	4	5	2	3	18	16	5	5	5	5	4	5	2	3	5	5	4	5
$\pi_{12\mid2}$	21	22	27	29	10	12	17	17	16	15	13	14	31	33	18	19	13	14
$\pi_{13\mid2}$	20	32	26	39	11	13	16	23	2236	1582	4	26	32	45	17	23	4	5
$\pi_{21\mid2}$	21	19	27	21	10	10	15	13	33	28	30	27	27	23	16	13	35	28
$\pi_{22\mid2}$	4	4	3	4	17	18	8	8	5	6	5	6	3	4	7	7	5	6
$\pi_{23\mid2}$	22	29	28	36	10	11	15	19	12	17	12	16	26	35	13	17	13	17
$\pi_{31\mid2}$	21	17	26	20	10	10	13	12	306	297	561	316	18	15	9	8	670	358
$\pi_{32\mid2}$	21	20	26	24	11	11	13	12	29	26	38	35	17	18	9	9	31	29
$\pi_{33\mid2}$	4	3	7	6	16	21	35	45	2	1	2	2	5	5	22	30	2	2
P_1	13	14	13	13	21	22	15	16	13	13	12	12	15	13	16	16	13	13
P_2	14	14	14	14	22	21	19	18	17	17	16	17	15	14	19	19	18	18
P_3	14	15	17	19	21	21	24	26	15	16	18	21	19	18	22	26	19	19
coefficients of variation of pooled estimates																		
M_1	13	13	13	12	12	13	13	12	12	14	12	13	14	13	14	13	13	15
M_2	26	28	25	25	26	26	26	26	27	27	26	27	22	21	24	24	21	24
F_1	5	6	5	6	5	8	5	7	5	7	5	5	7	7	6	7	6	7
F_2	4	5	4	5	4	5	4	5	4	6	4	5	4	4	4	4	4	5
F_3	7	8	7	7	7	7	7	8	8	8	7	8	6	6	6	6	5	6
P	7	7	7	7	7	7	7	7	7	7	7	7	8	7	7	7	8	7
coefficients of variation of estimates from the pooled data																		
M_1	13	13	13	11	12	13	13	12	12	14	12	12	14	14	13	15	13	12
M_2	26	27	25	24	26	24	26	23	27	27	26	26	22	23	23	41	21	17
F_1	5	5	5	5	5	5	5	5	5	6	5	5	7	6	6	6	6	6
F_2	4	4	4	4	4	4	4	4	4	5	4	4	4	4	4	4	4	4
F_3	7	7	7	7	7	7	7	6	8	7	7	7	6	5	6	5	5	6
P	7	7	7	8	7	7	7	8	7	7	7	8	8	8	7	8	8	8

There are some interesting general patterns. The fishing mortalities are more precisely estimated than the natural mortalities, and the diagonal entries of the transition matrices are more precisely estimated than the off-diagonal entries.

4 Model 3: the generic spatial Brownie model with observers

When recapture information comes from commercial fisheries, we do not expect 100% of recaptured tags to be returned/reported, nor do we expect the reporting

rates to be known; thus, they must be estimated. Although tag-recapture data by themselves contain some information about reporting rates, the information is generally weak and insufficient to distinguish non-reporting from natural and fishing mortality without making some fairly restrictive assumptions. Thus, auxiliary data for estimating reporting rates generally needs to collected. In this section, we expand the generic spatial Brownie model to incorporate the estimation of reporting rates for the situation in which observers monitor a portion of the catches. We assume that each region has a group of boats with observers on them, and that 100% of recaptured tags are returned from these boats.

We suppose that in each region in each time period the fishery can be divided into G components. In component $g = 1$, observers are present and all tags are returned. In the other components, a proportion λ_{trg} of tags are returned (note that we are allowing for the reporting rate to vary between time periods, regions and components). For consistency, we set $\lambda_{tr1} = 1$. We also know the proportion of the fishery δ_{trg} fished by each component. In practice, the proportions δ_{trg} are inferred from the size of the catch returned by each component. In theory, these are not known exactly, but we argue that the error attached to the δ's is much smaller than other sources of variation in the data, so we treat the δ_{trg} values as known. Of course

$$\sum_{g=1}^{G} \delta_{trg} = 1$$

for any t and r.

4.1 Tagging data

For our cohort of interest, focus on the group tagged in region r' just prior to time period t'. All the calculations are the same as before, except the final calculation of the probability that the tag is returned from each component.

We simply need to modify equation (4). Thus the probability that the fish is caught by component g during time period t in region r and the tag is returned is

$$q_{trg} \quad = \quad \lambda_{trg}\delta_{trg}q_{tr} \,. \tag{6}$$

The probability that the tag is not returned is

$$\bar{q} \;=\; 1 - \sum_{t=1}^{T}\sum_{r=1}^{R}\sum_{g=1}^{G} q_{trg}\,. \tag{7}$$

We therefore obtain a set of probabilities:

$$
\begin{aligned}
q \;=\; &\big(\bar{q}, q_{111}, q_{112}, \ldots, q_{11G}, q_{121}, q_{122}, \ldots, q_{12G}, \ldots, q_{1R1}, q_{1R2}, \ldots, q_{1RG}, \\[4pt]
&\; q_{211}, q_{212}, \ldots, q_{21G}, q_{221}, q_{222}, \ldots, q_{22G}, \ldots, q_{2R1}, q_{2R2}, \ldots, q_{2RG}, \\[4pt]
&\; \ldots, \\[4pt]
&\; q_{T11}, q_{T12}, \ldots, q_{T1G}, q_{T21}, q_{T22}, \ldots, q_{T2G}, \ldots, q_{TR1}, q_{TR2}, \ldots, q_{TRG}\big)
\end{aligned}
$$

For the tagging group $G_{t'r'}$, the corresponding vector q is denoted $q_{|t'r'}$, with elements $\bar{q}_{|t'r'}$ and $q_{trg|t'r'}$. If, for example, there are $R = 3$ regions, $T = 3$ years and $G = 2$ groups, there will be $RT = 9$ separate vectors q, each with 19 elements.

4.2 Catch-at-age data

The catch-at-age data are treated slightly differently than in the base model because we assume that only the component with observers will be aged. As such, we only have catch-at-age data from the observer component (component 1) of each region. The rest of the argument is the same as in Section 2.2. Denote the expected catch of this cohort in region r during time period t from component 1 as μ_{tr}. Then

$$\mu_{tr} = \sum_{r_0=1}^{R} q_{tr1|1r_0} P_{r_0}\,.$$

Note that $q_{tr1|1r0}$ in μ_{tr} refers to component 1.

4.3 Inference

The principles of the method of inference are the usual ones. The tag-recapture data are treated as multinomial, and the catch-at-age data as Gaussian with known coefficient of variation, $c = 0.2$. The parameters are estimated by maximum likelihood, where the parameters to be estimated are the same as for the generic base model except we now have $TR(G-1)$ reporting rate parameters to estimate as well.

4.4 Numerical example

In order to crystallise the nature of the data, we simulate some data from the generic spatial model with observers. We assume $R = 3$ regions, $T = 3$ years and $G = 2$ components. The true parameter values used to simulate the data are the same as those listed in Table 3, along with $\delta_{tr1} = 0.1$ and $\delta_{tr2} = 0.9$, and $\lambda_{tr2} = 0.25$ for all t and r.

Table 13: Tag-recapture data from the generic spatial model with observers

t'	r'	$N_{t'r'}$	unreturned tags																		
				returned tags																	
				$t=1$						$t=2$						$t=3$					
				$r=1$		$r=2$		$r=3$		$r=1$		$r=2$		$r=3$		$r=1$		$r=2$		$r=3$	
				component (g)																	
				1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
1	1	1415	1284	11	26	0	0	0	0	17	27	0	1	1	2	16	20	1	2	4	3
1	2	1415	1287	0	0	11	20	0	0	0	3	18	29	1	3	1	4	12	21	1	4
1	3	1415	1272	0	0	0	0	16	31	0	2	1	1	14	36	3	6	5	3	7	18
2	1	949	857	0	0	0	0	0	0	9	34	0	0	0	0	9	27	1	6	1	5
2	2	949	853	0	0	0	0	0	0	0	0	18	31	0	0	1	2	14	26	2	2
2	3	949	863	0	0	0	0	0	0	0	0	0	0	19	26	2	3	2	2	9	23
3	1	636	584	0	0	0	0	0	0	0	0	0	0	0	0	16	36	0	0	0	0
3	2	636	593	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	33	0	0
3	3	636	588	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	29

In Table 13, t' and r' are the time and region of the tagging group, and $N_{t'r'}$ fish are tagged in region r' just prior to time period t'. There are 1415 fish of age 1 released in each of the 3 regions, 949 fish of age 2 and 636 of age 3. The declining number of tags reflects the declining population size with age. Of the 1415 tags released in region 1 just prior to time period 1, 11 are returned from region 1 during time period 1 by component 1 (the observers), and 26 by component 2 (the non-observers), even though component 1 represents only 10% of the fishery. Overall, 8181/9000=90.9% of tags are not returned, substantially more than for the base model.

The corresponding catch-at-age data are presented in Table 14. Here we have used a coefficient of variation of $c = 0.2$.

Table 14: Catch-at-age data from the generic spatial model
with observers

	$t=1$			$t=2$			$t=3$	
$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$
809.92	616.10	791.28	1246.91	837.581	1259.12	1480.48	1095.63	1438.53

The total catch is 9576 fish, considerably less than the 90770 in the base model because our catch-at-age data merely come from component 1 of the fishery.

We simulated 500 such data sets and estimated the parameters by maximum likelihood, assuming that the coefficient of variation in the catch-at-age data, $c = 0.2$, is known. (In the base model we used only 100 simulations. The number was increased to 500 here because the expected returns are smaller.) In Table 15 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 15: True parameter values, mean maximum likelihood
estimates and summary statistics for 500 sets of data gener-
ated from a model in which there is tagging of a single cohort
of fish in all regions for three years and partial reporting of
tags by some components of the fishery

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.151	0.293	0.459	$\pi_{13\mid2}$	0.100	0.030	0.098	0.176
M_2	0.200	0.010	0.194	0.385	$\pi_{21\mid2}$	0.100	0.029	0.100	0.182
F_{11}	0.100	0.063	0.101	0.149	$\pi_{22\mid2}$	0.800	0.684	0.798	0.896
F_{21}	0.200	0.138	0.200	0.277	$\pi_{23\mid2}$	0.100	0.031	0.102	0.191
F_{31}	0.300	0.198	0.302	0.444	$\pi_{31\mid2}$	0.100	0.031	0.100	0.188
F_{12}	0.100	0.060	0.102	0.145	$\pi_{32\mid2}$	0.100	0.030	0.101	0.182
F_{22}	0.200	0.138	0.201	0.278	$\pi_{33\mid2}$	0.800	0.698	0.799	0.911
F_{32}	0.300	0.189	0.302	0.438	$\log_{10} P_1$	5.000	4.796	4.982	5.132
F_{13}	0.100	0.064	0.101	0.149	$\log_{10} P_2$	5.000	4.813	4.981	5.144
F_{23}	0.200	0.142	0.199	0.274	$\log_{10} P_3$	5.000	4.809	4.981	5.124
F_{33}	0.300	0.184	0.300	0.446	λ_{11}	0.250	0.137	0.256	0.412
$\pi_{11\mid1}$	0.900	0.815	0.898	0.968	λ_{12}	0.250	0.171	0.255	0.362
$\pi_{12\mid1}$	0.050	0.000	0.053	0.128	λ_{13}	0.250	0.179	0.253	0.348
$\pi_{13\mid1}$	0.050	0.000	0.049	0.119	λ_{21}	0.250	0.142	0.260	0.457
$\pi_{21\mid1}$	0.050	0.000	0.052	0.119	λ_{22}	0.250	0.171	0.255	0.393

Table 15: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
$\pi_{22|1}$	0.900	0.812	0.895	0.972	λ_{23}	0.250	0.188	0.255	0.364
$\pi_{23|1}$	0.050	0.000	0.053	0.119	λ_{31}	0.250	0.144	0.260	0.442
$\pi_{31|1}$	0.050	0.000	0.049	0.112	λ_{32}	0.250	0.174	0.255	0.367
$\pi_{32|1}$	0.050	0.000	0.050	0.114	λ_{33}	0.250	0.183	0.257	0.347
$\pi_{33|1}$	0.900	0.810	0.902	0.980	F_1	0.100	0.079	0.099	0.121
$\pi_{11|2}$	0.800	0.698	0.802	0.895	F_2	0.200	0.162	0.197	0.236
$\pi_{12|2}$	0.100	0.033	0.099	0.184	F_3	0.300	0.213	0.297	0.395

Most parameter estimates are reasonably unbiased, although the initial population counts are slightly too small on average, and the reporting rates ($\hat{\lambda}$s) slightly too large.

5 Model 4: pooled data from the generic spatial Brownie model with observers

In this model the data are generated from the generic spatial Brownie model with observers, and then pooled over regions. The data are analysed ignoring regions. The idea is to see whether any substantial biases are incurred by treating the data as arising from a single region in the case where reporting rates are estimated from observer data.

5.1 Tagging data

The model used to analyse the tagging data is essentially the same as that in Section 3.1, but with the added complication of components. Let δ_{tg} be the proportion of the fishery at time t fished by component g, and λ_{tg} be the proportion of tags returned by component g, where $\lambda_{t1} = 1$. For $T = 3$ years of releases and returns and $G = 2$ components, the Brownie model probabilities of tag returns are set out in Table 16.

Table 16: Probabilities of tag return by year after release.

t'	$N_{t'.}$	component	$t = 1$	$t = 2$	$t = 3$
1	N_1	1	$(1 - S_1)f_1\lambda_{11}\delta_{11}$	$S_1(1 - S_2)f_2\lambda_{21}\delta_{21}$	$S_1 S_2(1 - S_3)f_3\lambda_{31}\delta_{31}$
		2	$(1 - S_1)f_1\lambda_{12}\delta_{12}$	$S_1(1 - S_2)f_2\lambda_{22}\delta_{22}$	$S_1 S_2(1 - S_3)f_3\lambda_{32}\delta_{32}$
2	N_2	1		$(1 - S_2)f_2\lambda_{21}\delta_{21}$	$S_2(1 - S_3)f_3\lambda_{31}\delta_{31}$
		2		$(1 - S_2)f_2\lambda_{22}\delta_{22}$	$S_2(1 - S_3)f_3\lambda_{32}\delta_{32}$
3	N_3	1			$(1 - S_3)f_3\lambda_{31}\delta_{31}$
		2			$(1 - S_3)f_3\lambda_{32}\delta_{32}$

Note that we could simplify Table 16 by omitting all λ_{t1} parameters since they equal 1.

The λ_{tg} parameters (for $g > 1$) are estimated from the data, but the δ_{tg} parameters (proportions of the fishery in each component) are assumed known. In reality, we know the δ values by region (δ_{trg}), so we need to average these sensibly to carry out analysis of the pooled data. It is important to distinguish S_t here, the pooled conditional survival probability during time t, from S_t in (3), the diagonal matrix of regional survival probabilities during time t.

We suggest weighting the δ_{trg} values by the catch-at-age data. This parallels the way the overall δ_{tg} values would be calculated in practice. Let C_{tr} be the catch-at-age from the cohort of interest reported by the observer component $g = 1$ during time interval t in region r. The estimated total catch from this cohort during time interval t in region r is therefore C_{tr}/δ_{tr1}. We therefore suggest the weighted value

$$\delta_{tg} = \frac{\sum_{r=1}^{R} \delta_{trg} C_{tr}/\delta_{tr1}}{\sum_{r=1}^{R} C_{tr}/\delta_{tr1}} . \tag{8}$$

It is immediate that $\sum_{g=1}^{G} \delta_{tg} = 1$. Note that δ_{t1} is a weighted harmonic mean. If $\delta_{tr1} = 0$ for any t and r, then $C_{tr} = 0$, so (8) will contain 0/0 terms. There seems to be little we can do about this, except to be alert to the possibility.

It should be emphasised that the use of (8) implies that, in general, the tag-recapture data cannot be analysed unless catch-at-age data are available. Also, the

δ_{tg} will be regarded as known and fixed, even though they rely on quantities (catch-at-age values) that are subject to error. However, if δ_{trg} does not vary with r for given t and g, then δ_{tg} does not rely on catch-at-age data.

In many of the scenarios we shall study, δ_{trg} does not depend on t; however, in principle, δ_{tg} still depends on t through C_{tr}.

It is also important to understand the link between the λ_{tg} parameters and their regional equivalents λ_{trg}. This is not trivial. Essentially we have to use a *counting argument* again. This leads to

$$\sum_{r=1}^{R} P_{tr}(1 - S_{tr}) f_{tr} \delta_{trg} \lambda_{trg} = P_t(1 - S_t) f_t \delta_{tg} \lambda_{tg} \,. \tag{9}$$

Although this can be solved explicitly for λ_{tg}, the salient point is that λ_{tg} essentially depends, through P_{tr}, on all parameters of the model relating to time periods 1 to t.

5.2   Catch-at-age data

For the pooled catch-at-age data we have only a single number C_t for each time period for the target cohort. This represents the catch taken by the observer component $g = 1$. If P is the total initial size of the cohort (at the start of time period 1), then the expected number of fish caught by component $g = 1$ during time period t is

$$\mu_t = PS_1 S_2 \ldots S_{t-1}(1 - S_t) f_t \delta_{t1} \,.$$

These are P multiplied by the probabilities in line 1 of Table 16 (for the case $R = T = 3$). Note that the λ_{tg} values in this line are all 1, since all tags are assumed to be returned from the observer component.

5.3   Inference

Each row of the tag-recapture data matrix is treated as multinomial with the probabilities in the relevant table, such as Table 16. The catch-at-age data are treated as Gaussian with known coefficient of variation.

5.4   Numerical example

The data in Table 13, when pooled over regions, results in Table 17.

Table 17: Pooled tag-recapture data from the generic spatial model with observers

t'	$N_{t'.}$	unreturned tags	returned tags					
			$t = 1$		$t = 2$		$t = 3$	
			component (g)					
			1	2	1	2	1	2
1	4245	3843	38	77	52	104	50	81
2	2847	2573	0	0	46	91	41	96
3	1908	1765	0	0	0	0	45	98

Similarly, the corresponding catch-at-age data from Table 14, when pooled over regions, results in Table 18.

Table 18: Pooled catch-at-age data from the generic spatial model with observers

$t = 1$	$t = 2$	$t = 3$
2217.30	3343.60	4014.64

We repeated this pooling of data sets over regions for the same 500 data sets simulated in the numerical example section for the generic spatial model with observers, then estimated the parameters by maximum likelihood using the pooled model with observers. In Table 19 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 19: True parameter values, mean maximum likelihood estimates and summary statistics for 500 sets of *pooled* data, generated from a model in which there is tagging of a single cohort of fish in all regions for three years and partial reporting of tags by some components of the fishery

Par	True	2.5%	Mean	97.5%
M_1	0.300	0.156	0.294	0.459
M_2	0.200	0.010	0.198	0.401
F_1	0.100	0.079	0.101	0.126
F_2	0.200	0.162	0.200	0.241
F_3	0.300	0.216	0.302	0.411
$\log_{10} P$	5.477	5.384	5.462	5.543
λ_{12}	0.250	0.172	0.252	0.348
λ_{22}	0.250	0.197	0.252	0.313
λ_{32}	0.250	0.208	0.252	0.304

For this example, there do not appear to be any significant biases incurred by pooling the data over regions.

5.5 Comparison of estimates from unpooled and pooled data

We re-examine the 18 scenarios explored in Section 3.5, augmented with 2 components, one with and one without observers. We assume $\delta_{tr1} = 0.1$ and $\delta_{tr2} = 0.9$ (i.e., the proportion of observers in each region and time period is 0.1), and $\lambda_{tr2} = 0.25$ for all t and r (i.e., the reporting rate in the component without observers is 0.25 for all time periods and regions).

For each of the 18 scenarios, we ran 500 simulations of the data and estimated the parameters by maximum likelihood. For each simulation, the parameters were estimated from applying the generic spatial model with observers to the regional data, and then population-wide fishing mortality and abundance estimates were calculated from the regional estimates. Population-wide parameter were also estimated from applying the non-spatial model with observers to the pooled data

for comparison. The coefficient of variation of the catch-at-age data was assumed to be 0.2 in both the simulations and estimation phases.

The relative biases (as defined previously) are set out in Table 20. Note that because there is only one component without observers, we have dropped the component subscript on the λ parameters. Thus, for the spatial model, λ_{tr} refers to the reporting rate in component 2 for year t and region r, and for the pooled model, λ_t refers to the reporting rate in component 2 for year t.

Table 20: Relative biases of the parameter estimates in the 18 scenarios

	releases proportional to abundance												releases equal					
	balanced								unbalanced $\to$ area 3				balanced				unbal. $\to$ 3	
	low				high								low		high			
	$\overset{P}{=}$		$\overset{P}{\neq}$		$\overset{P}{=}$		$\overset{P}{\neq}$		$\overset{P}{=}$		$\overset{P}{\neq}$		$\overset{P}{\neq}$		$\overset{P}{\neq}$		$\overset{P}{\neq}$	
	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$	$\overset{F}{=}$	$\overset{F}{\neq}$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	-2.3	-1.6	0.1	-1.7	-2.2	0.2	-2.9	-1.8	-0.7	-2.6	-0.2	-0.9	-0.5	-1.5	-2.1	-0.3	-0.7	-0.7
M_2	-2.9	-0.1	-3.6	-1.1	-0.5	-2.0	1.5	-3.4	-1.2	1.7	-5.2	-2.8	0.8	0.3	-0.7	-2.1	-2.9	-1.2
F_{11}	0.8	0.5	0.3	-0.4	3.9	1.4	2.4	1.5	0.8	0.7	1.6	-1.4	0.5	1.2	3.4	2.8	-0.6	-0.5
F_{21}	-0.1	0.1	1.3	-0.8	-0.3	1.0	-0.6	-0.7	1.0	-1.2	0.2	-0.6	0.5	0.9	-0.3	0.2	-1.4	0.7
F_{31}	1.0	2.0	1.3	1.3	2.6	3.0	2.3	1.0	1.3	0.2	1.3	1.2	-0.1	-0.1	1.4	0.9	0.4	2.0
F_{12}	1.6	0.1	1.5	2.1	1.6	3.7	2.4	3.8	2.1	-0.5	0.7	3.1	0.1	0.6	3.1	5.7	2.6	2.9
F_{22}	0.5	0.6	-0.5	1.3	0.2	1.0	0.4	-0.5	1.0	-0.2	-0.8	0.4	0.1	0.3	-0.6	0.0	0.3	2.1
F_{32}	0.8	-0.2	-0.3	1.4	1.8	1.2	1.1	0.1	2.0	0.3	0.2	-0.6	1.0	0.4	0.6	-0.8	-1.1	0.2
F_{13}	0.8	1.5	-0.2	4.2	3.1	4.4	2.4	6.7	0.3	2.7	3.5	8.5	0.3	1.5	2.0	1.0	3.9	3.8
F_{23}	-0.5	-0.5	-0.2	-1.6	2.1	2.1	-0.2	3.9	-0.1	-0.7	0.6	-0.5	1.0	-1.1	0.2	1.8	0.8	-1.7
F_{33}	0.3	-0.2	2.0	1.5	0.5	0.6	1.2	1.2	0.4	1.9	-1.1	-0.3	0.5	0.6	-0.9	0.2	-0.3	0.3
$\pi_{11|1}$	-0.2	0.2	-0.1	0.0	1.5	-0.5	-0.2	0.3	0.0	0.1	0.2	0.1	0.2	0.7	0.0	0.0	0.1	0.1
$\pi_{12|1}$	5.0	-2.3	-1.1	-1.2	0.4	-1.3	-1.5	-2.6	-0.1	-1.1	-0.5	-0.6	6.8	-9.2	-1.0	-0.8	-0.6	-2.2
$\pi_{13|1}$	-1.1	-0.6	3.7	1.1	-2.4	3.6	2.4	1.1	~0.0	~0.0	~0.0	~0.0	-9.9	-4.2	0.9	1.0	~0.0	~0.0
$\pi_{21|1}$	4.8	-2.4	-1.6	0.9	0.4	0.3	-0.6	0.4	-5.6	4.2	3.1	-2.1	4.1	0.5	-0.9	0.7	-2.5	-1.2
$\pi_{22|1}$	-0.6	0.1	0.2	-0.5	-0.8	-0.9	-1.3	-0.3	0.4	0.4	0.1	-0.1	-0.4	-0.4	-0.4	-0.5	0.0	-0.2
$\pi_{23|1}$	5.4	0.6	-0.4	5.0	0.7	2.0	3.8	0.2	-0.1	-2.0	-1.0	0.8	0.7	4.2	1.9	0.4	0.4	0.7
$\pi_{31|1}$	-2.7	3.5	-0.6	0.6	-0.1	0.4	0.6	-0.9	~0.0	~0.0	~0.0	~0.0	1.8	2.8	-0.9	-0.2	~0.0	~0.0
$\pi_{32|1}$	-0.8	0.2	1.6	0.0	1.4	-0.3	-1.2	1.7	-0.5	-4.2	-0.3	-0.9	2.3	-2.4	0.5	0.2	-5.3	-4.6
$\pi_{33|1}$	0.2	-0.2	-0.2	-0.1	-1.0	-0.3	2.6	-3.8	0.0	0.2	-0.1	0.0	-0.9	-0.1	1.5	0.0	0.2	0.2
$\pi_{11|2}$	0.3	-0.2	-0.4	-0.3	2.2	-1.3	-1.0	-1.0	-0.3	-0.1	-0.2	-0.3	-0.2	0.0	-0.1	-0.3	0.1	-0.2
$\pi_{12|2}$	-0.8	1.7	6.4	4.2	-0.8	1.0	1.3	1.6	0.6	-0.3	0.4	0.6	0.5	3.8	-1.9	0.5	-0.3	0.5
$\pi_{13|2}$	-1.6	-0.5	0.7	1.7	-0.3	5.1	3.2	2.9	~0.0	~0.0	~0.0	~0.0	2.9	-3.7	2.5	1.0	~0.0	~0.0
$\pi_{21|2}$	0.0	-0.6	1.0	-2.3	-0.4	0.0	-0.3	-1.0	-1.2	-0.8	-7.4	1.6	1.4	0.4	-0.5	-0.7	-4.1	-0.5
$\pi_{22|2}$	-0.2	0.4	0.3	0.2	-0.6	0.1	-1.0	-1.8	-0.1	0.6	-0.4	0.6	-0.3	-0.4	-0.3	0.0	0.4	0.8
$\pi_{23|2}$	1.9	-2.5	-3.9	0.5	0.7	-0.2	2.7	5.4	0.4	-1.5	2.6	-1.9	1.7	4.4	1.2	0.7	-0.4	-2.2
$\pi_{31|2}$	-0.5	-0.8	1.8	-0.6	-1.9	-1.6	-3.0	0.7	~0.0	~0.0	~0.0	~0.0	-1.2	1.3	-0.8	-1.1	~0.0	~0.0

Table 20: relative biases: continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\pi_{32\mid 2}$	1.2	-0.9	-0.5	1.9	0.8	0.6	1.6	-2.4	-1.2	1.5	-1.5	-3.0	-2.0	3.5	0.8	1.6	-1.3	5.1
$\pi_{33\mid 2}$	-0.1	0.2	-0.3	-0.3	2.1	4.6	6.5	7.4	0.0	-0.1	0.0	0.1	0.7	-1.0	-0.1	-2.0	0.0	-0.3
P_1	-2.3	-2.6	-1.6	-2.0	-2.2	-1.0	-3.3	-2.0	-1.3	-1.4	-1.9	-1.9	-1.6	-3.0	-3.0	-3.9	-1.3	-1.4
P_2	-2.4	-1.2	-1.8	-2.3	1.5	-3.2	-2.4	-2.6	-2.5	1.1	-0.6	-1.6	-0.9	-1.0	-0.7	-1.8	-0.7	-1.3
P_3	-2.4	-0.7	2.4	4.5	-1.1	-0.1	8.6	9.3	-1.2	-1.3	-0.7	-0.2	0.2	3.9	4.6	14.6	-1.6	1.4
λ_{11}	2.2	3.0	3.6	2.6	2.6	4.3	0.5	0.9	4.4	3.2	1.7	4.5	2.9	5.0	3.7	3.4	6.7	6.7
λ_{21}	1.9	1.5	0.4	2.1	3.2	0.9	2.3	2.0	3.4	3.2	1.0	2.3	1.9	0.5	3.0	2.0	4.4	1.2
λ_{31}	1.2	1.6	0.2	0.9	1.7	1.6	0.7	1.5	2.1	3.4	2.2	0.7	2.0	2.1	0.7	1.6	1.9	0.9
λ_{12}	3.9	4.7	3.9	4.2	4.7	2.8	4.2	2.8	3.7	5.2	5.5	3.4	5.8	4.4	6.1	0.8	4.5	4.6
λ_{22}	2.1	1.5	2.7	2.7	1.8	2.6	2.4	4.3	1.5	2.0	3.6	2.0	2.1	2.5	3.7	3.5	2.5	1.8
λ_{32}	2.2	2.0	1.6	1.0	2.0	2.2	2.7	3.2	1.6	3.4	1.9	3.3	1.9	2.0	1.4	2.6	3.3	2.7
λ_{13}	3.9	5.0	9.6	10.0	5.1	5.5	12.1	15.6	4.3	4.7	4.9	1.7	7.5	9.4	10.9	17.5	2.8	8.1
λ_{23}	2.2	5.2	4.0	8.4	1.0	2.7	3.0	6.1	3.0	4.0	2.7	6.6	2.5	5.6	2.5	4.0	1.3	5.8
λ_{33}	2.6	3.9	2.6	5.3	1.4	5.0	3.1	5.9	1.5	1.7	2.7	3.6	1.4	3.3	2.6	3.6	1.6	2.7
								relative biases of pooled estimates										
M_1	-2.3	-1.6	0.1	-1.7	-2.2	0.2	-2.9	-1.8	-0.7	-2.6	-0.2	-0.9	-0.5	-1.5	-2.1	-0.3	-0.7	-0.7
M_2	-2.9	-0.1	-3.6	-1.1	-0.5	-2.0	1.5	-3.4	-1.2	1.7	-5.2	-2.8	0.8	0.3	-0.7	-2.1	-2.9	-1.2
F_1	-1.1	-1.6	-1.7	-1.8	-2.6	-2.5	-2.4	-2.2	-1.3	-1.7	-0.7	-1.8	-2.2	-1.9	-2.9	-3.7	-2.0	-2.2
F_2	-1.3	-1.4	-1.0	-1.9	-1.3	-0.7	-1.9	-1.9	-0.7	-2.0	-1.2	-1.6	-0.9	-1.2	-2.0	-1.6	-1.6	-1.0
F_3	-0.8	-1.0	-0.7	-0.6	-0.3	0.0	-0.3	-1.7	-0.3	-0.6	-1.3	-1.6	-0.7	-1.3	-0.5	-1.4	-1.6	-0.9
P	-2.3	-1.5	-1.0	-1.0	-0.6	-1.4	-1.0	-0.3	-1.6	-0.5	-1.3	-1.5	-1.0	-1.2	-1.0	-0.1	-1.2	-0.9
							relative biases of estimates from the pooled data											
M_1	-2.0	-1.3	0.6	-1.2	-0.8	-15.2	-1.8	-0.6	-0.2	-2.1	0.2	-0.3	-0.1	-9.2	-1.5	-16.0	-0.2	21.2
M_2	-1.1	1.9	-1.9	0.6	3.5	-34.7	5.2	2.3	0.3	3.1	-3.3	-1.4	1.8	-7.6	0.9	-40.3	-1.5	38.1
F_1	0.6	0.0	-0.1	0.3	0.9	-1.7	0.0	0.7	0.3	0.3	0.9	-0.1	-0.6	-12.6	0.4	-11.1	0.3	-9.9
F_2	-0.2	0.2	0.3	-0.4	0.4	-11.5	-0.4	-0.1	0.3	-0.9	-0.1	-0.1	0.1	-13.8	-0.7	-14.6	-0.1	-2.4
F_3	0.6	0.6	0.7	1.3	1.8	-14.5	1.9	1.2	1.1	0.6	0.0	0.0	0.4	-13.8	0.5	-18.2	-0.4	9.3
P	-3.0	-2.3	-1.7	-2.1	-1.7	-2.6	-2.0	-1.3	-2.3	-1.4	-2.0	-2.5	-1.6	6.3	-2.0	2.2	-2.0	9.5
λ_1	0.8	1.5	2.2	1.5	0.9	1.2	1.3	0.9	1.8	1.2	1.1	1.9	2.2	-0.5	2.1	-4.5	1.7	0.9
λ_2	0.6	0.6	0.3	1.6	0.5	-0.3	0.9	1.6	1.2	1.4	0.8	1.1	0.5	0.2	1.3	0.9	0.7	0.8
λ_3	0.7	0.9	0.0	0.1	0.4	1.3	0.4	1.0	0.3	1.4	1.1	0.8	0.6	1.7	0.1	2.0	1.0	0.8

Note that in order to calculate the relative bias of the λ_t estimates from the pooled data, we needed to calculate 'true' λ_t values from the true region-specific λ_{tr} values. This was done using equation (9).

There are several points of interest. The λ values are consistently estimated with a small positive bias. Scenarios 6, 14, 16 and 18 yield poor estimates from the pooled data. In scenario 6, for example, the sampling distribution of $\hat{M}_2$ (from the pooled data) is quite skewed with many estimates near 0. It seems that there is not

enough information to estimate M_2. Once one parameter cannot be estimated well, others are affected also.

The coefficients of variation are reported in Table 21. The very large values occur when the true parameter value is zero.

Table 21: Coefficients of variation of the parameter estimates in the 18 scenarios

	releases proportional to abundance												releases equal					
	balanced								unbalanced $\to$ area 3				balanced				unbal. $\to$ 3	
	low				high								low		high			
	$P \atop =$		$P \atop \neq$		$P \atop =$		$P \atop \neq$		$P \atop =$		$P \atop \neq$		$P \atop \neq$		$P \atop \neq$		$P \atop \neq$	
	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$	$F \atop =$	$F \atop \neq$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	27	26	25	24	26	26	26	25	24	29	25	25	29	27	29	27	29	32
M_2	49	51	51	49	50	48	50	51	47	53	52	53	43	42	47	50	44	46
F_{11}	22	19	18	16	26	23	20	17	21	19	18	16	22	21	25	23	22	21
F_{21}	17	16	15	13	20	16	15	14	18	17	17	14	17	15	19	15	18	16
F_{31}	22	22	20	19	23	20	19	18	24	24	24	22	18	17	19	19	20	19
F_{12}	22	22	23	22	26	26	24	25	23	24	22	25	24	23	29	29	27	27
F_{22}	18	17	17	18	19	18	19	20	17	18	17	17	18	19	18	19	18	18
F_{32}	21	22	21	21	23	22	23	23	22	22	21	23	19	19	20	18	20	20
F_{13}	21	26	30	38	27	32	39	45	22	27	29	37	27	34	34	41	26	33
F_{23}	18	22	24	30	19	27	26	33	16	19	19	26	19	23	21	28	18	22
F_{33}	22	23	28	31	22	29	30	36	18	21	21	24	17	23	20	25	19	21
$\pi_{11|1}$	5	6	4	4	19	10	8	9	8	8	6	6	5	6	11	12	9	9
$\pi_{12|1}$	60	63	47	47	24	34	29	30	24	24	19	19	68	73	41	39	26	27
$\pi_{13|1}$	61	86	48	73	24	42	30	42	1040	433	1041	578	73	95	37	50	523	289
$\pi_{21|1}$	57	53	44	38	23	24	27	23	60	48	57	48	53	45	30	29	70	60
$\pi_{22|1}$	5	5	5	7	20	15	14	14	9	11	8	10	7	7	16	16	11	12
$\pi_{23|1}$	57	81	49	69	23	36	28	37	22	30	21	28	56	72	31	39	27	33
$\pi_{31|1}$	63	47	41	40	24	15	21	19	539	781	534	685	34	31	19	18	469	514
$\pi_{32|1}$	62	59	46	43	25	17	22	21	59	60	84	83	36	39	18	18	69	76
$\pi_{33|1}$	5	4	12	12	20	59	60	89	3	3	4	4	10	10	52	67	3	4
$\pi_{11|2}$	6	8	4	5	29	11	8	10	9	9	8	8	5	6	9	10	8	8
$\pi_{12|2}$	39	42	48	49	19	35	28	30	27	28	23	23	61	58	35	35	25	25
$\pi_{13|2}$	39	61	48	70	19	40	30	39	711	528	1262	636	58	88	27	41	10	2236
$\pi_{21|2}$	40	31	49	39	19	24	27	24	61	50	58	49	45	39	27	24	60	54
$\pi_{22|2}$	7	7	6	7	29	13	14	15	9	12	9	11	6	7	12	12	9	11
$\pi_{23|2}$	40	54	47	69	19	32	29	34	22	33	23	31	44	62	23	29	23	31
$\pi_{31|2}$	40	31	44	38	20	17	24	21	411	302	569	371	33	26	18	15	658	386
$\pi_{32|2}$	39	36	46	45	19	17	24	23	53	48	72	64	34	32	17	15	58	54
$\pi_{33|2}$	7	6	13	12	29	59	65	88	3	3	4	3	9	9	43	54	3	3
P_1	19	19	17	16	31	26	22	20	17	17	15	15	20	17	25	23	17	17
P_2	20	19	19	20	29	30	26	26	23	24	23	25	22	20	29	31	27	26
P_3	19	23	30	39	31	33	44	50	22	26	29	34	27	33	42	49	26	34
λ_{11}	28	25	25	20	34	26	25	22	27	26	24	20	33	29	35	30	33	30
λ_{21}	20	17	16	13	19	16	16	14	21	20	18	16	20	17	21	16	22	17
λ_{31}	18	16	15	13	16	13	14	12	20	19	18	14	15	13	15	13	17	16
λ_{12}	30	31	29	30	34	33	30	34	31	31	30	33	33	34	37	40	36	38
λ_{22}	21	19	20	20	19	19	20	20	19	18	18	18	21	20	21	20	21	21
λ_{32}	18	16	17	16	16	15	17	17	18	16	15	17	16	16	16	15	16	17
λ_{13}	29	41	40	57	35	43	54	59	29	40	40	54	37	50	48	58	39	51
λ_{23}	20	28	26	39	18	30	26	36	20	25	22	31	22	29	23	32	21	28
λ_{33}	16	21	22	29	16	27	22	28	15	18	17	22	16	22	17	24	16	20

coefficients of variation of pooled estimates

Table 21: coefficients of variation: continued

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	27	26	25	24	26	26	26	25	24	29	25	25	29	27	29	27	29	32
M_2	49	51	51	49	50	48	50	51	47	53	52	53	43	42	47	50	44	46
F_1	11	13	12	13	13	13	13	14	12	13	12	13	13	14	13	14	12	14
F_2	10	11	10	11	10	10	10	11	10	11	10	11	10	11	11	11	11	11
F_3	16	17	16	16	16	16	16	17	15	17	16	16	13	13	14	14	14	14
P	10	10	10	10	10	10	10	10	10	10	10	10	11	11	11	11	10	11

coefficients of variation of estimates from the pooled data

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M_1	27	26	25	24	26	30	26	25	24	28	25	25	29	28	28	30	29	25
M_2	48	50	50	48	48	65	48	46	47	51	51	52	42	44	45	73	43	32
F_1	12	13	13	13	13	13	13	13	13	13	13	13	13	14	14	14	14	14
F_2	11	11	10	10	10	10	10	10	10	11	10	10	10	10	11	10	11	10
F_3	16	16	16	15	16	15	15	15	15	16	16	16	12	13	13	13	14	14
P	10	10	10	11	10	10	10	10	10	10	10	11	11	12	11	11	11	11
λ_1	17	18	17	16	18	16	17	17	17	17	17	16	19	21	21	19	20	20
λ_2	12	12	12	11	11	11	12	12	12	12	11	11	12	12	12	12	12	13
λ_3	10	10	10	10	10	9	10	10	10	10	10	10	9	9	9	9	9	10

Comparing with the results in Table 12, the same general patterns are observed here as for the models with 100% reporting rates. In particular, the fishing mortalities are estimated more precisely than the natural mortalities, and the probabilities of staying in a region are estimated more precisely than the probabilities of transitioning between regions. However, the overall level of precision of the estimates is much poorer. For example, the coefficients of variation have approximately doubled for all common parameters except the initial population sizes (for which the increases are still reasonably large).

6 Model 5: including archival tagging data in the generic spatial Brownie model

Over the last 15 years, archival tags have been developed as a more informative alternative to conventional tags. Archival tags come in several forms, including some that relay information to satellites, but we shall focus only on the original style that are returned to the laboratory when the fish is caught, just like conventional tags. However, whereas a recaptured conventional tag merely tells the researcher where and when the fish was tagged and recaptured, an archival tag conveys the

intermediate times and locations as well. We shall assume that these are known without error. While this is not a very realistic assumption with regard to location, which is estimated from ambient light measurements, it will suffice for evaluating the maximum potential gain from augmenting a convential tagging experiment with archival tags.

Archival tags exploit modern micro-electronics and are therefore very expensive. The number of archival tags that can be deployed is small, in practice. If archival tags were released separately from the conventional tags, we would have to estimate fishing mortality and transition parameters specifically for the archival tags. This would render them practically useless. In this exercise we assume that the archival tags are released in parallel with the conventional tags. In fact, we assume that, given a number $N_{t'r'}$ of tags released in region r' at the beginning of time period t', there is a fixed proportion p_A of archival tags, so that $N_{t'r'}p_A$ are archival and $N_{t'r'}(1 - p_A)$ are conventional.

In this section, we ignore the need for observers by assuming that reporting rates are 100%, so that we can concentrate on the effectiveness of archival tags independent of reporting rate issues.

6.1   Archival tagging data

The probability of an archival tag not being recaptured is exactly the same as that of a conventional tag released from the same region at the same time. Thus, the probability is $\bar{q}$, as in equation (7).

The case of recaptured tags is more interesting because the tag provides information about where the fish was during each time period between release and recapture. Suppose that the fish is tagged at the beginning of time period t' and is recaptured in time period t^*, and that it was in region r_t for time period t, where $t' \leq t \leq t^*$. The probability of this event is

$$Q = \left(\prod_{t=t'}^{t^*-1} S_{tr_t} \pi_{t,r_t \, r_{t+1}} \right) \; (1 - S_{t^* r_{t^*}}) f_{t^* r_{t^*}} \; .$$

6.2   Inference

The parameters are estimated by maximising the likelihood. The conventional tag-recapture data and the catch-at-age data contribute the same likelihood components discussed for the generic base model (model 1).

For an unreturned archival tag, the contribution to the likelihood is $\bar{q}$, and for a returned archival tag, it is $\log Q$, where $\bar{q}$ and Q are derived for the appropriate tagging region and time and recapture history.

6.3   Numerical example

We give an example of the data to flesh out the concepts. This is structurally the same as the numerical example section for model 1, in that the true parameter values and the total numbers of tag releases are the same, although it is a different simulation. Let $N_{t'r'}$ denote the total tags released in region r' at the beginning of time period t', and let $M_{t'r'} = (1 - p_A)N_{t'r'}$ denote the number of conventional tags. In this example, we assume $p_A = 0.05$, so that only 5% of the releases are archival tags.

The conventional tagging data are set out in Table 22.

Table 22: Conventional tag-recapture data from the generic base model with archival tags

t'	r'	$N_{t'r'}$	$M_{t'r'}$	unreturned tags	returned tags								
					$t=1$			$t=2$			$t=3$		
					$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$
1	1	1415	1344	964	109	0	0	128	5	4	100	16	18
1	2	1415	1344	959	0	98	0	12	130	10	17	102	16
1	3	1415	1344	934	0	0	118	11	4	137	19	25	96
2	1	949	901	609	0	0	0	145	0	0	118	13	16
2	2	949	901	600	0	0	0	0	152	0	13	124	12
2	3	949	901	605	0	0	0	0	0	149	10	18	119
3	1	636	604	463	0	0	0	0	0	0	141	0	0
3	2	636	604	460	0	0	0	0	0	0	0	144	0
3	3	636	604	467	0	0	0	0	0	0	0	0	137

We now turn to the archival tags. Let $A_{t'r'} = p_A N_{t'r'}$ be the number of archival tags released in region r' at the beginning of time period t'. Of course

$$N_{t'r'} = M_{t'r'} + A_{t'r'} \,,$$

as can be checked for this example from the appropriate entries in Tables 22 and 23.

We need to treat the unreturned tags separately from the returned. The numbers of unreturned tags are set out in Table 23.

Table 23: Archival tag-recapture data from the generic base model with archival tags: total tag releases and unreturned tags

t'	r'	$A_{t'r'}$	unreturned tags
1	1	71	49
1	2	71	49
1	3	71	48
2	1	48	23
2	2	48	31
2	3	48	31
3	1	32	27
3	2	32	28
3	3	32	28

For the returned tags, we know the full history of regions visited by the fish. The first column of Table 24 gives the region that the fish was in during time period 1; the second gives the region it was in during time period 2; and the third gives the region it was in during time period 3. A zero in column 1 means that the fish has not yet been tagged — this, in fact, is the meaning of all leading zeroes. A trailing zero means that the fish has previously been caught. The final column is the number of recaptured tags, n_a, with the given history in the first three columns.

Table 24: Archival tag-recapture data from the generic base model with archival tags: history of returned tags (columns 1 to 3 specify the regions the fish were in during time periods 1 to 3 respectively)

$t = 1$	$t = 2$	$t = 3$	n_a
1	0	0	4
1	1	0	8
1	1	1	5*
1	1	2	4*
1	3	0	1
2	0	0	10
2	1	1	1*
2	2	0	7
2	2	2	4*
3	0	0	6
3	1	0	1
3	2	0	1
3	2	2	1*
3	3	0	8
3	3	3	6*
0	1	0	16
0	1	1	6
0	1	3	3
0	2	0	11
0	2	1	1
0	2	2	3
0	2	3	2
0	3	0	12
0	3	2	2
0	0	1	5
0	0	2	4
0	0	3	4

It should be noted that only the 21 cases marked with an asterisk yield more information than conventional tags. There are 9000 tags in total, so these represent less than 0.3% of the total. Even of the 2625 returned tags, it is only 0.8%, so it is difficult to imagine that these will have much impact on the estimation.

Note also that there are 71 archival tags deployed in region 1 at the start of

time period 1. Of these, 49 are not returned (Table 23) and 4+8+5+4+1=22 are returned (Table 24). Similar calculations apply to other times and periods.

We simulated 100 such data sets, with $p_a = 0.05$ of tags being archival, and estimated the parameters by maximum likelihood, assuming that the coefficient of variation in the catch-at-age data, $c = 0.2$, is known. In Table 25 we report summary statistics of the estimates.

Table 25: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of data generated from a model in which there is tagging (95% conventional, 5% archival) of a single cohort of fish in all regions for three years

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.212	0.296	0.357	$\pi_{32\mid1}$	0.050	0.019	0.050	0.088
M_2	0.200	0.105	0.194	0.307	$\pi_{33\mid1}$	0.900	0.844	0.899	0.942
F_{11}	0.100	0.087	0.100	0.118	$\pi_{11\mid2}$	0.800	0.753	0.800	0.856
F_{21}	0.200	0.176	0.200	0.231	$\pi_{12\mid2}$	0.100	0.060	0.102	0.140
F_{31}	0.300	0.253	0.296	0.346	$\pi_{13\mid2}$	0.100	0.059	0.098	0.132
F_{12}	0.100	0.081	0.099	0.118	$\pi_{21\mid2}$	0.100	0.067	0.103	0.145
F_{22}	0.200	0.175	0.198	0.225	$\pi_{22\mid2}$	0.800	0.736	0.794	0.853
F_{32}	0.300	0.251	0.298	0.337	$\pi_{23\mid2}$	0.100	0.065	0.103	0.141
F_{13}	0.100	0.084	0.100	0.122	$\pi_{31\mid2}$	0.100	0.060	0.104	0.149
F_{23}	0.200	0.176	0.202	0.229	$\pi_{32\mid2}$	0.100	0.063	0.099	0.134
F_{33}	0.300	0.260	0.299	0.355	$\pi_{33\mid2}$	0.800	0.738	0.797	0.855
$\pi_{11\mid1}$	0.900	0.865	0.903	0.949	$\log_{10} P_1$	5.000	4.877	4.995	5.104
$\pi_{12\mid1}$	0.050	0.019	0.048	0.076	$\log_{10} P_2$	5.000	4.872	4.990	5.099
$\pi_{13\mid1}$	0.050	0.022	0.049	0.082	$\log_{10} P_3$	5.000	4.876	4.991	5.092
$\pi_{21\mid1}$	0.050	0.021	0.047	0.078	F_1	0.100	0.091	0.100	0.110
$\pi_{22\mid1}$	0.900	0.863	0.905	0.939	F_2	0.200	0.185	0.200	0.219
$\pi_{23\mid1}$	0.050	0.020	0.048	0.084	F_3	0.300	0.265	0.297	0.339
$\pi_{31\mid1}$	0.050	0.013	0.051	0.085					

The mean, 2.5% and 97.5% quantiles of the parameter estimates are very similar to those in Table 3, which contains analogous results except using 100% conventional tags.

As explained previously, we did not expect including 5% archival tags to have much impact on the results. Thus, we repeated the exercise with 5%, 50% and 95% of the total tag releases being archival. We increased the number of replicates to 250 in each case instead of 100 to ensure that any differences are real and not just due to insufficient sample sizes. We do not expect the means of the parameter estimates to differ significantly (because they are already unbiased using only conventional tags); instead, we are interested in whether the precision of any of the estimates is improved. Table 26 presents a comparison of the coefficients of variation of the parameter estimates from the three scenarios.

Table 26: Coefficients of variation ($\times 100\%$) of parameter estimates from 250 simulations using 1) 5% archival tags; 2) 50% archival tags; and 3) 95% archival tags. True parameter values are also given.

Par	True	Percent archival 5%	50%	95%	Par	True	Percent archival 5%	50%	95%
M_1	0.3	12.6	12.5	12.6	$\pi_{32\|1}$	0.05	36.9	29.2	24.5
M_2	0.2	24.8	25.5	26.6	$\pi_{33\|1}$	0.9	2.8	2.3	1.9
F_{11}	0.1	8.7	7.8	9.5	$\pi_{11\|2}$	0.8	3.6	3.1	3.3
F_{21}	0.2	6.9	6.2	6.2	$\pi_{12\|2}$	0.1	20.2	18.0	19.3
F_{31}	0.3	8.2	8.8	8.8	$\pi_{13\|2}$	0.1	20.8	18.4	18.6
F_{12}	0.1	8.6	9.0	9.2	$\pi_{21\|2}$	0.1	20.4	19.5	18.0
F_{22}	0.2	6.2	6.0	6.4	$\pi_{22\|2}$	0.8	3.6	3.4	3.0
F_{32}	0.3	8.0	8.3	9.2	$\pi_{23\|2}$	0.1	21.6	19.3	19.0
F_{13}	0.1	9.1	9.3	8.9	$\pi_{31\|2}$	0.1	21.5	20.6	19.0
F_{23}	0.2	6.7	5.3	6.1	$\pi_{32\|2}$	0.1	20.2	19.4	18.4
F_{33}	0.3	8.7	8.5	8.2	$\pi_{33\|2}$	0.8	3.5	3.3	3.1
$\pi_{11\|1}$	0.9	2.2	2.1	1.9	$\log_{10} P_1$	5	1.1	1.2	1.2
$\pi_{12\|1}$	0.05	33.6	27.8	23.7	$\log_{10} P_2$	5	1.3	1.2	1.2
$\pi_{13\|1}$	0.05	31.4	29.5	25.8	$\log_{10} P_3$	5	1.2	1.2	1.3
$\pi_{21\|1}$	0.05	36.6	25.5	24.8	F_1	0.1	5.1	5.4	5.4
$\pi_{22\|1}$	0.9	2.7	2.3	1.9	F_2	0.2	4.1	3.7	4.1
$\pi_{23\|1}$	0.05	33.4	30.0	26.0	F_3	0.3	6.6	7.1	7.3
$\pi_{31\|1}$	0.05	33.9	31.0	24.3					

We see that increasing the percent of archival tags leads to improvements in precision

of all the transition probabilities, and significant improvements for some of them. Somewhat surprisingly, the precision of the other parameter estimates is largely unaffected.

While these findings are specific to the situation being considered, we carried out simulations for several other situations and reached much the same conclusion. Our investigation was not exhaustive, however, and there are perhaps situations in which archival tags could improve the mortality and abundance estimates as well as the transition rates.

7 Model 6: allowing for overdispersion in the tag-recapture data in the generic spatial Brownie model

Tag-recapture data are commonly modelled as multinomial, although in practice they exhibit extra-multinomial variation. In the context of binomial data, the seminal paper is Williams (1982). He, and subsequent authors, have discussed a variety of models for overdispersed binomial data. The extra variation has a variety of causes: fish swim in schools; fishing is a commercial enterprise; weather can affect spotting ability; and so on. In short, fishing is not an exercise in random sampling.

We do not have many degrees of freedom to model overdispersion. Consider just the tag-recapture data in the generic base model — that is, ignore the catch-at-age data. We are estimating 23 independent parameters from 36 independent data values. We cannot afford to use many degrees of freedom on overdispersion. We therefore choose the most parsimonious and tractable model we can for the overdispersion.

7.1 The Dirichlet-multinomial distribution

The Dirichlet-multinomial distribution appears to be due to Mosimann (1962), as a model for multinomial style data with overdispersion.

The multinomial distribution describes n trials, in which an observation in any given trial falls into one of k categories. The probability that observation falls into category i is p_i, and the probability of data $n_1, n_2, \ldots, n_k$ for the category counts over the n trials is

$$\left(\begin{array}{c} n \\ n_1, n_2, \ldots, n_k \end{array} \right) p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k} \,.$$

Note that $\sum n_i = n$. We will denote this distribution as $\Pr\{\underline{n}|\underline{p}\}$, where $\underline{v} = (v_1, v_2, \ldots, v_k)'$ is generic notation for a vector.

The Dirichlet distribution has been invented to describe the variation of proportions that add to 1. The proportions are presumed to vary randomly, and the probability (density) of a set of proportions $p_1, p_2, \ldots, p_k$ is

$$\frac{\Gamma(a)}{\prod_{i=1}^{k} \Gamma(a_i)} \prod_{i=1}^{k} p_i^{a_i - 1} \,,$$

where $a_i > 0$ for all i. Here $\Gamma(x)$ denotes the Gamma function, with the well-known property that $\Gamma(n) = (n-1)!$ when n is an integer. Also $a = \sum_{i=1}^{k} a_i$. Denote the parameters of the Dirichlet distribution by $\underline{a} = (a_1, a_2, \ldots, a_k)'$. We denote this distribution by $\Pr\{\underline{p}|\underline{a}\}$. If we let P_i be the random variable corresponding to the proportion in category i, then it is easy to show that $\mathrm{E}[P_i] = a_i/a$. If $k = 2$, the Dirichlet distribution becomes the beta distribution.

The Dirichlet-multinomial distribution is the compound distribution

$$\Pr\{\underline{n}\} = \int \Pr\{\underline{n}|\underline{p}\} \, \Pr\{\underline{p}|\underline{a}\} \, d\underline{p} \,.$$

The integral is over all values $\underline{p}$ on the k-dimensional simplex, that is, over all values $\underline{p}$ such that $p_i \geq 0$ and $\sum_{i=1}^{k} p_i = 1$. It is easy to show that the resulting distribution is

$$\left(\begin{array}{c} n \\ n_1, n_2, \ldots, n_k \end{array} \right) \frac{\Gamma(a) \left[\prod_{i=1}^{k} \Gamma(n_i + a_i) \right]}{\left[\prod_{i=1}^{k} \Gamma(a_i) \right] \Gamma(n + a)} \,.$$

Part of the appeal of this distribution is its explicit and easily-computed functional form. It is easy to fit the parameters by maximising the log-likelihood, for example.

The Dirichlet-multinomial distribution generalises the beta-binomial distribution, which is often used in statistical circles for analysing overdispersed binomial data.

It is important to emphasise here exactly what this model means. For a given set of n trials (in our case, each trial equates to a tagged fish), the probabilities of occurrence $\underline{p}$ are a fixed set of numbers. However, if we were to take a second set of n trials (i.e. a second set of n tagged fish), the $\underline{p}$ would vary. Here n is the same in each set. Note that $\underline{p}$ does *not* vary randomly from trial to trial within a given set. If it did, the probability of the data would be

$$\prod_{j=1}^{n} \frac{\Gamma(a) \left[\prod_{i=1}^{k} \Gamma(n_{ij} + a_i)\right]}{\left[\prod_{i=1}^{k} \Gamma(a_i)\right] \Gamma(1+a)} = \frac{\Gamma(a)^n \left[\prod_{j=1}^{n} \prod_{i=1}^{k} \Gamma(n_{ij} + a_i)\right]}{\left[\prod_{i=1}^{k} \Gamma(a_i)\right]^n \Gamma(1+a)^n},$$

where $n_{ij} = 1$ if individual j is in group i, and is 0 otherwise. This is not a realistic model: it would be hard to capture the wholesale deviations from the multinomial probabilities that operate on a group of individuals.

Let N_i be the random variable corresponding to the count in category i in a Dirichlet-multinomial distribution. It is easy to show that

$$\begin{aligned}
\mathrm{E}[N_i] &= n\frac{a_i}{a} = n\mathrm{E}[P_i]\,; \\
\mathrm{Var}[N_i] &= n\mathrm{E}[P_i](1 - \mathrm{E}[P_i])\left[\frac{n+a}{1+a}\right]\,; \\
\mathrm{Cov}[N_i, N_j] &= -n\mathrm{E}[P_i]\mathrm{E}[P_j]\left[\frac{n+a}{1+a}\right]\,.
\end{aligned}$$

The variance-covariance matrix of the Dirichlet-multinomial distribution is proportional to that of a multinomial distribution — the scaling factor is $(n+a)/(1+a)$. The overdispersion is greatest for small a, and reduces to 0 as $a \to \infty$.

If we let $\mu_i = \mathrm{E}[P_i]$, then we can write the Dirichlet-multinomial distribution as

$$\begin{pmatrix} n \\ n_1, n_2, \ldots, n_k \end{pmatrix} \frac{\Gamma(a) \left[\prod_{i=1}^{k} \Gamma(n_i + a\mu_i)\right]}{\left[\prod_{i=1}^{k} \Gamma(a\mu_i)\right] \Gamma(n+a)},$$

where μ_i (which play the role of the p_i in the multinomial distribution) will be modelled in terms of natural and fishing mortalities and transition probabilities, and a represents overdispersion. Note that since $\sum_{i=1}^{k} \mu_i = 1$, $a = \sum_{i=1}^{k} a\mu_i$.

It is useful to develop an intuitive understanding of the parameters. Consider the case $E[P_i] = 1/2$. For the multinomial case, the coefficient of variation of N_i is $1/\sqrt{n}$, but for the Dirichlet-multinomial it is $\sqrt{(n+a)/(1+a)}/\sqrt{n}$. For n large (> 1000, say), the CV of the multinomial is only a few per cent. This seems unlikely to be true for real data, and we assume that the CV of the Dirichlet-multinomial is $\gg 1/\sqrt{n}$. If we solve for a for a given n and CV, we get $a = n(1-\text{CV}^2)/(n\text{CV}^2-1)$. If we assume $n\text{CV}^2 \gg 1$, then $a \approx (1-\text{CV}^2)/\text{CV}^2$ or $\text{CV} = 1/\sqrt{1+a}$. This means that the parameter a has a simple interpretation. An equivalent way to look at this is to assume that $a \ll n$ in practice.

This suggests a sensible value to use for a in practice. We would often want $CV = 0.1$ for the tagging data. This equates to $a = 99$ exactly, so we might set $a = 101$. It may, in fact, be better to reparameterise, so that we use $1/\sqrt{1+a}$ rather than a itself.

7.2   Simulating Dirichlet-multinomial data

Probably the simplest way to simulate Dirichlet-multinomial data is to first simulate $X_1, \ldots, X_k$ as n independent Gamma variates, in which the density of X_i is

$$\frac{e^{-x}x^{a_i-1}}{\Gamma(a_i)}$$

(where $a_i = a\mu_i$) and then set $p_i = X_i/(X_1 + \cdots + X_k)$. Then simulate multinomial data from the multinomial distribution with proportions $(p_1, \ldots, p_k)$. Note that if $\mu_i = 0$, as might happen in the context of a real problem, we simply set $X_i = 0$.

For a statement of the result that the p_i have a Dirichlet distribution, see Aitchison (1986), Property 3.3.

7.3   Tagging data

Suppose, as in Section 2.1, we consider a particular tagging group that generates a corresponding vector of multinomial probabilities

$$q = (\bar{q}, q_{11}, q_{12}, \ldots, q_{1R}, q_{21}, q_{21}, \ldots, q_{2R}, \ldots, q_{T1}, q_{T2}, \ldots, q_{TR}).$$

Write this as

$$q = (\mu_1, \mu_2, \ldots, \mu_{1+TR}).$$

Write the counts as

$$(n_1, n_2, \ldots, n_{1+TR}),$$

and $n = \sum_{i=1}^{1+TR} n_i$. If we assume that the counts are a realisation of a Dirichlet-multinomial distribution, then their probability is

$$\binom{n}{n_1, n_2, \ldots, n_{1+TR}} \frac{\Gamma(a) \left[\prod_{i=1}^{1+TR} \Gamma(n_i + a\mu_i)\right]}{\left[\prod_{i=1}^{1+TR} \Gamma(a\mu_i)\right] \Gamma(n + a)}, \tag{10}$$

where a is an overdispersion parameter.

7.4 Inference

The log-likelihood of the tag-recapture data is, apart from an additive constant,

$$\log \Gamma(a) + \sum_{i=1}^{1+TR} \left(\log \Gamma(n_i + a\mu_i) - \log \Gamma(a\mu_i)\right) - \log \Gamma(n + a).$$

Most programming environments provide a $\log \Gamma(x)$ function, so it is easy to compute the likelihood for a given set of parameter values. Note that for a structural zero, $n_i = \mu_i = 0$. But $\log \Gamma(0) = \infty$, and we strike numerical problems. Unfortunately, we cannot use a similar trick to that described in Section 2.3. Note that if $n_i = 0$, we have $\log \Gamma(n_i + a\mu_i) - \log \Gamma(a\mu_i) = 0$, whether $\mu_i = 0$ or not, so the simplest way around the problem is simply to omit the data with $n_i = 0$ from the log-likelihood.

Abramowitz and Stegun (1964), paragraph 6.1.46 on page 257, claim that

$$\lim_{x \to \infty} x^{b-a} \frac{\Gamma(x + a)}{\Gamma(x + b)} = 1. \tag{11}$$

This suggests that we should group the log-likelihood into terms. Thus, we write the log-likelihood as

$$[n \log a + \log \Gamma(a) - \log \Gamma(n + a)]$$
$$+ \sum_{i=1}^{1+TR} [-n_i \log(a\mu_i) + \log \Gamma(n_i + a\mu_i) - \log \Gamma(a\mu_i)]$$
$$+ \sum_{i=1}^{1+TR} n_i \log \mu_i. \tag{12}$$

From (11) the first two lines of (12) each $\to 0$ as $a \to \infty$, and hence log-likelihood reduces to the multinomial likelihood (the final line of (12)) as $a \to \infty$.

7.5 Numerical example

The data have the same structure as the base case, so we do not present any examples of data sets. Note, however, that the number of tags released are the same as presented in Table 1, namely 1415 for the one-year-olds, 949 for the two-year-olds and 636 for the three-year-olds.

7.5.1 Low overdispersion

We commence by simulating 100 data sets with the true values presented in Table 27. The true value of a is 9999. In fact, we prefer to reparameterise a as $1/\sqrt{1+a}$, which equals 0.01 when $a = 9999$. If $n = 1000$, the variance-inflation factor over the multinomial is $(1000 + 9999)/(9999 + 1) \approx 1.01$, or, in terms of standard deviations, about 5% overdispersion. If $n = 500$, it is about 2.5% overdispersion, and if $n = 1500$, about 7.2%. This therefore represents quite modest overdispersion. The summary statistics for the maximum likelihood estimators are presented in Table 27.

Table 27: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of minimally overdispersed tag-recapture and catch-at-age data generated from a model in which there is tagging of a single cohort of fish in all regions for three years

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.219	0.287	0.350	$\pi_{32\mid1}$	0.050	0.015	0.049	0.085
M_2	0.200	0.077	0.188	0.295	$\pi_{33\mid1}$	0.900	0.857	0.905	0.956
F_{11}	0.100	0.083	0.099	0.117	$\pi_{11\mid2}$	0.800	0.736	0.782	0.853
F_{21}	0.200	0.172	0.195	0.223	$\pi_{12\mid2}$	0.100	0.058	0.102	0.144
F_{31}	0.300	0.249	0.299	0.353	$\pi_{13\mid2}$	0.100	0.072	0.115	0.165
F_{12}	0.100	0.082	0.101	0.113	$\pi_{21\mid2}$	0.100	0.065	0.100	0.136
F_{22}	0.200	0.175	0.197	0.224	$\pi_{22\mid2}$	0.800	0.748	0.800	0.859

Table 27: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
F_{32}	0.300	0.260	0.300	0.361	$\pi_{23\|2}$	0.100	0.059	0.100	0.143
F_{13}	0.100	0.083	0.099	0.116	$\pi_{31\|2}$	0.100	0.064	0.098	0.144
F_{23}	0.200	0.168	0.190	0.214	$\pi_{32\|2}$	0.100	0.054	0.096	0.136
F_{33}	0.300	0.239	0.289	0.334	$\pi_{33\|2}$	0.800	0.757	0.805	0.856
$\pi_{11\|1}$	0.900	0.856	0.898	0.944	$1/\sqrt{1+a}$	0.010	0.000	0.000	0.000
$\pi_{12\|1}$	0.050	0.026	0.053	0.088	$\log_{10} P_1$	5.000	4.866	4.986	5.078
$\pi_{13\|1}$	0.050	0.012	0.049	0.081	$\log_{10} P_2$	5.000	4.832	4.993	5.091
$\pi_{21\|1}$	0.050	0.021	0.047	0.073	$\log_{10} P_3$	5.000	4.835	4.987	5.115
$\pi_{22\|1}$	0.900	0.837	0.887	0.941	F_1	0.100	0.088	0.099	0.109
$\pi_{23\|1}$	0.050	0.029	0.066	0.102	F_2	0.200	0.177	0.194	0.208
$\pi_{31\|1}$	0.050	0.019	0.046	0.071	F_3	0.300	0.255	0.296	0.336

The parameter estimates appear to be reasonable, although some exhibit small apparent biases, which may or may not be real. An interesting feature is that $1/\sqrt{1+a} = 0$ (corresponding to $a = \infty$) for all 100 trials — this makes it clear why we reparameterised a. Essentially the small amount of overdispersion is lost.

7.5.2 Medium overdispersion

We now consider a situation with increased overdispersion. We set $a = 999$, equivalent to $1/\sqrt{1+a} = 0.032$. For $n = 500$, the overdispersion is 22%, for $n = 1000$ it is 41%, and for $n = 1500$ it is 58%. The results are presented in Table 28.

Table 28: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of moderately overdispersed tag-recapture and catch-at-age data generated from a model in which there is tagging of a single cohort of fish in all regions for three years

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.220	0.286	0.347	$\pi_{32\|1}$	0.050	0.009	0.040	0.073
M_2	0.200	0.075	0.180	0.279	$\pi_{33\|1}$	0.900	0.822	0.866	0.914
F_{11}	0.100	0.084	0.100	0.117	$\pi_{11\|2}$	0.800	0.752	0.796	0.853
F_{21}	0.200	0.167	0.190	0.215	$\pi_{12\|2}$	0.100	0.068	0.112	0.152

Table 28: continued

Par	True	2.5%	Mean	97.5%	Par		True	2.5%	Mean	97.5%
F_{31}	0.300	0.247	0.295	0.349	$\pi_{13\|2}$		0.100	0.055	0.092	0.125
F_{12}	0.100	0.088	0.105	0.119	$\pi_{21\|2}$		0.100	0.060	0.095	0.139
F_{22}	0.200	0.173	0.197	0.225	$\pi_{22\|2}$		0.800	0.751	0.813	0.884
F_{32}	0.300	0.264	0.309	0.366	$\pi_{23\|2}$		0.100	0.047	0.092	0.136
F_{13}	0.100	0.081	0.097	0.113	$\pi_{31\|2}$		0.100	0.040	0.080	0.119
F_{23}	0.200	0.163	0.179	0.200	$\pi_{32\|2}$		0.100	0.065	0.106	0.152
F_{33}	0.300	0.236	0.283	0.327	$\pi_{33\|2}$		0.800	0.762	0.814	0.868
$\pi_{11\|1}$	0.900	0.850	0.888	0.929	$1/\sqrt{1+a}$		0.032	0.000	0.005	0.022
$\pi_{12\|1}$	0.050	0.046	0.080	0.117	$\log_{10} P_1$		5.000	4.855	4.977	5.076
$\pi_{13\|1}$	0.050	0.009	0.032	0.063	$\log_{10} P_2$		5.000	4.802	4.976	5.074
$\pi_{21\|1}$	0.050	0.022	0.047	0.078	$\log_{10} P_3$		5.000	4.875	5.011	5.143
$\pi_{22\|1}$	0.900	0.810	0.869	0.927	F_1		0.100	0.090	0.100	0.109
$\pi_{23\|1}$	0.050	0.042	0.084	0.132	F_2		0.200	0.174	0.188	0.201
$\pi_{31\|1}$	0.050	0.055	0.094	0.135	F_3		0.300	0.254	0.295	0.336

Some modest biases are observed in many of the parameter estimates.

7.5.3 High overdispersion

We now try an extreme case with very high overdispersion. We set $1/\sqrt{1+a} = 0.05$. For $n = 500$, the overdispersion is 50%, for $n = 1000$ it is 87%, and for $n = 1500$ it is 118%. The results are presented in Table 29.

Table 29: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of highly overdispersed tag-recapture and catch-at-age data generated from a model in which there is tagging of a single cohort of fish in all regions for three years

Par	True	2.5%	Mean	97.5%	Par		True	2.5%	Mean	97.5%
M_1	0.300	0.188	0.258	0.322	$\pi_{32\|1}$		0.050	0.079	0.124	0.176
M_2	0.200	0.146	0.244	0.345	$\pi_{33\|1}$		0.900	0.784	0.846	0.897
F_{11}	0.100	0.046	0.059	0.075	$\pi_{11\|2}$		0.800	0.721	0.766	0.823
F_{21}	0.200	0.157	0.181	0.205	$\pi_{12\|2}$		0.100	0.095	0.135	0.175
F_{31}	0.300	0.273	0.321	0.380	$\pi_{13\|2}$		0.100	0.062	0.100	0.140
F_{12}	0.100	0.091	0.109	0.123	$\pi_{21\|2}$		0.100	0.051	0.085	0.120

Table 29: continued

Par	True	2.5%	Mean	97.5%	Par		True	2.5%	Mean	97.5%
F_{22}	0.200	0.165	0.186	0.211	$\pi_{22\|2}$		0.800	0.718	0.782	0.852
F_{32}	0.300	0.298	0.351	0.402	$\pi_{23\|2}$		0.100	0.082	0.134	0.182
F_{13}	0.100	0.080	0.094	0.113	$\pi_{31\|2}$		0.100	0.078	0.112	0.156
F_{23}	0.200	0.161	0.184	0.211	$\pi_{32\|2}$		0.100	0.030	0.060	0.106
F_{33}	0.300	0.205	0.252	0.291	$\pi_{33\|2}$		0.800	0.775	0.828	0.873
$\pi_{11\|1}$	0.900	0.879	0.912	0.949	$1/\sqrt{1+a}$		0.050	0.000	0.021	0.036
$\pi_{12\|1}$	0.050	0.000	0.022	0.045	$\log_{10} P_1$		5.000	4.987	5.099	5.186
$\pi_{13\|1}$	0.050	0.038	0.065	0.093	$\log_{10} P_2$		5.000	4.757	4.941	5.041
$\pi_{21\|1}$	0.050	0.019	0.045	0.073	$\log_{10} P_3$		5.000	4.887	5.020	5.140
$\pi_{22\|1}$	0.900	0.876	0.931	0.961	F_1		0.100	0.074	0.084	0.094
$\pi_{23\|1}$	0.050	0.000	0.024	0.057	F_2		0.200	0.168	0.183	0.198
$\pi_{31\|1}$	0.050	0.010	0.031	0.059	F_3		0.300	0.266	0.306	0.343

We now see considerable biases in most parameter estimates, particularly in the natural mortalities. Oddly enough, the population sizes are estimated reasonably well. Even in this case $1/\sqrt{1+a}$ is estimated to be 0 occasionally. Overall the maximum likelihood estimator of this parameter is very biased, with a mean of 0.021 and a 97.5 percentile of 0.036, well below the true value of 0.05.

This last point raised the question of whether the estimation would improve if we set $1/\sqrt{1+a}$ at its true value of 0.05. The results are set out in Table 30.

Table 30: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of highly overdispersed tag-recapture and catch-at-age data generated from a model in which there is tagging of a single cohort of fish in all regions for three years, fixing the overdispersion parameter, $1/\sqrt{1+a}$, at its true value

Par	True	2.5%	Mean	97.5%	Par		True	2.5%	Mean	97.5%
M_1	0.300	0.184	0.259	0.321	$\pi_{32\|1}$		0.050	0.082	0.125	0.177
M_2	0.200	0.147	0.240	0.343	$\pi_{33\|1}$		0.900	0.779	0.838	0.893
F_{11}	0.100	0.053	0.067	0.085	$\pi_{11\|2}$		0.800	0.713	0.757	0.809
F_{21}	0.200	0.154	0.179	0.202	$\pi_{12\|2}$		0.100	0.097	0.138	0.180
F_{31}	0.300	0.272	0.318	0.375	$\pi_{13\|2}$		0.100	0.067	0.105	0.144

Table 30: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
F_{12}	0.100	0.091	0.110	0.130	$\pi_{21|2}$	0.100	0.056	0.089	0.124
F_{22}	0.200	0.164	0.189	0.212	$\pi_{22|2}$	0.800	0.710	0.780	0.846
F_{32}	0.300	0.301	0.351	0.403	$\pi_{23|2}$	0.100	0.078	0.131	0.182
F_{13}	0.100	0.076	0.093	0.116	$\pi_{31|2}$	0.100	0.081	0.115	0.160
F_{23}	0.200	0.157	0.186	0.218	$\pi_{32|2}$	0.100	0.032	0.061	0.102
F_{33}	0.300	0.209	0.255	0.293	$\pi_{33|2}$	0.800	0.772	0.824	0.868
$\pi_{11|1}$	0.900	0.864	0.899	0.939	$1/\sqrt{1+a}$	0.050	0.050	0.050	0.050
$\pi_{12|1}$	0.050	0.000	0.030	0.053	$\log_{10} P_1$	5.000	4.955	5.067	5.158
$\pi_{13|1}$	0.050	0.046	0.071	0.098	$\log_{10} P_2$	5.000	4.753	4.936	5.036
$\pi_{21|1}$	0.050	0.025	0.050	0.078	$\log_{10} P_3$	5.000	4.876	5.013	5.137
$\pi_{22|1}$	0.900	0.873	0.919	0.949	F_1	0.100	0.075	0.088	0.102
$\pi_{23|1}$	0.050	0.000	0.031	0.063	F_2	0.200	0.169	0.184	0.200
$\pi_{31|1}$	0.050	0.016	0.037	0.064	F_3	0.300	0.267	0.306	0.343

It is clear that setting a at its true value results in negligible improvement.

Finally we estimated the parameters assuming that the tag-recapture data were multinomial — that is, we ignored the overdispersion. The results are displayed in Table 31.

Table 31: True parameter values, mean maximum likelihood estimates and summary statistics for 100 sets of highly overdispersed tag-recapture and catch-at-age data generated from a model in which there is tagging of a single cohort of fish in all regions for three years, assuming multinomial recapture data (i.e., assuming no overdispersion)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.187	0.254	0.319	$\pi_{32|1}$	0.050	0.079	0.125	0.174
M_2	0.200	0.143	0.244	0.345	$\pi_{33|1}$	0.900	0.786	0.846	0.899
F_{11}	0.100	0.044	0.056	0.068	$\pi_{11|2}$	0.800	0.721	0.767	0.825
F_{21}	0.200	0.160	0.184	0.208	$\pi_{12|2}$	0.100	0.093	0.134	0.174
F_{31}	0.300	0.276	0.322	0.381	$\pi_{13|2}$	0.100	0.061	0.099	0.139
F_{12}	0.100	0.090	0.108	0.120	$\pi_{21|2}$	0.100	0.050	0.083	0.118
F_{22}	0.200	0.164	0.186	0.211	$\pi_{22|2}$	0.800	0.712	0.777	0.843
F_{32}	0.300	0.297	0.351	0.404	$\pi_{23|2}$	0.100	0.093	0.140	0.185
F_{13}	0.100	0.081	0.094	0.113	$\pi_{31|2}$	0.100	0.077	0.110	0.151

Table 31: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
F_{23}	0.200	0.161	0.182	0.209	$\pi_{32\vert 2}$	0.100	0.031	0.063	0.110
F_{33}	0.300	0.203	0.250	0.289	$\pi_{33\vert 2}$	0.800	0.767	0.826	0.875
$\pi_{11\vert 1}$	0.900	0.879	0.914	0.952	$1/\sqrt{1+a}$	0.050	0.000	0.000	0.000
$\pi_{12\vert 1}$	0.050	0.000	0.020	0.039	$\log_{10} P_1$	5.000	4.998	5.114	5.216
$\pi_{13\vert 1}$	0.050	0.037	0.066	0.096	$\log_{10} P_2$	5.000	4.756	4.941	5.040
$\pi_{21\vert 1}$	0.050	0.017	0.044	0.073	$\log_{10} P_3$	5.000	4.882	5.020	5.140
$\pi_{22\vert 1}$	0.900	0.878	0.935	0.965	F_1	0.100	0.073	0.082	0.093
$\pi_{23\vert 1}$	0.050	0.000	0.022	0.055	F_2	0.200	0.167	0.184	0.198
$\pi_{31\vert 1}$	0.050	0.007	0.029	0.057	F_3	0.300	0.265	0.305	0.344

This also failed to improve matters.

7.6 Overdispersion: concluding remarks

The Dirichlet-multinomial model is capable of modelling overdispersion. However, the degree of overdispersion is always underestimated by maximum likelihood. As the degree of overdispersion increases, substantial biases are seen in some parameter estimators. Remarkably, the population size estimates seem to be fairly reasonable, even for high overdispersion cases.

If we focus on $\hat{M}_1$, we find that in Tables 3, 27, 28, and 29 the width of the (2.5%,97.5%) interval of estimates remains fairly constant at about 0.13. However, between the medium and high overdispersion cases (Tables 28 and 29), the mean suddenly changes from about 0.29 to about 0.26 (recall the true value is 0.30). It would be interesting to see if this behaviour is maintained if the level of overdispersion is estimated in a less biased way.

References

AITCHISON, J. (1986) *The statistical analysis of compositional data.* Chapman and Hall: London.

CATCHPOLE, E.A. & MORGAN, B.J.T. (1997) Detecting parameter redundancy. *Biometrika,* **84,** 187-196.

MOSIMANN, J.E. (1962) On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika,* **49,** 65-82.

WILLIAMS, D.A. (1982) Extra-binomial variation in logistic linear models. *Applied Statistics,* **31,** 144-148.

Appendix 11:

Incorporating spatial structure into an integrated tag-recapture and catch model using movement dynamics representative of southern bluefin tuna

Geoff M. Laslett, Tom Polacheck and J. Paige Eveson

1   Introduction

In Appendix 10, we expanded the integrated multi-year tagging and catch-at-age models developed in Appendices 5 and 7 to incorporate spatial structure. The spatial structure allowed for fish abundance and fishing mortality to differ among defined regions and for fish to move among these regions, without any restrictions on the movement patterns (i.e., transition probabilities) between regions. In the current appendix, instead of using such a generic spatial framework, we develop two spatial models that are intended to approximate the spatial/movement dynamics and fishery dynamics of southern bluefin tuna (SBT).

We define four main fishing regions in which juvenile (immature) SBT are caught:

1. Great Australian Bight (GAB)

2. South African fishery

3. South-East Indian Ocean

4. Tasman Sea

The GAB fishery consists of purse seine vessels catching juveniles for transportion into tuna farms, and is generally referred to as the surface fishery. The other three areas are composed of various longline fisheries, dominated by Japanese vessels but also including Taiwanese, Korean, South African and New Zealand vessels.

Fishing generally occurs in these regions in two distinct seasons:

1. Austral summer for the GAB surface fishery

2. Austral winter for the longline fisheries

Juvenile SBT can be found in abundance in the GAB during the austral summer up until around age 5 (few fish older than age 5 are found in the GAB). The surface fishery, operating out of Port Lincoln on the eastern extreme of the GAB, is at its

peak from December to April, although the fishing season can extend outside these months. By late autumn, essentially all tuna have left the GAB and migrated to one of the other three regions. The Tasman Sea, South-East Indian Ocean and the South African fisheries generally operate in the austral winter, even though tuna may be found in these areas in the austral summer.

In both the SBT spatial models described in this appendix, we divide time into $t = 1, 2, \ldots, T$ time periods, and space into $r = 1, 2, \ldots, R$ regions. Here time period t refers to the interval $(t - 1, t]$. There are two periods per year, one for each season. There are $R = 4$ regions, corresponding to the four fishing areas. Thus the model captures the structure of the SBT fishery in broad detail. For computational convenience, we shall impose a slightly artificial structure on the fishery. During time period t (that is, during the period $(t - 1, t)$), the fish in a region stay within the region, but they may be caught or die naturally. Exactly at the end of each time period (that is, at time t exactly), the fish will move between regions, according to the rules of the model. This model conveniently separates the fishing and natural mortality processes from the movement processes.

In the first spatial model, we allow for Markovian (i.e. memoryless) movements between regions at the end of each season. In the second model, the movements between regions are not Markovian. Rather, each fish adopts either the Tasman Sea, the South-East Indian Ocean or the South African region as its winter base. It either stays there for the summer or moves to the GAB. In the autumn it remains in or returns to its winter base. Thus, the movement of each fish is described by a two-state Markov chain, with the GAB and the winter region as the two states. The overall transition model for a randomly chosen fish at time 0 is not Markovian because each fish remembers its winter base. The regional transition model is a superposition of three Markov chains.

The generic model described in Appendix 10 is different from either of these models. In the generic model, a fish remains in each region for a year. In the SBT models, a fish remains in each region for a season. The generic model and the

first SBT model can be equated if generic years are regarded as SBT seasons and suitable constraints are imposed on the generic transition probabilities. The second SBT model is not a special case of the generic model.

There is a second important difference between the generic and SBT models, which we introduce to reflect actual fishing practice. In the generic model we assume that some vessels in all regions carry observers, and that all tags are returned from such boats. This allows us to estimate reporting rates for vessels without observers. We carry this feature over to the winter fisheries in the SBT models, but for the summer surface fishery we assume instead that tag seeding data are available for estimating reporting rates.

Ultimately even the first SBT spatial model and the generic model are quite different and are probably best regarded as separate models.

An outline of the current appendix is as follows. We first describe the population dynamics equations underlying the two models. We then present the first SBT spatial model, starting with a description of the Markov movement dynamics, then describing the data sets and corresponding likelihood components that make up the integrated tag-recapture and catch-at-age model. The data used in both SBT models are of the same three types: tag-recapture data, tag seeding data and catch-at-age data. For convenience, we describe the data types in terms of the first SBT model. Next, we present the second SBT spatial model (the site-fidelity model) following a similar outline as for the first model.

For a given cohort of tagged fish, both models provide age-specific estimates of natural mortality, age- and region-specific estimates of fishing mortality and region-specific estimates of initial abundance. Often, however, it is not the regional parameter estimates of fishing mortality and abundance that are of primary interest, but rather the population-wide estimates (i.e., over all regions). Such population-wide estimates can be derived from the regional estimates, as will be discussed. They can also be obtained using a non-spatial model applied to the data pooled across all regions; however, we would expect the estimates to be biased if spatial heterogeneity exists. To investigate these biases, we first describe the non-spatial model for

the pooled data (referred to as the pooled model), and then compare estimates of overall fishing mortality and abundance obtained from the pooled model with those obtained from the spatial model. Results are presented for a range of scenarios for both spatial SBT models.

Finally, we investigate what happens if data are generated according to the second (site-fidelity) model but analysed using the first (Markov) model to see whether misspecification of the underlying movement dynamics produces unreliable estimates of the mortality rate and abundance parameters.

2 Survival and non-survival in both SBT models

During a given time period t, the survival, natural mortality and fishing processes in both SBT models are the same. We describe them in this section.

The probability of survival of a fish during a time period, conditional on its being alive at the beginning of the period, depends on both the natural and fishing mortalities. For a given cohort, we assume that natural mortality depends only on age: M_a. We assume that the natural mortality is equally divided between the summer season and the winter season, so that the natural mortality in each season is $M_a/2$. We assume that the fishing mortality of a fish of age a in region r in season s is F_{asr}, where $r = 1, \ldots, R$. For the SBT fishery with $R = 4$, the F_{asr} values will have the structure:

$$F_{a1} = \begin{pmatrix} F_{a11} \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad F_{a2} = \begin{pmatrix} 0 \\ F_{a22} \\ F_{a23} \\ F_{a24} \end{pmatrix}$$

Note that fishing only takes place in area 1 in season 1 and in areas 2, 3, 4 in season 2, so the fishing mortality is 0 in the complementary areas.

In the interest of parsimony, it is often desirable to put constraints on the fishing mortality parameters. For example, for SBT, we may assume that $F_{as} = F_{3s}$

for $a > 3$. We do not consider fish older than age 3 in the simulations we present in this appendix, so we do not actually apply this constraint.

We now state explicitly the meaning of M_a and F_{asr}. The probability that a fish of age a, alive in region r at the beginning of season s, survives until the end of the time period, is

$$S_{asr} = e^{-M_a/2 - F_{asr}} \ .$$

The probability that it does not survive, and is caught, is $(1 - S_{asr})f_{asr}$, where

$$f_{asr} = \frac{F_{asr}}{M_a/2 + F_{asr}}$$

is the conditional probability that it is caught rather than dies naturally. Likewise the probability that it does not survive, and dies naturally, is $(1 - S_{asr})(1 - f_{asr})$. This is a type of competing risks model.

Recall that the above statements apply to a particular cohort of fish. When considering multiple cohorts (as in Appendix 16), we assume the natural mortality still depends only on age, but that fishing mortality in each region and season varies not only with age but also year.

3   The first SBT model: Markovian transitions

The two SBT models differ in the way fish move between regions at the end of each season. We now describe the first model. In our presentation of the model and the data we shall focus on a single cohort of fish.

Since there are two seasons per year, it is convenient to work in terms of periods t, taking values $1, 2, 3, 4, \ldots, T$, where there are two periods per year. The fish are tagged at the beginning of each period, in the regions where fishing occurs, starting at age 1 in time period 1. The season is given by $s = t - 2[(t-1)/2]$, where $[x]$ means the integer part of x. Thus, $s = 1$ when t is odd and $s = 2$ when t is even. Year is given by $y = [(t+1)/2]$, and since we are considering a single cohort of fish, we can also express the age of a fish using the same equation, namely $a = [(t+1)/2]$.

Note that we can also write an expression for the time period in terms of a fish's age and the season as $t = 2(a - 1) + s$.

Consider a fish alive in region r_C at the beginning of period t_C (for example, a fish tagged in region r_C at the start of t_C). The fish's age is given by $a_C = [(t_C + 1)/2]$ and the season is $s_{t_C} = t_C - 2[(t_C - 1)/2]$. Let R_t be the region that the fish is in at the start (and during) time period t ($t \geq t_C$), and let $L_t = 1$ if the fish is alive at the start of time interval t. Then the probability that the fish is alive and in region r during time period $t + 1$ is

$$\Pr\{R_{t+1} = r, L_{t+1} = 1 | t_C, r_C\} =$$
$$\sum_{r'=1}^{R} \Pr\{R_t = r', L_t = 1 | t_C, r_C\} e^{-M_a/2 - F_{asr'}} \pi_{r'r|as} , \tag{1}$$

where $\pi_{r'r|as}$ represents the probability of moving from region r' to r at the end of time period t. Looking at the right-hand side of this equation, we read that the fish was alive in region r' at time t with probability $\Pr\{R_t = r', L_t = 1 | t_C, r_C\}$, survived the time period with probability $e^{-M_a/2 - F_{asr'}}$, and then moved from region r' to region r with transition probability $\pi_{r'r|as}$. Note that these transition probabilities are a function of a fish's age $a = a(t)$ and season $s = s(t)$, rather than t directly (this distinction only becomes important when multiple cohorts are being modelled). Furthermore, they do not depend on any of the fish's history prior to time t.

Let Π_{as} denote the matrix of transition probabilities for a fish of age a at the end of season s. We assume the following structure for these matrices ($R = 4$). For season 1:

$$\Pi_{a1} = \begin{pmatrix} 0 & \pi_{12|a1} & \pi_{13|a1} & 1 - \pi_{12|a1} - \pi_{13|a1} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

For season 2:

$$
\Pi_{a2} \;=\; \begin{pmatrix}
1 & 0 & 0 & 0 \\[2mm]
\pi_{21|a2} & 1 - \pi_{21|a2} & 0 & 0 \\[2mm]
\pi_{31|a2} & 0 & 1 - \pi_{31|a2} & 0 \\[2mm]
\pi_{41|a2} & 0 & 0 & 1 - \pi_{41|a2}
\end{pmatrix}
$$

These transition matrices essentially say that at the end of season 1, a fish in region 1 moves to either region 2, 3 or 4, whereas a fish in region 2, 3 or 4 stays in that region. At the end of season 2, a fish stays in its current region or moves to region 1. Note however that a fish could make the following transitions: regions 1, 2, 1, 3 in seasons 1, 2, 1, 2 in time periods 1, 2, 3, 4 respectively. Thus, it need not remain faithful to its winter region. For SBT, we expect the transition probabilities from regions 2 to 4 back to region 1 at the end of season 2 to be very small, if not zero, once a fish reaches age 5.

Suppose S_{as} is the $R \times R$ diagonal matrix with rth diagonal element $S_{asr} = e^{-M_a/2 - F_{asr}}$. Then, the equations given in (1) can be expressed in matrix form as

$$
p_{t+1|t_C r_C} = \Pi'_{as}\, S_{as}\, p_{t|t_C r_C}
$$

where $p_{t|t_C r_C}$ is a vector of length R whose rth entry is the probability that a fish, alive in region r_C at the beginning of time t_C (i.e., season s_C and age a_C), is alive in region r at the beginning of time period t (season s and age a). Of course, the rth entry of $p_{t_C|t_C r_C}$ is 1 for $r = r_C$ and 0 for $r \neq r_C$.

The probability that a fish, alive in region r_C at the beginning of time t_C, is caught during time period t in region r is

$$
q_{tr|t_C r_C} = f_{asr}(1 - S_{asr})p_{tr|t_C r_C} \tag{2}
$$

where $p_{tr|t_C r_C}$ (the rth entry of vector $p_{t|t_C r_C}$ above) is the probability that the fish is alive in region r at the beginning of time period t (season s and age a), $1 - S_{asr}$ is the probability that the fish does not survive the time period, and f_{asr} is the probability that it is caught rather than dies naturally (see previous section).

3.1 Tagging data

3.1.1 Structure of the tagging data

We consider tag-recapture data for a single cohort of fish that is tagged in T consecutive time periods starting at age 1 in time period 1. Tagging occurs at the beginning of each odd time period (i.e., season 1) in region 1, and at the beginning of each even time period (i.e., season 2) in regions $2, \ldots, R$. Tags are subsequently returned from the surface fishery operating in region 1 in season 1, and from the longline fisheries operating in regions $2, \ldots, R$ in season 2.

In commercial fishery situations it is unlikely that all recaptured tags will be returned, so we need to allow for the possibility of unreturned tags in the model. This means that we need to estimate additional reporting rate parameters. For the surface fishery in region 1, we assume that tag seeding data are available for estimating reporting rates (see next section). For the longline fisheries in regions $2, \ldots, R$, we assume that each region has a group of boats with observers on them, and that for this subset of the data, 100% of recaptured tags are returned.

For generality, we suppose that in each region at each time the fishery can be divided into two components. In component $g = 1$ (the component with observers), all tags are returned. In the other component, a proportion λ_{tr2} of tags are returned. For consistency, we set $\lambda_{tr1} = 1$. We also know the proportion of the fishery δ_{trg} fished by each component. In practice, the proportions δ_{trg} are inferred from the size of the catch returned by each component. In theory, these are not known exactly, but we argue that the error attached to the δ's is much smaller than other sources of variation in the data, so we treat the δ_{trg} values as known. Of course

$$\delta_{tr1} + \delta_{tr2} = 1$$

for any t and r.

In time periods $t = 1, 3, 5, \ldots$, corresponding to season 1, fishing only occurs in region 1. Because we assume there are no observers in region 1, we set $\delta_{t11} = 0$ and $\delta_{t12} = 1$. For the other regions, it does not matter what we set δ_{trg} to, because

fishing mortality in these regions in season 1 is 0. For convenience, we arbitrarily set $\delta_{t+1,rg} = \delta_{trg}$ for regions $2, \ldots, R$.

In time periods $t = 2, 4, 6, \ldots$, corresponding to season 2, fishing only occurs in regions $2, \ldots, R$. It does not matter what we set δ_{t1g} to, because fishing mortality in region 1 in season 1 is 0. Again, for convenience we set $\delta_{t+1,1g} = \delta_{t1g}$ (which means we set $\delta_{t11} = 0$ and $\delta_{t12} = 1$ for even as well as odd time periods).

Consider a fish tagged in region r_C at the beginning of period t_C. We first generate the vectors $p_{t|t_C r_C}$, as described in the previous section. Once these have been generated, we can calculate the probability that the tag is returned from component g in region r during time interval t. This is

$$q_{trg|t_C r_C} = \lambda_{trg} \delta_{trg} q_{tr|t_C r_C} \tag{3}$$

where $q_{tr|t_C r_C}$, given by equation (2), is the probability that the fish is caught in region r at time t, δ_{trg} is the probability that the fish is caught by a vessel in component g given that it is caught, and λ_{trg} is the probability that the tag is returned, given that it is caught by component g. The probability that the tag is not returned is

$$\bar{q}_{|t_C r_C} = 1 - \sum_{r=1}^{R} \sum_{t=1}^{T} \sum_{g=1}^{2} q_{trg|t_C r_C} \,.$$

We therefore obtain a set of multinomial probabilities for fish tagged in region r_C at the beginning of period t_C:

$$
\begin{aligned}
q = \ & (\bar{q}, q_{111}, q_{112}, q_{121}, q_{122}, \ldots, q_{1R1}, q_{1R2}, \\
& q_{211}, q_{212}, q_{221}, q_{222}, \ldots, q_{2R1}, q_{2R2}, \\
& \ldots, \\
& q_{T11}, q_{T12}, q_{T21}, q_{T22}, \ldots, q_{TR1}, q_{TR2})
\end{aligned} \tag{4}
$$

Many of these probabilities will be 0. For readability we have dropped the conditional expression $(|t_C r_C)$ in the subscript of all terms.

3.1.2 Inference for the tagging data

Inference for these data is fairly straightforward. We refer to the fish tagged in region r_C at the beginning of period t_C as tagging group (t_C, r_C). If tagging occurs in R regions and T time periods, then if we assume complete years of tagging (T even), there will be $I = RT/2$ tagging groups — $T/2$ corresponding to fish tagged in region 1 in odd time periods and $T/2$ corresponding to fish tagged in each of regions $2, \ldots, R$ in even time periods.

Suppose we put the tag-recapture data into a matrix n of dimension $I \times J$, for which the data in row i are the numbers of recaptures by time period, region and component from tagging group i. The first entry of each row is the number of unrecaptured tags, so that $J = 1 + TR2$. The probabilities are placed into a corresponding matrix Q. Row i of Q is the vector q from (4) for the relevant tagging group. Each row of data is treated independently as multinomial data — this implies that we regard sampling variation as the only significant source of variation in the data.

The log-likelihood is (apart from an additive constant)

$$ l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{g=1}^{2} n_{ijg} \log Q_{ijg} \, . $$

If there are structural zeroes in the data, this method of computing the log-likelihood will fail, because $Q_{ijg} = 0$ for structural zeroes, and $\log Q_{ijg} = -\infty$ then. The likelihood for structural zeroes is $0^0 = 1$, and the log-likelihood is 0. We can achieve this result by the following simple device. Let

$$ Q_{ijg}^* = \begin{cases} Q_{ijg} & \text{if } n_{ijg} > 0; \\ 0.01 & \text{otherwise.} \end{cases} $$

Then

$$ l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ijg} \log Q_{ijg}^* \, . $$

3.2 Tag seeding data

3.2.1 Structure of the tag seeding data

Because we are assuming there are no observers in the surface fishery in region 1 (i.e., $\delta_{t11} = 0$ and $\delta_{t12} = 1$), we cannot estimate the reporting rates (λ_{t12}) in this region; they are confounded with fishing mortality. So for region 1 we assume that we have some independent tag seeding data. In region 1 and time periods $t = 1, 3, 5, \ldots$, suppose that D_t tags are seeded into the fish after being caught. Of these, d_t are reported.

3.2.2 Inference for the tag seeding data

We assume that these data are binomial, so that the log-likelihood is (apart from an additive constant)

$$l_{seed} = \sum_{t=1,3,\ldots} \left(d_t \log \lambda_{t12} + (D_t - d_t) \log(1 - \lambda_{t12}) \right).$$

3.3 Catch-at-age data

3.3.1 Structure of the catch-at-age data

We now look at the model for catch-at-age data. Again we concentrate on a single cohort, the same as for the tag-recapture data. A complication with catch-at-age data is that, since the fish do not have tags, we do not know their origins. However, we do know their ages, hence we can work out in which year they were one-year-olds. Let P_{tr} be the number of fish in the cohort alive in region r at the beginning of time period t. Suppose the population size of age 1 fish at the beginning of time period $t = 1$ in region r_0 is P_{1r_0}. Then the probability of a fish originating in region r_0 being captured in region r in time period t is $q_{tr|1r_0}$, as defined in equation (2). Thus, the probability of being captured in region r during time period t regardless of its source is

$$\frac{\sum_{r_0=1}^{R} q_{tr|1r_0} P_{1r_0}}{\sum_{r_0=1}^{R} P_{1r_0}} .$$

In the longline regions $(r = 2, \ldots, R)$, we assume that only the fish caught by vessels with observers are aged. Therefore the expected number of fish of this cohort caught in the observer component, and thus aged, in region r $(r = 2, \ldots, R)$ during time period t is

$$\mu_{tr} = \delta_{tr1} \sum_{r_0=1}^{R} q_{tr|1r_0} P_{1r_0} \,.$$

In the surface fishery $(r = 1)$, we assume that a proportion δ_{t11}^* of the fish caught are sampled and aged, so the same style of equation can be used:

$$\mu_{t1} = \delta_{t11}^* \sum_{r_0=1}^{R} q_{t1|1r_0} P_{1r_0} \,.$$

3.3.2   Inference for the catch-at-age data

For the catch-at-age data, we adopt a Gaussian model with known coefficient of variation c_r different for each region. The catches in each region and time are assumed to be statistically independent.

If the catch in region r and time t is C_{tr}, the log-likelihood is (apart from an additive constant)

$$l_{catch} = \sum_{t=1}^{T} \sum_{r=1}^{R} \left(-\log \mu_{tr} - \frac{1}{2} \frac{(C_{tr} - \mu_{tr})^2}{c_r^2 \mu_{tr}^2} \right) \,.$$

The Gaussian model is intended to be a catch-all for all sources of variance in the catch-at-age data, such as sampling variability and heterogeneity in capture probabilities of different vessels (refer to Appendix 5 for a complete discussion).

3.4   Overall log-likelihood

The parameters are estimated by maximising the total log-likelihood

$$l = l_{tag} + l_{seed} + l_{catch} \,.$$

We are implicitly assuming that all three types of data are statistically independent, as seems reasonable.

Estimates are obtained for the age-specific natural mortality parameters, the age-, season- and region-specific fishing mortality parameters, the initial region-specific population sizes, the age-specific transition probabilities between regions, and the time- and region-specific reporting rates. Note, however, that when a cohort is tagged at A consecutive ages, only $A-1$ natural mortality rate parameters can be estimated (this is shown in Appendix 13, and is a well-known feature in non-spatial Brownie models).

3.5 Overall (population-wide) fishing mortality

It is often of interest to compute population-wide fishing mortality parameters, averaged over all regions. There are several ways in which overall fishing mortality can be defined, and we describe three such ways below. All of our methods involve P_{tr}, the number of fish in the cohort alive in region r at the beginning of time period t, and $P_t = \sum_{r=1}^{R} P_{tr}$, the total number of fish in the cohort (over all regions) at the beginning of time period t.

Furthermore, in the methods below, t and $t+1$ constitute a year, where a year starts in season 1; thus, t is an odd integer corresponding to season 1 and $t+1$ is an even integer corresponding to season 2. Furthermore, a fish is age a throughout time periods t and $t+1$, and will become age $a+1$ at the start of time period $t+2$.

3.5.1 Average yearly fishing mortality

The simplest possibility is to define the overall fishing mortality by

$$P_t \exp(-F_a^* - M_a) = P_{t+2} \tag{5}$$

where, recall, $P_t = \sum_{r=1}^{R} P_{tr}$. Here we are saying that the total number of fish alive at the beginning of $t+2$ consists of those alive at the beginning of time period t and that survive both time periods t and $t+1$, and we assume an overall fishing mortality rate in this equation.

In practice, the simplest way to compute F_a^* is to estimate all parameters first, and then use these estimates to compute P_{tr} for each t and r. Then, using the estimated M_a values, compute F_a^* as

$$F_a^* = \log(P_t / P_{t+2}) - M_a \,.$$

In our simulations, we not only calculate estimated overall fishing mortalities, but for comparison we also calculate 'true' overall fishing mortalities, which are calculated in the same way except using the true parameter values instead of the estimated values.

It is possible to avoid the explicit calculation of P_{t+2} and the need for M_a in the computation of F_a^*. We examine survivors from the beginning of time period t to the end of time period $t+1$. In season 1 the fish can be in region 1 and survive. If so, it will move to region r for $r = 2, \ldots, R$, where (if lucky) it will also survive. Or in season 1 it can be in region r for $r > 1$, survive and stay there. Again, if lucky, it will survive to the end of $t+1$. The number of fish that are still surviving at the end of the year is

$$
\begin{aligned}
P_{t+2} &= \sum_{r=2}^{R} \Big(P_{t1} \exp(-F_{a11} - M_a/2)\pi_{1r|a1} + P_{tr} \exp(-M_a/2) \Big) \exp(-F_{a2r} - M_a/2) \\
&= \exp(-M_a) \sum_{r=2}^{R} (P_{t1}\exp(-F_{a11})\pi_{1r|a1} + P_{tr}) \exp(-F_{a2r}) \,. \tag{6}
\end{aligned}
$$

Here we are using the fact that the *total* population size at the end of time $t+1$ is identical to the total population size at the beginning of time $t+2$. Substituting equation (5) for P_{t+2}, we obtain

$$F_a^* = -\log\left(\frac{\sum_{r=2}^{R}(P_{t1}\exp(-F_{a11})\pi_{1r|a1} + P_{tr})\exp(-F_{a2r})}{P_t} \right) \,.$$

Thus we have expressed the average fishing mortality F_a^* in terms of the spatial fishing mortalities F_{a1r} and F_{a2r}, the population sizes at the beginning of the year and the transfer rates from region 1 to r at the end of season 1. This means that explicit calculation of P_{t+2} is avoided.

In fact, the most convenient formula from a computational viewpoint is to use

$$P_{t+2} = \sum_{r=2}^{R} P_{t+1,r} \exp(-F_{a2r} - M_a/2)$$

in (5). Hence

$$F_a^* = -\log\left(\frac{\sum_{r=2}^{R} P_{t+1,r} \exp(-F_{a2r})}{P_t}\right) - M_a/2 \, . \tag{7}$$

3.5.2 Average spatial fishing mortality

We could define an overall fishing mortality F_a by setting the non-zero fishing mortality parameters to be equal across regions within a year. Thus, we could set $F_{a11} = F_{a22} = F_{a23} = \ldots = F_{a2R} \equiv F_a$ in equation (6). By doing so, we are asking what (spatially) constant F_a would give rise to P_{t+2}. We obtain

$$\sum_{r=2}^{R} (P_{t1} \exp(-F_{a11})\pi_{1r|a1} + P_{tr}) \exp(-F_{a2r})$$

$$= \sum_{r=2}^{R} (P_{t1} \exp(-F_a)\pi_{1r|a1} + P_{tr}) \exp(-F_a)$$

$$= (P_{t1} \exp(-F_a) + \sum_{r=2}^{R} P_{tr}) \exp(-F_a) \, ,$$

since $\sum_{r=2}^{R} \pi_{1r|a1} = 1$. Set $\Sigma_1 = \sum_{r=2}^{R}(P_{t1} \exp(-F_{a11})\pi_{1r|a1} + P_{tr}) \exp(-F_{a2r})$, $\Sigma_2 = P_{t1}$, $\Sigma_3 = \sum_{r=2}^{R} P_{tr}$ and $\phi = \exp(-F_a)$. This equation becomes

$$\Sigma_1 = \Sigma_2 \phi^2 + \Sigma_3 \phi$$

with solution

$$\phi = \frac{-\Sigma_3 \pm \sqrt{\Sigma_3^2 + 4\Sigma_2 \Sigma_1}}{2\Sigma_2} \, .$$

Since $\Sigma_i > 0$ for $i = 1, 2$ and 3, and $\phi > 0$, only the solution involving the positive square-root is of interest to us, and we obtain

$$F_a = -\log\left(\frac{-\Sigma_3 + \sqrt{\Sigma_3^2 + 4\Sigma_2 \Sigma_1}}{2\Sigma_2}\right) \, .$$

Note that $\Sigma_1 = P_{t+2}/\exp(-M_a)$, where $P_t = \sum_{r=1}^{R} P_{tr}$. This is useful from a computational viewpoint. We would normally just compute P_t for all t, then choose the relevant ones for computing the overall fishing mortalities.

Unlike the average yearly fishing mortality F_a^* defined in the previous subsection, F_a takes into account the fact that the fish in region 1 in season 1 are subject to fishing pressure in both seasons. Thus F_a^* solves $\Sigma_1 = \Sigma_2 \phi + \Sigma_3 \phi$ and F_a solves $\Sigma_1 = \Sigma_2 \phi^2 + \Sigma_3 \phi$.

3.5.3   Average seasonal fishing mortality

Another possible summary statistic is the average seasonal fishing mortality. In season 1 this will be defined by

$$P_t \exp(-F_{a1}^* - M_a/2) \;=\; P_{t+1}\,. \tag{8}$$

Equating P_{t+1} to the total number of fish alive at the end of time period t (season 1), we obtain

$$P_t \exp(-F_{a1}^* - M_a/2) \;=\; P_{t1} \exp(-F_{a11} - M) + \sum_{r=2}^{R} P_{tr} \exp(-M_a/2)\,,$$

so that

$$F_{a1}^* \;=\; -\log\left(\frac{P_{t1}\exp(-F_{a11}) + \sum_{r=2}^{R} P_{tr}}{P_t}\right)\,.$$

Similarly for period $t+1$ (season 2) we obtain

$$F_{a2}^* \;=\; -\log\left(\frac{\sum_{r=2}^{R} P_{t+1,r}\exp(-F_{a2r})}{P_{t+1}}\right)\,.$$

Note that

$$F_{a1}^* \;=\; -\log\left(\frac{P_{t+1}}{P_t}\right) - M_a/2\,,$$

and hence that

$$
\begin{aligned}
F_{a1}^* + F_{a2}^* \;&=\; -\log\left(\frac{P_{t+1}}{P_t}\right) - M_a/2 - \log\left(\frac{\sum_{r=2}^{R} P_{t+1,r}\exp(-F_{a2r})}{P_{t+1}}\right) \\
&=\; -\log\left(\frac{\sum_{r=2}^{R} P_{t+1,r}\exp(-F_{a2r})}{P_t}\right) - M_a/2 \\
&=\; F_a^*
\end{aligned}
$$

from equation (7). Thus the average yearly fishing mortality is the sum of its average seasonal fishing mortalities.

3.6   Numerical example

In order to illustrate the model, we simulated data for a tagging experiment involving a single cohort of fish tagged and recaptured over 3 consecutive years ($T = 6$ time periods) in $R = 4$ regions. In accordance with the model, we simulated data with multinomial tag-recapture data, binomial tag-seeding data, and Gaussian catch-at-age data. We decided to start with good quality data, in which we had large sample sizes and the coefficient of variation of the catch-at-age data was $c_r = 0.1$ for each region. The population (cohort) sizes at the beginning of time period $t = 1$ in each region were $P_{11} = 850000$ and $P_{12} = P_{13} = P_{14} = 500000$. The tagging numbers N_{tr} in region r at the beginning of time interval t were $N_{11} = 1350$, $N_{22} = N_{23} = N_{24} = 500$, $N_{31} = 1275$, $N_{42} = N_{43} = N_{44} = 750$, $N_{51} = 1200$ and $N_{62} = N_{63} = N_{64} = 500$. The proportion of the fishery with observers (component $g = 1$) was $\delta_{tr1} = 0.1$ for all t and $r > 1$. Recall that $\delta_{t11} = 0$. In region 1 and season 1, the number of tags seeded into the catches was set at $D_t = 50$ for $t = 1, 3, 5$, and the proportion of the catch sampled to be aged was set at $\delta^*_{t11} = 0.05$ for $t = 1, 3, 5$. For those parameters estimated by the model, the true parameter values used in simulating the data are given in Table 1.

We simulated 100 such data sets and estimated the parameters by maximum likelihood, assuming that the coefficient of variation in the catch-at-age data, the proportions of observer coverage in regions 2, 3 and 4, and the proportion of catch sampled in region 1 were known. In Table 1 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 1: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the Markov model to 100 sets of good quality data generated from the Markov model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.093	0.289	0.492	$\pi_{41|32}$	0.700	—	—	—
M_2	0.200	0.010	0.181	0.409	$P_{11}/10^5$	8.500	5.979	8.426	11.894
F_{111}	0.150	0.114	0.151	0.201	$P_{12}/10^5$	5.000	1.683	5.072	8.092
F_{211}	0.100	0.082	0.103	0.126	$P_{13}/10^5$	5.000	2.075	5.003	8.780
F_{311}	0.200	0.159	0.196	0.247	$P_{14}/10^5$	5.000	2.147	5.026	8.205
F_{122}	0.200	0.130	0.203	0.299	λ_{112}	0.650	0.534	0.658	0.776
F_{222}	0.100	0.059	0.100	0.156	λ_{222}	0.250	0.134	0.248	0.412
F_{322}	0.100	0.060	0.099	0.175	λ_{232}	0.250	0.165	0.257	0.396
F_{123}	0.200	0.127	0.209	0.290	λ_{242}	0.250	0.154	0.262	0.406
F_{223}	0.100	0.067	0.100	0.156	λ_{312}	0.650	0.572	0.658	0.769
F_{323}	0.100	0.056	0.103	0.183	λ_{422}	0.250	0.171	0.257	0.385
F_{124}	0.200	0.126	0.208	0.315	λ_{432}	0.250	0.160	0.253	0.384
F_{224}	0.100	0.071	0.101	0.140	λ_{442}	0.250	0.165	0.259	0.426
F_{324}	0.100	0.057	0.100	0.160	λ_{512}	0.650	0.578	0.662	0.756
$\pi_{12|11}$	0.333	0.200	0.339	0.500	λ_{622}	0.250	0.154	0.264	0.385
$\pi_{13|11}$	0.333	0.213	0.327	0.490	λ_{632}	0.250	0.172	0.257	0.355
$\pi_{12|21}$	0.333	0.199	0.341	0.488	λ_{642}	0.250	0.180	0.260	0.353
$\pi_{13|21}$	0.333	0.188	0.332	0.454	F_1^*	0.252	0.204	0.249	0.290
$\pi_{12|31}$	0.333	0.184	0.340	0.531	F_2^*	0.190	0.160	0.184	0.214
$\pi_{13|31}$	0.333	0.182	0.332	0.485	F_3^*	0.257	0.199	0.246	0.314
$\pi_{21|12}$	0.900	0.625	0.886	0.990	F_{11}^*	0.052	0.037	0.051	0.066
$\pi_{31|12}$	0.900	0.612	0.863	0.990	F_{12}^*	0.200	0.158	0.198	0.237
$\pi_{41|12}$	0.900	0.652	0.870	0.990	F_{21}^*	0.090	0.076	0.088	0.104
$\pi_{21|22}$	0.800	0.617	0.790	0.947	F_{22}^*	0.100	0.082	0.096	0.114
$\pi_{31|22}$	0.800	0.611	0.797	0.989	F_{31}^*	0.157	0.126	0.151	0.186
$\pi_{41|22}$	0.800	0.560	0.785	0.990	F_{32}^*	0.100	0.072	0.095	0.131
$\pi_{21|32}$	0.700	—	—	—	$P_1/10^5$	23.500	19.740	23.527	28.007
$\pi_{31|32}$	0.700	—	—	—	ϵ	0.000	0.000	0.000	0.000

All of the parameters are reasonably well estimated; however, there is suggestion of some small biases. In particular, both natural mortality parameters are slightly underestimated on average. The initial population count (at the start of time period 1) in region 1 is slightly too small on average, but seems to be compensated by

small positive biases in the initial population estimates in the other three regions. Indeed, the overall initial population size estimate (P_1) has a very slight positive bias. The mean reporting rate (λ_{tr2}) estimates are slightly too large, and there are small negative biases in the transition probabilities back to region 1 at the end of season 2 ($\pi_{r1|12}$ and $\pi_{r1|22}$ for $r = 2, 3, 4$). Nevertheless, it is reassuring to know that if we have enough good quality data all parameters are estimable with reasonable precision. Note that the three $\pi_{r1|32}$ parameters refer to transitions at the end of period 6 (age 3, season 2), at which stage we have stopped collecting data, and thus are not estimable from these data.

Also given in Table 1 are overall fishing mortality estimates and their true values. In particular, F_a^* is the yearly average fishing mortality for fish of age a outlined in Section 3.5.1, and F_{as}^* is the average seasonal fishing mortality for fish of age a in season s outlined in Section 3.5.3. These are all slightly underestimated on average.

The final figure in the table, ϵ, is the maximum absolute derivative of the log-likelihood with respect to each parameter, and should be close to 0 if the maximum likelihood estimation routine has converged satisfactorily.

Next we considered an example with poorer quality data, likely to be more representative of the situation for SBT. We increased the coefficient of variation for the catch-at-age data to $c_r = 0.3$ for each region. We assumed the initial population sizes in the longline regions were smaller and that the numbers of fish tagged in these regions were less. Specifically, we set the initial population sizes in each region to be $P_{11} = 850000$ and $P_{12} = P_{13} = P_{14} = 50000$, and the numbers of fish tagged in region r at the beginning of time interval t (N_{tr}) to be $N_{11} = 1350$, $N_{22} = N_{23} = N_{24} = 50$, $N_{31} = 1275$, $N_{42} = N_{43} = N_{44} = 75$, $N_{51} = 1200$ and $N_{62} = N_{63} = N_{64} = 50$. The fishing mortality rates were also lowered quite significantly at ages 1 and 2 to be more representative of fishing rates for SBT (see Table 2). All other parameters were kept the same as in the first example. Both the smaller release numbers and lower fishing mortality rates will result in lower numbers of returns.

We again simulated 100 data sets and estimated the parameters by maximum likelihood, assuming that the coefficient of variation in the catch-at-age data, the proportions of observer coverage in regions 2, 3 and 4, and the proportion of catch sampled in region 1 were known. In Table 2 we report the mean and the 2.5% and 97.5% quantiles of the estimates.

Table 2: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the Markov model to 100 sets of poor quality data generated from the Markov model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.046	0.295	0.502	$\pi_{41\|32}$	0.700	—	—	—
M_2	0.200	0.010	0.167	0.442	$P_{11}/10^5$	8.500	3.117	6.981	10.904
F_{111}	0.020	0.014	0.021	0.029	$P_{12}/10^5$	0.500	0.030	0.812	4.572
F_{211}	0.080	0.067	0.087	0.109	$P_{13}/10^5$	0.500	0.030	0.932	4.347
F_{311}	0.200	0.152	0.204	0.279	$P_{14}/10^5$	0.500	0.030	1.257	6.063
F_{122}	0.001	0.000	0.004	0.017	λ_{112}	0.650	0.519	0.648	0.787
F_{222}	0.020	0.008	0.029	0.090	λ_{222}	0.250	0.050	0.143	1.000
F_{322}	0.100	0.042	0.113	0.241	λ_{232}	0.250	0.050	0.098	1.000
F_{123}	0.001	0.000	0.003	0.011	λ_{242}	0.250	0.050	0.174	1.000
F_{223}	0.020	0.007	0.026	0.080	λ_{312}	0.650	0.514	0.643	0.766
F_{323}	0.100	0.038	0.115	0.251	λ_{422}	0.250	0.050	0.275	0.660
F_{124}	0.001	0.000	0.003	0.013	λ_{432}	0.250	0.052	0.298	0.824
F_{224}	0.020	0.007	0.031	0.122	λ_{442}	0.250	0.050	0.308	0.763
F_{324}	0.100	0.032	0.105	0.235	λ_{512}	0.650	0.508	0.649	0.782
$\pi_{12\|11}$	0.333	0.000	0.371	0.990	λ_{622}	0.250	0.113	0.272	0.510
$\pi_{13\|11}$	0.333	0.000	0.325	0.990	λ_{632}	0.250	0.110	0.258	0.457
$\pi_{12\|21}$	0.333	0.005	0.326	0.860	λ_{642}	0.250	0.143	0.288	0.564
$\pi_{13\|21}$	0.333	0.003	0.332	0.838	F_1^*	0.018	0.008	0.016	0.027
$\pi_{12\|31}$	0.333	0.099	0.335	0.664	F_2^*	0.092	0.052	0.085	0.114
$\pi_{13\|31}$	0.333	0.087	0.320	0.609	F_3^*	0.257	0.158	0.247	0.374
$\pi_{21\|12}$	0.900	0.010	0.716	0.990	F_{11}^*	0.017	0.007	0.015	0.026
$\pi_{31\|12}$	0.900	0.010	0.636	0.990	F_{12}^*	0.001	0.001	0.001	0.001
$\pi_{41\|12}$	0.900	0.010	0.725	0.990	F_{21}^*	0.072	0.037	0.066	0.093
$\pi_{21\|22}$	0.800	0.464	0.780	0.990	F_{22}^*	0.020	0.012	0.019	0.028
$\pi_{31\|22}$	0.800	0.297	0.722	0.990	F_{31}^*	0.157	0.099	0.153	0.229
$\pi_{41\|22}$	0.800	0.329	0.746	0.990	F_{32}^*	0.100	0.057	0.094	0.159
$\pi_{21\|32}$	0.700	—	—	—	$P_1/10^5$	10.000	6.815	9.982	13.622
$\pi_{31\|32}$	0.700	—	—	—	ϵ	0.000	0.000	0.000	0.000

The higher coefficients of variation and reduced population sizes and numbers of tag returns have resulted in much poorer estimates in general. The precision of all parameter estimates has decreased substantially; this is especially true for the age 1 reporting rates in regions 2 to 4 (time period 2) and the transfer probabilities for age 1 fish, due to the age 1 fishing mortality being so low that very little tag-return data exists for estimating these parameters. Some of the biases are also quite large. In comparison to the good quality data example, the negative bias in natural mortality at ages 2 and above (M_2) has increased, the negative biases in the transition probabilities back to region 1 at the end of season 2 have become more pronounced, and the biases in the reporting rate estimates are larger (and not always positive). The negative bias in the initial population size in region 1 is much greater, as are the positive biases in regions 2, 3 and 4. Interestingly, though, the total initial population size estimate is essentially unbiased despite the regional population estimates being so poorly estimated. Finally, some small positive biases have also become apparent in all of the regional fishing mortality estimates; however, the average yearly and average seasonal fishing mortality estimates are slightly underestimated (presumably resulting from the large negative bias in the region 1 population size estimate).

4   The second SBT model: the winter region fidelity model

This is a variation on the SBT model with 4 regions and two seasons per year. Again, region 1 is representative of the GAB with a surface fishery operating only in the austral summer (season 1). The remaining 3 regions represent the longline fisheries that operate in the winter (season 2) in waters off South Africa, in the South East Indian Ocean, and in Tasman Sea. As explained in the Introduction, in the second SBT model, each fish remains faithful to its winter region, and thus has 'memory'.

Note that the model can easily be generalized to any number of winter regions but for notational convienience we will use 3 winter regions in our presentation.

For the second SBT model, we classify all fish from the cohort of interest that are alive at the beginning of the time period 1 as belonging to one of three categories depending on the winter region to which they are faithful: category 2 fish are those which are faithful to region 2; category 3 fish are those which are faithful to region 3; and category 4 fish are those fish which are faithful to region 4. Let P_{1r} represent the population size of the cohort in region r at the start of time period 1. Then the P_{12} fish in region 2 at the start of time period 1 clearly belong to category 2, and similar statements apply for regions 3 and 4. For the P_{11} fish in region 1 at the start of time period 1, we assume each fish can be classified as belonging to category 2, 3 or 4 according to whether it migrates to region 2, 3 or 4 for the winter season, or would do so if it survived. In particular, we assume each fish in region 1 can be placed into category r with probability w_r ($\sum_{r=2}^{4} w_r = 1$). We can think of w_r as the probability of a fish moving from region 1 to region r at the end of time period 1 (age 1 and season 1); using the notation from the first SBT model, this means $w_r = \pi_{1r|11}$.

Before launching into a general argument, we consider those fish from a particular category in order to clarify the basic ideas.

4.1 Category 2 fish

Focus on a fish belonging to category 2. At the end of each season, this fish moves according to a two state Markov chain, in which the states are regions 1 and 2. Let $\Pi_{as}^{(2)}$ be the transition matrix at the end of time period t corresponding to age a and season s for fish in category 2. Then, for season 1, we define

$$\Pi_{a1}^{(2)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

In other words, at the end of season 1, the category 2 fish in region 1 always move to region 2 and those fish already in region 2 stay there. Note that the (3,3) and (4,4) diagonal entries of $\Pi_{a1}^{(2)}$ are set to 1. In fact, only the top left 2×2 matrix is used by fish in category 2. It is possible to do all calculations in terms of 2×2 matrices and vectors of length 2. This is the most efficient computational method, and it is how our computer programs operate. It does, however, require that you keep track of whether you are dealing with fish in category 2, 3 or 4. For this reason, it is easier to write all the equations in terms of 4×4 matrices, and vectors of length 4.

For season 2, we define

$$
\Pi_{a2}^{(2)} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
\pi_{21|a2} & 1 - \pi_{21|a2} & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
$$

Thus, at the end of season 2, a fish from category 2 either moves to region 1 or stays in region 2. The selection of fish that move is random (i.e., whether or not a fish migrates to region 1 at the end of season 2 is independent of whether it has previously been in region 1). For SBT, we expect the transitions back to region 1 at the end of season 2 to be close to zero, if not zero, for fish of age 5 and older.

When analysing the data we shall need to consider fish tagged at the beginning of time t_C in region r_C. This implies that we shall need to know the expected number of fish from each source alive in region r_C at the beginning of time period t_C. We shall also need the probability that a fish, known to be alive in region r_C at the beginning of time period t_C, is caught subsequently in the various regions.

For a category 2 fish starting out in region 1, the probability that it is alive at the beginning of time period 1 in region r_C equals 1 for region 1 and equals 0 for

all other regions. We denote this in vector notation by

$$p^{(2)}_{1|11} \;=\; \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Similarly, for a category 2 fish starting out in region 2, the probability that it is alive at the beginning of time period 1 in region r_C equals 1 for region 2 and equals 0 for all other regions. We denote this in vector notation by

$$p^{(2)}_{1|12} \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

No other regions are relevant for category 2 fish. We can make similar statements for fish from categories 3 and 4. In general, for fish from category r $(r = 2, 3, 4)$, the first element of $p^{(r)}_{1|11}$ equals 1 and all other elements equal 0; and the rth element of $p^{(r)}_{1|1r}$ equals 1 and all other elements equal 0.

Denote the probability that a category 2 fish starting out in region 1 is alive in region r_C at the beginning of time period t_C by $p^{(2)}_{t_C r_C|11}$. This can be calculated iteratively as follows. Using the starting vector $p^{(2)}_{1|11}$ and the updating equation

$$p^{(2)}_{t+1|11} \;=\; \Pi'^{(2)}_{as} \, S_{as} \, p^{(2)}_{t|11} \,,$$

calculate $p^{(2)}_{t_C|11}$, then the rth element is $p^{(2)}_{t_C r_C|11}$. Here S_{as} is the 4×4 diagonal matrix whose rth diagonal entry $S_{asr} = \exp(-M_a/2 - F_{asr})$ is the probability of a fish in region r at the start of time period t corresponding to age a and season s surviving to the end of the time period.

The expected number of category 2 fish starting out in region 1 that are alive at the beginning of time period t_C in region r_C is

$$P_{11} \, w_2 \, p^{(2)}_{t_C r_C|11} \,.$$

Recall that P_{11} is the number of fish from the cohort in region 1 at the start of time period 1 and that w_2 is the proportion of fish in region 1 at the start of time period 1 that belong to category 2.

Now we can make similar statements for category 2 fish starting out in region 2. In particular, the probability that a category 2 fish starting out in region 2 is alive in region r_C at the beginning of time period t_C can be calculated iteratively using the starting vector $p^{(2)}_{1|12}$ and the updating equation

$$p^{(2)}_{t+1|12} \;\; = \;\; \Pi'^{(2)}_{as}\, S_{as}\, p^{(2)}_{t|12}$$

to calculate $p^{(2)}_{t_C|11}$ and then extracting the rth element.

The expected number of category 2 fish starting out in region 2 that are alive at the beginning of time period t_C in region r_C is

$$P_{12}\, p^{(2)}_{t_C r_C|12}\,.$$

Recall that P_{12} is the number of fish from the cohort in region 2 at the start of time period 1 (all belonging to category 2).

Given that a category 2 fish is alive at the beginning of time period t_C in region r_C, denote the probability that it is recaptured in time period t ($\geq t_C$) in region r by $p^{(2)}_{tr|t_C r_C}$. This can be calculated iteratively in a similar way as above using the updating equation

$$p^{(2)}_{t+1|t_C r_C} \;\; = \;\; \Pi'^{(2)}_{as}\, S_{as}\, p^{(2)}_{t|t_C r_C}\,.$$

Here we are using for the first iteration the fact that the rth element of $p^{(2)}_{t_C|t_C r_C}$ equals 1 for $r = r_C$ and 0 for $r \neq r_C$. Note that this statement applies to all category 2 fish (regardless of whether they started time period 1 in region 1 or 2). The only real difference between category 2 fish originating in region 1 versus region 2 is what happens in time period 1. After that they behave the same in a statistical sense.

4.2 All regions

Following the principles outlined in the previous section, the expected number of fish from the cohort (from all categories) that are alive at the beginning of time period t_C in region r_C is given by

$$P_{t_C r_C} = P_{11} \sum_{r=2}^{4} w_r \, p_{t_C r_C|11}^{(r)} + \sum_{r=2}^{4} P_{1r} \, p_{t_C r_C|1r}^{(r)} \,. \tag{9}$$

Equation (9) is a general expression and often some of the terms will be 0. For example, if $t_C > 1$ and $r_C = 2$, then only terms involving $p_{t_C r_C|11}^{(2)}$ and $p_{t_C r_C|12}^{(2)}$ can be non-zero; similarly for $r_C = 3$ and $r_C = 4$. But when t_C (> 1) is odd and $r_C = 1$, all terms may be non-zero, whereas when t_C is even and $r_C = 1$, all terms will be 0 since no fish are found in region 1 in the winter season.

The expected number of fish alive at the beginning of time period t in region r that were also alive in region r_C at the beginning of time period t_C is

$$P_{tr, t_C r_C} = P_{11} \sum_{r'=2}^{4} w_{r'} \, p_{t_C r_C|11}^{(r')} \, p_{tr|t_C r_C}^{(r')} + \sum_{r'=2}^{4} P_{1r'} \, p_{tr|1r'}^{(r')} \, p_{tr|t_C r_C}^{(r')} \,. \tag{10}$$

Then the probability that a fish, alive in region t_C at the beginning of period t_C, is alive in region r at the beginning of period t is obtained by dividing (10) by (9). Thus

$$p_{tr|t_C r_C} = \frac{P_{11} \sum_{r'=2}^{4} w_{r'} \, p_{t_C r_C|11}^{(r')} \, p_{tr|t_C r_C}^{(r')} + \sum_{r'=2}^{4} P_{1r'} \, p_{t_C r_C|1r'}^{(r')} \, p_{tr|t_C r_C}^{(r')}}{P_{11} \sum_{r'=2}^{4} w_{r'} \, p_{t_C r_C|11}^{(r')} + \sum_{r'=2}^{4} P_{1r'} \, p_{t_C r_C|1r'}^{(r')}} \,. \tag{11}$$

It is not always the case in statistics that $E[X/Y] = E[X]/E[Y]$ for random variables X and Y, however it is true here. Without going into great detail, if we assume that the P_{1r}'s are independent Poisson variables, and that the allocation of fish to categories and movement of fish within categories is random, then (using thinning theorems for Poisson processes) the number of fish in a given region at the start of a given time period will also be Poisson. Then, applying the standard arguments for independent Poisson variables, it follows that the conditional probability of a fish being in region r at time t given that it was alive in region r_C at time t_C is binomial with 'success' probability given by equation (11).

The quantity in (11) is directly analogous to $p_{tr|t_C r_C}$ in the first SBT model; however, there is an important difference. The quantities $p_{tr|t_C r_C}$ in SBT model 1 do not depend on the initial population sizes P_{1r}, but they do in SBT model 2. This is important because it means, unlike in SBT model 1, in SBT model 2 the tagging data cannot be analysed independently of the catch-at-age data.

The probability of a fish, alive in region t_C at the beginning of period t_C, being caught in region r during time period t (corresponding to age a and season s) is

$$q_{tr|t_C r_C} = \frac{F_{asr}}{M_a/2 + F_{asr}}(1 - S_{asr})\, p_{tr|t_C r_C} \,. \tag{12}$$

Here $p_{tr|t_C r_C}$ is given by (11), $(1 - S_{asr}) = \exp(-M_a/2 - F_{asr})$ is the probability that the fish does not survive the time period, and $F_{asr}/(M_a/2 + F_{asr})$ is the probability that it is caught (rather than dies naturally) given that it does not survive. This is directly analogous to $q_{tr|t_C r_C}$ from equation (2) in SBT model 1.

We assume that fish are tagged randomly, and that tagged and untagged fish behave in the same way, so that the probability statements given in (11) and (12) apply equally to tagged and untagged fish.

4.3 Tagging data

The structure of the tag-return data and the likelihood for these data are essentially the same as for the first SBT model, as described in Sections 3.1.1 and 3.1.2. The only difference is that in equation (4), the term $q_{tr|t_C r_C}$ is now given by equation (12) rather than equation (2).

4.4 Tag seeding data

The structure of the tag seeding data and the likelihood for these data are the same as in Sections 3.2.1 and 3.2.2.

4.5 Catch-at-age data

The structure of the catch-at-age data and the likelihood for these data are essentially the same as described in Sections 3.3.1 and 3.3.2. Similar to the tagging data, the only difference is that the term $q_{tr|1r_0}$ in the expression for μ_{tr}, the expected number of fish from the cohort caught in region r during time period t, is now defined by equation (12) rather than equation (2).

4.6 Overall log-likelihood

The parameters are again estimated by maximising the total log-likelihood

$$l = l_{tag} + l_{seed} + l_{catch} \, .$$

The parameters to be estimated are the same as for the first SBT model except that there are less transition probabilities (π parameters) because after the first time period, we no longer need to estimate the transitions from region 1 to regions 2, 3 and 4 at the end of season 1 as they are determined by the proportion of category 2, 3 and 4 fish present.

4.7 Overall (population-wide) fishing mortality

As with the first SBT model, we outline three possible ways in which an overall fishing mortality parameter, averaged over regions, can be defined and calculated. We again use the notation that P_{tr} is the number of fish in the cohort alive in region r at the beginning of time period t, and $P_t = \sum_{r=1}^{R} P_{tr}$ is the total number of fish from the cohort alive at the beginning of time period t. We also introduce the notation that $P_{t1}^{(r)}$ is the number of fish from category r ($r = 2, 3, 4$) that are in region 1 at the beginning of time period t.

The three methods are essentially the same as for the first SBT model so we will only describe where there are differences.

4.7.1 Average yearly fishing mortality

Refer to Section 3.5.1 for SBT model 1. The only difference for SBT model 2 is equation (6) should be replaced with

$$
\begin{aligned}
P_{t+2} &= \sum_{r=2}^{4} \left(P_{t1}^{(r)} \exp(-F_{a11} - M_a/2) + P_{tr} \exp(-M_a/2) \right) \exp(-F_{a2r} - M_a/2) \\
&= \exp(-M_a) \sum_{r=2}^{4} (P_{t1}^{(r)} \exp(-F_{a11}) + P_{tr}) \exp(-F_{a2r}) .
\end{aligned} \tag{13}
$$

Equation (13) should then be substituted for P_{t+2} in equation (5) to solve for F_a^*. In words, the only difference between (13) and (6) is that the number of fish in region 1 in time period t, where t corresponds to age a and season 1, that migrate to region r at the end of the season no longer depends on the transition probability $\pi_{1r|a1}$ but is instead determined by the number of category r fish in region 1 at the start of time period t, $P_{t1}^{(r)}$, that survive to the end of the period.

4.7.2 Average spatial fishing mortality

Analogously to Section 3.5.2, we could also define an overall fishing mortality F_a by setting $F_{a11} = F_{a22} = F_{a23} = F_{a24} \equiv F_a$ in (13). We obtain

$$
\sum_{r=2}^{4} (P_{t1}^{(r)} \exp(-F_{a11}) + P_{tr}) \exp(-F_{a2r}) = \sum_{r=2}^{4} (P_{t1}^{(r)} \exp(-F_a) + P_{tr}) \exp(-F_a) .
$$

If we set $\Sigma_1 = \sum_{r=2}^{4} (P_{t1}^{(r)} \exp(-F_{a11}) + P_{tr}) \exp(-F_{t+1,r})$, $\Sigma_2 = \sum_{r=2}^{R} P_{t1}^{Cr}$, $\Sigma_3 = \sum_{r=2}^{R} P_{tr}$ and $\phi = \exp(-F_a)$, and solve for F_a (see Section 3.5.2), we obtain

$$
F_a = -\log \left(\frac{-\Sigma_3 + \sqrt{\Sigma_3^2 + 4\Sigma_2 \Sigma_1}}{2\Sigma_2} \right) .
$$

4.7.3 Average seasonal fishing mortality

Another possible summary statistic is the average seasonal fishing mortality; that is, the average fishing mortality for a given time period over all regions. The description and equations given in Section 3.5.3 for SBT model 1 are all valid for SBT model 2.

4.8 Numerical example

In order to illustrate the second SBT model, we examined the same two scenarios presented for the first SBT model, one with good quality data and one with poor quality data. Of course, not all of the transition probabilities are relevant in the site-fidelity model. Thus, we set $w_2 = \pi_{12|11}$, $w_3 = \pi_{13|11}$ and $w_4 = 1 - w_2 - w_3$ and ignored $\pi_{12|21}$, $\pi_{13|21}$, $\pi_{12|31}$ and $\pi_{13|31}$. Otherwise, we took all parameter values to be the same.

First consider the good quality data scenario, with true parameter values specified in Table 1 and the two paragraphs preceeding it. Using these parameter values, we simulated 100 data sets in accordance with the site-fidelity model and estimated the relevant parameters by maximizing the overall likelihood for this model. As usual, we assumed that the coefficients of variation for the catch-at-age data and the proportions of observer coverage/catch sampling were known. Summary results are given in Table 3.

Table 3: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the site-fidelity model to 100 sets of good quality data generated from the site-fidelity model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.088	0.285	0.496	$P_{11}/10^5$	8.500	5.979	8.438	11.865
M_2	0.200	0.010	0.190	0.395	$P_{12}/10^5$	5.000	2.624	5.061	7.695
F_{111}	0.150	0.114	0.151	0.201	$P_{13}/10^5$	5.000	2.696	5.010	7.166
F_{211}	0.100	0.083	0.104	0.131	$P_{14}/10^5$	5.000	2.793	4.952	7.963
F_{311}	0.200	0.163	0.199	0.251	λ_{112}	0.650	0.532	0.658	0.775
F_{122}	0.200	0.144	0.198	0.265	λ_{222}	0.250	0.147	0.250	0.358
F_{222}	0.100	0.062	0.097	0.132	λ_{232}	0.250	0.179	0.257	0.369
F_{322}	0.100	0.063	0.096	0.141	λ_{242}	0.250	0.154	0.260	0.384
F_{123}	0.200	0.143	0.205	0.279	λ_{312}	0.650	0.571	0.655	0.774
F_{223}	0.100	0.074	0.100	0.136	λ_{422}	0.250	0.170	0.260	0.373
F_{323}	0.100	0.062	0.102	0.156	λ_{432}	0.250	0.170	0.252	0.383
F_{124}	0.200	0.146	0.204	0.267	λ_{442}	0.250	0.151	0.258	0.389
F_{224}	0.100	0.070	0.100	0.139	λ_{512}	0.650	0.551	0.659	0.759
F_{324}	0.100	0.062	0.099	0.148	λ_{622}	0.250	0.170	0.263	0.389

Table 3: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
w_2	0.333	0.238	0.340	0.466	λ_{632}	0.250	0.163	0.256	0.376
w_3	0.333	0.217	0.327	0.453	λ_{642}	0.250	0.175	0.259	0.380
w_4	0.334	0.222	0.333	0.442	F_1^*	0.252	0.210	0.250	0.297
$\pi_{21|12}$	0.900	0.602	0.857	0.990	F_2^*	0.190	0.159	0.184	0.210
$\pi_{31|12}$	0.900	0.576	0.859	0.990	F_3^*	0.257	0.197	0.248	0.300
$\pi_{41|12}$	0.900	0.516	0.863	0.990	F_{11}^*	0.052	0.037	0.051	0.065
$\pi_{21|22}$	0.800	0.607	0.789	0.990	F_{12}^*	0.200	0.168	0.199	0.235
$\pi_{31|22}$	0.800	0.602	0.798	0.988	F_{21}^*	0.090	0.075	0.088	0.105
$\pi_{41|22}$	0.800	0.569	0.786	0.986	F_{22}^*	0.100	0.082	0.097	0.114
$\pi_{21|32}$	0.700	—	—	—	F_{31}^*	0.157	0.126	0.152	0.188
$\pi_{31|32}$	0.700	—	—	—	F_{32}^*	0.100	0.072	0.096	0.127
$\pi_{41|32}$	0.700	—	—	—	$P_1/10^5$	23.500	19.752	23.462	27.769
					ϵ	0.000	0.000	0.000	0.000

We then repeated the exercise for the poor quality data scenario, with true parameter values specified in Table 2 and the two paragraphs preceeding it. The summary results from 100 simulations are given in Table 4.

Table 4: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the site-fidelity model to 100 sets of poor quality data generated from the site-fidelity model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.049	0.296	0.503	$P_{11}/10^5$	8.500	3.141	6.881	10.254
M_2	0.200	0.010	0.163	0.446	$P_{12}/10^5$	0.500	0.081	0.961	4.015
F_{111}	0.020	0.014	0.021	0.030	$P_{13}/10^5$	0.500	0.081	0.981	3.776
F_{211}	0.080	0.068	0.086	0.109	$P_{14}/10^5$	0.500	0.081	1.149	4.137
F_{311}	0.200	0.152	0.206	0.276	λ_{112}	0.650	0.522	0.648	0.782
F_{122}	0.001	0.000	0.001	0.002	λ_{222}	0.250	0.050	0.145	1.000
F_{222}	0.020	0.008	0.021	0.047	λ_{232}	0.250	0.050	0.098	1.000
F_{322}	0.100	0.042	0.102	0.210	λ_{242}	0.250	0.050	0.174	1.000
F_{123}	0.001	0.000	0.001	0.002	λ_{312}	0.650	0.519	0.645	0.768
F_{223}	0.020	0.009	0.022	0.048	λ_{422}	0.250	0.050	0.291	0.892
F_{323}	0.100	0.037	0.105	0.206	λ_{432}	0.250	0.053	0.273	0.597
F_{124}	0.001	0.000	0.001	0.003	λ_{442}	0.250	0.055	0.307	0.904
F_{224}	0.020	0.006	0.021	0.046	λ_{512}	0.650	0.502	0.647	0.780

Table 4: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
F_{324}	0.100	0.035	0.103	0.242	λ_{622}	0.250	0.120	0.281	0.557
w_2	0.333	0.114	0.338	0.669	λ_{632}	0.250	0.113	0.264	0.491
w_3	0.333	0.089	0.338	0.626	λ_{642}	0.250	0.140	0.289	0.557
w_4	0.334	0.092	0.324	0.643	F_1^*	0.018	0.008	0.016	0.026
$\pi_{21|12}$	0.900	0.299	0.804	0.990	F_2^*	0.092	0.061	0.087	0.115
$\pi_{31|12}$	0.900	0.351	0.785	0.990	F_3^*	0.257	0.167	0.243	0.371
$\pi_{41|12}$	0.900	0.352	0.826	0.990	F_{11}^*	0.017	0.007	0.015	0.024
$\pi_{21|22}$	0.800	0.486	0.790	0.990	F_{12}^*	0.001	0.001	0.001	0.001
$\pi_{31|22}$	0.800	0.323	0.728	0.990	F_{21}^*	0.072	0.044	0.068	0.093
$\pi_{41|22}$	0.800	0.389	0.756	0.990	F_{22}^*	0.020	0.012	0.019	0.030
$\pi_{21|32}$	0.700	—	—	—	F_{31}^*	0.157	0.102	0.150	0.221
$\pi_{31|32}$	0.700	—	—	—	F_{32}^*	0.100	0.056	0.092	0.151
$\pi_{41|32}$	0.700	—	—	—	$P_1/10^5$	10.000	6.804	9.972	13.252
					ϵ	0.000	0.000	0.000	0.000

For both the good quality and poor quality data scenarios, the results from for the site-fidelity model are essentially the same as the results for the Markov model. In both of these scenarios, the longline regions are 'symmetric' in that they have the same initial population sizes, the same fishing mortality rates, the same reporting rates, and the same transition rates; thus, the data corresponding to the two SBT models are almost equivalent. If the longline regions differed, then the data would differ between the two models. For example, all other things being equal, say that region 2 had a much higher fishing mortality than regions 3 and 4. Then in the site-fidelity model this region would become depleted of fish and the total population size would be determined by regions 3 and 4 only. On the contrary, in the Markov model fish would continue to redistribute to region 2, where they would be fished heavily, so the total population size would decline much more quickly than in the site-fidelity model. The effect of different scenarios on parameter estimation in the two SBT models will depend on the scenario.

5 Pooled data

Suppose that all of the data (i.e., the tag releases and returns and the catch-at-age data) are pooled over regions within each time period, and the pooled data are analysed ignoring regions. The idea is to see whether any substantial biases are incurred by ignoring the spatial structure of the data. The model for analysing the pooled data is the same whether the underlying spatial structure follows that of SBT model 1 or SBT model 2.

5.1 Tagging data

Consider a single cohort of fish. Suppose the N_t fish from this cohort are tagged at the beginning of each time period t. Recall that a year is split into two time periods, referred to as season 1 and season 2, and that fish are tagged at age 1 at the beginning of time period 1. Age a for the cohort of interest and season s are then direct functions of the time period (namely $s = s(t) = t - 2[(t-1)/2]$ and $a = a(t) = [(t+1)/2]$, where $[x]$ means the integer part of x).

Let $S_{as} = \exp(-F_{as} - M_a/2)$ denote the probability of a fish surviving to the end of time period t, where t corresponds to age a and season s. Here F_{as} is fishing mortality at age a in season s (no dependence on region) and M_a is natural mortality at age a. Also let $f_{as} = F_{as}/(F_{as} + M_a/2)$ denote the conditional probability that a fish, which does not survive the time period, is caught rather than dies naturally.

Tagged fish are subsequently recaptured in the fishery and a portion of the tags are returned. The fishery is divided into 2 components, where component 1 consists of vessels with observers. Let δ_{tg} be the proportion of fish caught by component g during time period t, and λ_{tg} be the proportion of recaptured tags that are returned by component g during time period t. We assume that $\lambda_{t1} = 1$ for all t. Analogous to the spatial models, we assume there are no observers in the fishery operating in season 1 (the GAB surface fishery), so that $\delta_{t1} = 0$ for $t = 1, 3, \ldots$. This means there is not enough information to estimate reporting rates

in odd time periods, so we again assume that tag seeding data are available in these periods.

For a fish tagged at the beginning of time period t_C, the probability of being recaptured in time period t by component g and its tag returned is

$$q_{tg|t_C} = \left(\prod_{t'=t_C}^{t-1} S_{a's'} \right) (1 - S_{as}) f_{as} \delta_{tg} \lambda_{tg}$$

where $a' = a(t')$ and $s' = s(t')$. Note that when $t = t_C$, $\prod_{t'=t_C}^{t-1} S_{a's'} \equiv 1$. The λ_{tg} parameters are estimated from the data, but the δ_{tg} parameters are assumed known. In practice, the proportions δ_{tg} would likely be inferred from the number of fish caught in each component. Although these are not known exactly (since the catch data are random), the size of the catch is generally large enough that the error will be small.

We may actually know (or have estimates of) the δ_{trg} values by region (as we assume in our simulations). If this is the case, we need to average them sensibly to calculate δ_{tg}. We suggest weighting the δ_{trg} values by the catch-at-age data. This parallels the way the overall δ_{tg} values would be calculated in practice. In season 1, no vessels with observers are operating, and so the concept is not meaningful. As specified above, we set $\delta_{t1} = 0$ for season 1. In season 2, there are no fish or fishing in region 1, but we assume that the other regions do have some vessels with observers. Let C_{tr} be the catch-at-age from the cohort of interest reported by the observer component $g = 1$ during time interval t in region r. The estimated total catch from this cohort during time interval t in region r is therefore C_{tr}/δ_{tr1}. We therefore suggest the weighted value

$$\delta_{tg} = \frac{\sum_{r=2}^{R} \delta_{trg} C_{tr}/\delta_{tr1}}{\sum_{r=2}^{R} C_{tr}/\delta_{tr1}} . \tag{14}$$

It is immediate that $\delta_{t1} + \delta_{t2} = 1$. Note that δ_{t1} is a weighted harmonic mean. If $\delta_{tr1} = 0$ for any t and r, then $C_{tr} = 0$, so (14) will contain $0/0$ terms. There seems to be little we can do about this, except to be alert to the possibility. In practice, we would need to replace C_{tr}/δ_{tr1} by some independent estimate of the total catch for region r.

It should be emphasised that the use of (14) implies that, in general, the tag-recapture data cannot be analysed unless catch-at-age data are available. Also, the δ_{tg} will be regarded as known and fixed, even though they rely on quantities (catch-at-age values) that are subject to error. However, if δ_{trg} does not vary with r for given t and g, then δ_{tg} does not rely on catch-at-age data.

In many of the scenarios we shall study, δ_{trg} does not depend on t. In principle, δ_{tg} still depends on t through C_{tr}.

It is also important to understand the link between the λ_{tg} parameters and their regional equivalents λ_{trg}. In particular, in order to compare the reporting rate estimates obtained from the spatial models with those obtained from the pooled model, we will need to calculate λ_{tg} estimates for the spatial model using the estimated λ_{trg} parameters. Moreover, in our simulations we assume true values for the λ_{trg} parameters and we need to be able to calculate true λ_{tg} values from these.

During season 1, fish are caught only in region 1. Note that tag return rates are conditional on the fish being caught. Hence the seasonal average tag return rate must equal the tag return rate from region 1 (i.e., $\lambda_{tg} = \lambda_{t1g}$ when $t = 1, 3, \ldots$).

During season 2 ($t = 2, 4, \ldots$), we adopt the same argument used for the generic model (Appendix 10). Let λ_{tg} be the seasonal average tag return rate. Remember that there are no fish in region 1. Using a *counting argument*, we obtain

$$\sum_{r=2}^{R} P_{tr}(1 - S_{asr})f_{asr}\delta_{trg}\lambda_{trg} \;\;=\;\; P_t(1 - S_{as})f_{as}\delta_{tg}\lambda_{tg}\,, \tag{15}$$

where $P_t = \sum_{r=2}^{R} P_{tr}$. This can be solved explicitly for λ_{tg}.

A problem with this argument is that it depends on δ_{tg}, which in turn depends on the catch-at-age data. There is an alternative counting argument that avoids δ_{tg}. We simply sum the expected number of tags returned over all regions, and divide by the expected number of tagged fish caught. Thus

$$\sum_{r=2}^{R} P_{tr}(1 - S_{asr})f_{asr}\delta_{trg}\lambda_{trg} \;\;=\;\; \left(\sum_{r=2}^{R} P_{tr}(1 - S_{asr})f_{asr}\delta_{trg}\right)\lambda_{tg}\,. \tag{16}$$

We investigated the difference between these two equations in practice and found they led to virtually the same answers; thus, we have decided to use equation (16) in our simulation results.

5.2 Tag-seeding data

The tag-seeding data apply only to season 1, for which we have separate data, so the usual model applies.

5.3 Catch-at-age data

For the pooled catch-at-age data we have only a single catch observation C_t for each time period for the target cohort.

In season 2 ($t = 2, 4, \ldots$), C_t represents the catch taken by observer group $g = 1$. If P_1 is the total population size for the cohort at the start of time period 1, then the expected number of fish caught by component $g = 1$ during time period t (corresponding to age a and season s) is

$$\mu_t = P_1 \left(\prod_{t'=1}^{t-1} S_{a's'} \right) (1 - S_{as}) f_{as} \delta_{t1}$$

where $a' = a(t')$ and $s' = s(t')$ and when $t = 1$, $\prod_{t'=1}^{t-1} S_{a's'} \equiv 1$.

In season 1 ($t = 1, 3, \ldots$), there are no observers but, as in the spatial models, we assume that a proportion δ_{t1}^* of the fish caught are sampled and aged so that a similar equation applies:

$$\mu_t = P_1 \left(\prod_{t'=1}^{t-1} S_{a's'} \right) (1 - S_{as}) f_{as} \delta_{t1}^* \,.$$

5.4 Inference

Each row of the tag-recapture data matrix is treated as multinomial, as usual. The tag-seeding data are treated as binomial. The catch-at-age data are treated as Gaussian with a known coefficient of variation c. Recall that we are also assuming that the proportion of observer catches from the longline fishery (δ_{t1}) in odd time periods and the proportion of catch sampling the surface fishery (δ_{t1}^*) in even time periods are known. The total log-likelihood for the three data sets is maximised to give parameter estimates for M_a, F_{as}, P_1 and λ_{tr}.

Note that in our simulations, we assume that the catch-at-age data in each region has a Gaussian distribution with known coefficient of variation c_r. Before applying the pooled model, we need to specify the coefficient of variation for the total catch-at-age data pooled across regions (c). We have simply chosen to use the average of the c_r's. Although this is not statistically correct, it is adequate for our purposes of comparing performance between the pooled and spatial models.

5.5   Numerical example

We took the same 100 simulated good quality data sets used in generating the results in Table 1 of the numerical example section for the first SBT model, pooled each of the data sets over regions, then estimated the parameters by maximum likelihood using the pooled model. We assumed the coefficient of variation (c) in the catch-at-age data was known and equal to the average of the true regional c_r's (see previous section). Similarly, we assumed the proportions of observer coverage in season 2 (δ_{t1}'s) were known and equal to the weighted average of the true regional δ_{tr1}'s as specified by equation (14), and that the proportion of catch sampling in season 1 (δ_{t1}^*) was known and equal to δ_{t11}^*. The results are summarised in Table 5.

Table 5: True parameter values, mean maximum likelihood estimates and summary statistics from applying the pooled model to 100 sets of data generated from the Markov model with good quality data

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.121	0.312	0.505	$P_1/10^5$	23.500	21.066	27.166	33.699
M_2	0.200	0.142	0.280	0.479	λ_{12}	0.650	0.534	0.659	0.776
F_{11}	0.052	0.114	0.152	0.202	λ_{22}	0.250	0.186	0.259	0.347
F_{12}	0.200	0.162	0.197	0.243	λ_{32}	0.650	0.520	0.661	0.773
F_{21}	0.090	0.079	0.097	0.121	λ_{42}	0.250	0.188	0.253	0.334
F_{22}	0.100	0.083	0.104	0.131	λ_{52}	0.650	0.529	0.662	0.793
F_{31}	0.157	0.144	0.178	0.228	λ_{62}	0.250	0.185	0.247	0.320
F_{32}	0.100	0.083	0.115	0.152	ϵ	0.000	0.000	0.000	0.000

We then repeated the exercise for the same 100 simulated poor quality data sets used in generating the results in Table 2 for the first SBT model. The results are summarised in Table 6.

Table 6: True parameter values, mean maximum likelihood estimates and summary statistics from applying the pooled model to 100 sets of data generated from the Markov model with poor quality data

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.129	0.347	0.553	$P_1/10^5$	10.000	6.640	10.510	16.448
M_2	0.200	0.108	0.323	0.544	λ_{12}	0.650	0.530	0.656	0.791
F_{11}	0.017	0.013	0.020	0.028	λ_{22}	0.250	0.010	0.251	1.000
F_{12}	0.001	0.001	0.001	0.002	λ_{32}	0.650	0.492	0.648	0.770
F_{21}	0.072	0.065	0.083	0.108	λ_{42}	0.250	0.118	0.272	0.592
F_{22}	0.020	0.012	0.022	0.037	λ_{52}	0.650	0.501	0.648	0.782
F_{31}	0.157	0.150	0.201	0.286	λ_{62}	0.250	0.160	0.249	0.373
F_{32}	0.100	0.083	0.127	0.196	ϵ	0.000	0.000	0.000	0.000

Comparing these results with the results in Tables 1 and 2 (obtained using the spatial Markov model) suggests that some large biases can occur from using the pooled data and model. For example, the estimates of M_1 and M_2, which showed small negative biases using the spatial model, exhibit large positives biases using the pooled model and are very imprecise; this is especially true for M_2 and for the scenario with poor quality data. The estimates of the average seasonal fishing mortalities from the spatial model (the F_{as}^* parameters) tended to have small negative biases, whereas the comparable estimates from the pooled model (the F_{as} parameters) tend to have positive biases with larger magnitude, particularly F_{31} and F_{32} in both scenarios and, to an even larger degree, F_{11} in the good quality data scenario. The total population size was well estimated using the spatial model, but is positively biased using the pooled model, and considerably so in the good quality data scenario. The reporting rate estimates are similar using the pooled model as using the spatial model.

We do not show the results from pooling the data simulated according to the site-fidelity model because they are essentially the same as for the Markov model. As discussed in the numerical example section of the site-fidelity model, the longline regions have equivalent population and movement dynamics in both of the scenarios considered, so the data from the two models are almost equivalent. The results of pooling the data over regions will not be so similar for the Markov model and site-fidelity model if the longline regions have different dynamics, and will be scenario-specific.

6   Cross-comparison of models

It is of interest to see what happens when the data are generated by one model and analysed using the other. In fact, if data are generated by the Markov Chain model (model 1), they cannot be analysed using the winter site-fidelity model (model 2). It is quite feasible for model 1 to generate data consisting of, say, regions 1, 2, 1, 3 in time periods 1, 2, 3, 4 respectively, but such data are impossible (have 0 probability) according to model 2. Hence the log-likelihood of such data is $-\infty$.

It is of more interest to see what happens if data are generated according to model 2 and analysed by model 1. The Markov chain model is more commonly used for spatial data, and we are investigating how robust it is to non-Markovian movement.

We took the same 100 simulated data sets used to produce the results in Table 3 (i.e. the good quality data sets generated in accordance with the site-fidelity model), and estimated the parameters by maximising the overall likelihood for the Markov model. The results are summarised in Table 7. Note that the Markov model provides estimates of the transfer probabilities from region 1 to the longline regions at the end of time periods 3 and 5, but we do not specify true values for these parameters for generating data using the site-fidelity model. This explains the dashes for these parameters in the true values column of the table.

Table 7: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the Markov model to 100 sets of good quality data generated from the site-fidelity model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.040	0.206	0.380	$\pi_{41\mid32}$	0.700	—	—	—
M_2	0.200	0.010	0.070	0.259	$P_{11}/10^5$	8.500	5.930	8.377	11.833
F_{111}	0.150	0.112	0.150	0.199	$P_{12}/10^5$	5.000	1.660	4.802	8.419
F_{211}	0.100	0.111	0.140	0.176	$P_{13}/10^5$	5.000	1.844	4.628	8.313
F_{311}	0.200	0.193	0.235	0.301	$P_{14}/10^5$	5.000	2.164	4.693	7.717
F_{122}	0.200	0.128	0.195	0.289	λ_{112}	0.650	0.525	0.655	0.772
F_{222}	0.100	0.048	0.087	0.136	λ_{222}	0.250	0.133	0.247	0.362
F_{322}	0.100	0.036	0.080	0.137	λ_{232}	0.250	0.157	0.251	0.377
F_{123}	0.200	0.137	0.207	0.287	λ_{242}	0.250	0.161	0.256	0.397
F_{223}	0.100	0.057	0.088	0.136	λ_{312}	0.650	0.500	0.596	0.729
F_{323}	0.100	0.035	0.085	0.161	λ_{422}	0.250	0.167	0.275	0.392
F_{124}	0.200	0.131	0.202	0.302	λ_{432}	0.250	0.183	0.269	0.410
F_{224}	0.100	0.055	0.088	0.120	λ_{442}	0.250	0.169	0.274	0.400
F_{324}	0.100	0.036	0.083	0.133	λ_{512}	0.650	0.534	0.648	0.752
$\pi_{12\mid11}(w_2)$	0.333	0.207	0.343	0.497	λ_{622}	0.250	0.195	0.295	0.432
$\pi_{13\mid11}(w_3)$	0.333	0.199	0.323	0.496	λ_{632}	0.250	0.186	0.289	0.419
$\pi_{12\mid21}$	—	0.130	0.336	0.633	λ_{642}	0.250	0.199	0.292	0.433
$\pi_{13\mid21}$	—	0.128	0.333	0.569	F_1^*	0.252	0.209	0.247	0.294
$\pi_{12\mid31}$	—	0.078	0.349	0.806	F_2^*	0.190	0.135	0.161	0.186
$\pi_{13\mid31}$	—	0.055	0.332	0.764	F_3^*	0.257	0.163	0.194	0.234
$\pi_{21\mid12}$	0.900	0.403	0.574	0.843	F_{11}^*	0.052	0.038	0.052	0.068
$\pi_{31\mid12}$	0.900	0.358	0.552	0.778	F_{12}^*	0.200	0.163	0.194	0.227
$\pi_{41\mid12}$	0.900	0.332	0.577	0.875	F_{21}^*	0.090	0.064	0.077	0.093
$\pi_{21\mid22}$	0.800	0.410	0.551	0.698	F_{22}^*	0.100	0.069	0.084	0.099
$\pi_{31\mid22}$	0.800	0.438	0.556	0.701	F_{31}^*	0.157	0.100	0.123	0.149
$\pi_{41\mid22}$	0.800	0.427	0.545	0.694	F_{32}^*	0.100	0.057	0.071	0.091
$\pi_{21\mid32}$	0.700	—	—	—	$P_1/10^5$	23.500	18.893	22.500	26.459
$\pi_{31\mid32}$	0.700	—	—	—	ϵ	0.000	0.000	0.000	0.000

Comparing the above table with Table 3, which contains the results from using the correct model on the same data sets, suggests that the consequences of using the incorrect movement dynamics model can be significant for some parameters. For example, the estimates of M_1 and M_2 are more severely underestimated, M_2

in particular, and the negative biases in the estimates of the transfer probabilities have become considerably larger. The region-specific fishing mortality estimates show larger biases (both positive and negative), and the average yearly and seasonal fishing mortality estimates show larger negative biases; although the estimates may still be adequate depending on the researcher's purposes. The region-specific and total initial population size estimates are slightly underestimated using the incorrect model, but are probably adequate for most purposes. Finally, the reporting rate estimates exhibit some larger biases, but are probably not large enough to be of great concern.

To see how much our findings could vary in a different situation, we repeated the exercise using the 100 simulated data sets used to produce the results in Table 4 (i.e. the poor quality data sets generated in accordance with the site-fidelity model). The results are summarised in Table 8.

Table 8: True parameter values, mean maximum likelihood estimates and summary statistics from fitting the Markov model to 100 sets of poor quality data generated from the site-fidelity model (see text for details)

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
M_1	0.300	0.025	0.283	0.479	$\pi_{41|32}$	0.700	—	—	—
M_2	0.200	0.010	0.106	0.376	$P_{11}/10^5$	8.500	3.058	6.903	10.997
F_{111}	0.020	0.014	0.021	0.029	$P_{12}/10^5$	0.500	0.030	0.843	4.691
F_{211}	0.080	0.066	0.087	0.108	$P_{13}/10^5$	0.500	0.030	1.061	4.644
F_{311}	0.200	0.153	0.209	0.282	$P_{14}/10^5$	0.500	0.030	1.048	5.811
F_{122}	0.001	0.000	0.004	0.015	λ_{112}	0.650	0.519	0.648	0.785
F_{222}	0.020	0.007	0.030	0.079	λ_{222}	0.250	0.050	0.145	1.000
F_{322}	0.100	0.038	0.106	0.226	λ_{232}	0.250	0.050	0.098	1.000
F_{123}	0.001	0.000	0.003	0.013	λ_{242}	0.250	0.050	0.174	1.000
F_{223}	0.020	0.007	0.026	0.092	λ_{312}	0.650	0.507	0.641	0.765
F_{323}	0.100	0.032	0.107	0.237	λ_{422}	0.250	0.050	0.280	0.685
F_{124}	0.001	0.000	0.003	0.016	λ_{432}	0.250	0.050	0.300	0.826
F_{224}	0.020	0.006	0.029	0.109	λ_{442}	0.250	0.050	0.313	0.781
F_{324}	0.100	0.027	0.101	0.246	λ_{512}	0.650	0.505	0.646	0.779
$\pi_{12|11}(w_2)$	0.333	0.000	0.392	0.990	λ_{622}	0.250	0.122	0.283	0.562
$\pi_{13|11}(w_3)$	0.333	0.000	0.339	0.990	λ_{632}	0.250	0.119	0.270	0.484
$\pi_{12|21}$	—	0.004	0.307	0.865	λ_{642}	0.250	0.142	0.296	0.617

Table 8: continued

Par	True	2.5%	Mean	97.5%	Par	True	2.5%	Mean	97.5%
$\pi_{13\mid21}$	—	0.005	0.347	0.836	F_1^*	0.018	0.008	0.016	0.026
$\pi_{12\mid31}$	—	0.092	0.340	0.716	F_2^*	0.092	0.050	0.082	0.113
$\pi_{13\mid31}$	—	0.078	0.315	0.706	F_3^*	0.257	0.149	0.220	0.344
$\pi_{21\mid12}$	0.900	0.010	0.676	0.990	F_{11}^*	0.017	0.007	0.015	0.025
$\pi_{31\mid12}$	0.900	0.010	0.572	0.990	F_{12}^*	0.001	0.001	0.001	0.001
$\pi_{41\mid12}$	0.900	0.010	0.618	0.990	F_{21}^*	0.072	0.035	0.065	0.091
$\pi_{21\mid22}$	0.800	0.271	0.646	0.990	F_{22}^*	0.020	0.011	0.018	0.027
$\pi_{31\mid22}$	0.800	0.166	0.610	0.981	F_{31}^*	0.157	0.092	0.137	0.213
$\pi_{41\mid22}$	0.800	0.192	0.621	0.990	F_{32}^*	0.100	0.055	0.083	0.139
$\pi_{21\mid32}$	0.700	—	—	—	$P_1/10^5$	10.000	6.717	9.856	13.326
$\pi_{31\mid32}$	0.700	—	—	—	ϵ	0.000	0.000	0.000	0.000

It would appear from comparing the above table with Table 4, which contains the results from using the correct model on the same data sets, that the consequences of using the incorrect movement dynamics model are less serious in the scenario with poor quality data. While M_2 and the transfer probabilities are again underestimated to a much greater extent when the wrong model is used, the fishing mortality estimates (regional and average), the initial population size estimates (regional and total), and the reporting rate estimates are all quite similar whether the correct or incorrect model is used.

7 Discussion

In this appendix we developed two spatial models with movement dynamics based roughly on our understanding of the behaviour of SBT. The two models are similar, except the first model allows for Markovian (or memoryless) movement between regions whereas the second model assumes fish remember their winter region and will always return to (or remain in) this region in the winter season. We showed that both models perform well when the data are of sufficient quality (e.g., large sample sizes and small uncertainty), but that large biases and imprecision can occur in many of the parameter estimates when the data are poor.

We also presented a non-spatial model for applying to the data pooled over regions within each time period, which we refer to as the pooled model. By comparing parameter estimates obtained from the spatial models with parameter estimates obtained from the pooled model, we demonstrated that ignoring spatial structure in the data can result in large biases in many of the parameter estimates.

In practice, it is unlikely that the spatial and movement dynamics for the species being studied will be known precisely. As such, we wanted to investigate the consequences of assuming the incorrect spatial model. We considered the case in which the true movement dynamics underlying the data corresponded to the site-fidelity spatial model, but the parameters were estimated using the Markov spatial model. We found that, relative to the parameter estimates using the correct model, some of the estimates using the incorrect model exhibited much larger biases, particularly the natural mortality rates and transfer probabilities. Moreover, the results depended on the scenario being considered, with the consequences of using the incorrect model being more serious in the scenario with good quality data than with poor quality data. This seems logical because if the data are relatively uninformative then the ability to distinguish between the two spatial models would be weak.

It is essential to bear in mind that the results in this appendix are specific to the examples presented, and will vary depending on the particular situation (as evidenced by the two scenarios we considered). The purpose of this appendix was to develop the theory and demonstrate how the models work. In Appendix 12, we more thoroughly investigate the performance of the SBT spatial models for a range of scenarios with spatial and temporal dynamics resembling those for juvenile SBT. In particular, we consider experimental design issues and attempt to address the question of when a non-spatial analysis may be adequate. We also include a more complete cross-comparison of models to see the consequences of assuming the incorrect underlying movement dynamics on parameter estimation.

Appendix 12:

Tagging in a spatial context: design and analysis considerations

Tom Polacheck, J. Paige Eveson and Geoff M. Laslett

FRDC Project 2002/015

Introduction

As shown in Appendices 5, 7, 9, 10 and 11, tagging experiments have the potential to provide direct quantitative estimates for key parameters modelling the population dynamics and assessing the stock status of fishery resources. For some of these parameters, such as movement parameters and natural mortality rates, tagging experiments probably provide the only approach that can yield direct estimates. In addition, they provide the only viable approach for obtaining fishery-independent (e.g., not reliant on CPUE) estimates of population sizes and fishing mortality rate estimates in many fisheries (particularly where the resource is geographically widespread). However, for tagging experiments to be successful there are a number of components requiring "adequate" data (e.g., tag releases, tag returns, catch data, reporting rates, etc.) in order to even be able to undertake the estimation, much less to achieve parameter estimates with acceptable/useful levels of precision and accuracy. It is essential that the overall data requirements of the anticipated analyses be carefully considered in the design and implementation phases of the experiments. If these are not obtainable, then careful consideration should be given to whether the tagging experiments are in fact worth undertaking. In addition, while it can be rather straightforward to list what types of data are required, the question of "how much" is required for each of the different types is complex and becomes more so if there is a need to incorporate complex spatial/temporal population and fishery dynamics.

In the design of any tagging experiment, decisions will be required on the relative amount of effort that should be devoted to the various components. Appendices 5, 7 and 9 present results that examine trade-offs between number of tags released, precision in the catch data and/or level of observer effort for tagging experiments in which systematic incomplete mixing of tagged and untagged fish is not an issue and in which the spatially related parameters are not of interest. Similar types of results can be generated within the context of a spatially and/or temporally disaggregated tagging experiment; however, the dimensionality of the design considerations becomes much larger (e.g., number of tag releases in each area and time period, level of observer coverage and/or catch sampling in each area and time period). In addition, the trade-offs will generally be dependent upon the true values of the population parameters, such as relative abundances in each area and transfer rates between areas. As such, it is not productive or even feasible to present extensive results on the relative trade-offs in design components for general tagging experiments involving complex spatial/temporal components. However, the general estimation and modelling frameworks developed in Appendices 10 and 11 can be used to examine design trade-offs in any potential application in conjunction with specific hypotheses about the spatial/temporal dynamics for the stock and fishery. We would recommend that such calculations be performed prior to undertaking an experiment in order to improve the efficiency of the design and ensure that the experiment has sufficient statistical power to provide useful results. The robustness of such simulations in predicting actual performance will depend upon the set of hypotheses tested. For example, results from an actual experiment may provide unanticipated information about the spatial/temporal dynamics that are outside of the range tested – particularly when there is little previous direct information about the spatial dynamics of the population being tagged.

Although the detailed design trade-offs and design issues will be specific to any particular tagging application, we can investigate more general design questions in situations where spatial effects may need to be considered. We have explored a wide range of simulations using the spatial modelling framework presented in Appendix 11 to gain some understanding of design issues in a spatial context. The purpose of this appendix is to summarize the results from these simulations and the general conclusions they suggest. We stress that the possible range of spatial dynamics that could be considered is vast and the relevance of the results need to be tied to concrete situations. This, combined with limits on what was feasible to explore within the scope of this project, resulted in our focusing on situations where the spatial and temporal dynamics generally resemble those for juvenile SBT (e.g., seasonal migrations between feeding/nursery areas with seasonally distinct fishing grounds). While the results provide general insights on design issues for spatially explicit tagging experiments, caution should be used in extrapolating any conclusions to situations with differing spatial dynamics or within a different experimental design/estimation framework.

Simulation Methods

Estimation Model

The basic model used for our simulations is the spatial tag-recapture and catch-at-age model developed in Appendix 11. The model was designed to have spatial and temporal dynamics resembling the general characteristics of the current SBT fishery and juvenile stock. Thus, the fisheries were divided into two seasonal components – a summer surface fishery and a winter longline fishery. The summer surface fishery occurs only in one area, namely South Australia/the Great Australian Bight (GAB), whereas the winter longline fishery occurs in three areas distinct from where the surface fishery occurs (South Africa, SE Indian Ocean and the Tasman Sea). All fish migrate out of the surface fishing area at the end of the summer season into one of the longline fishing areas, but only a fraction of the fish migrate from the longline fishing areas to the surface fishery area at the end of the winter season (for simplicity, we assume that all migrations between areas occur at the end of a season). Movement rates are assumed to vary among areas and with ages (e.g., a differing percent of fish migrate to the summer surface fishing area with age). Another simplification in the migration dynamics is that we do not allow for direct migration of fish between the longline areas. Thus, over a course of a single year, a fish either remains within a single longline area or it migrates between one of the longline areas and the surface fishery area. Two variations on the basic movement dynamics were considered: one in which movement is represented as a seasonal Markov process (i.e., a fish has no memory with respect to its previous migrations); and one in which there is site-fidelity in the movement dynamics (i.e., a fish that migrates out of a particular longline area will always return to that same area). These two movement models represent two limits of a continuum for modelling seasonal migration. Details of the structural models and the estimation frameworks are given in Appendix 11.

While the seasonal spatial dynamics for SBT are most likely more complex than represented by these models, the overall dynamics are broadly consistent with information from the fisheries, conventional tagging and limited archival tagging. For

example, the archival tagging results suggest a high degree of seasonal migration between the GAB and the winter longline areas, with fish tending to spend most of their winters within a single longline fishing ground. This is particularly true for the Tasman Sea area. For the South African area, fish clearly must move through the SE Indian Ocean when migrating between South Africa and the GAB. However, such movements appear to be relatively rapid based on the limited number of archival tag returns to date in which fish tagged in the GAB have moved to South Africa. Also, two very recent archival tag returns indicate that some older juveniles that spend their summers off South Africa may move into the SE Indian Ocean in the following winter. Nevertheless, the spatial dynamics considered here should provide an adequate framework for considering the major design issues for SBT-like situations.

Because the model was constructed to represent an SBT-like fishery situation, its estimation framework incorporates observer data for estimating reporting rates in the longline fisheries and tag seeding data for estimating reporting rates in the surface fishery. The model is comprised of three independent likelihood components – one for the tag-return data, one for the catch-at-age data, and one for the tag seeding data (the observer data are incorporated into both the tag-return and catch-at-age components). The overall likelihood is maximized to provide estimates of the fishing mortality rates by area and age, natural mortality rates by age, initial abundance (i.e., abundance at the time of first tagging) by area, and transition rates among areas by age and season. Note that the model has been developed for application to a single cohort of fish tagged at consecutive ages, and therefore year and age become equivalent terms (i.e., referring to parameters differing by age is the same as differing by year). Although it is possible to generalize the model to more cohorts (see Appendix 16), we only consider applications to data from one cohort of fish in our simulations.

Simulation Scenarios

In all our simulations, we assumed a tagging experiment with three release and recapture years involving a single cohort of fish. The population and movement dynamics and experimental designs that we considered are described below.

Population/movement dynamics

For generality and ease of reference, we will subsequently refer to the GAB as area 1 and to the South African, SE Indian Ocean and Tasman Sea areas as areas 2 to 4 respectively. Additionally, we will refer to the summer season as season 1 and the winter season as season 2. Thus, in summary, all fishing in area 1 occurs in season 1 and all fishing in the areas 2 to 4 occur in season 2; all fish move out of area 1 at the end of season 1 (no other transitions between areas occur at the end of season 1), and fish in areas 2 to 4 may choose to move into area 1 at the end of season 2 or stay in their current area.

The set of scenarios we explored cover a wide range of values for the spatial and fishery dynamics (i.e., for the area and age-specific fishing mortality rates, the initial abundance of fish in each area, and the age-specific transition probabilities governing the movements between areas). This was done because the underlying dynamics can affect the properties of the resulting estimates. As such, different dynamics can have implications for the design of tagging experiments. We explored scenarios for each of

the two different movement dynamics models developed in Appendix 11 (i.e., the Markov movement model and the site-fidelity movement model). In general, the results were similar with respect to design considerations. As such, results are only presented for the Markov movement model, except when we explore the importance of having the correct movement dynamics model. For the purposes of this appendix, we considered Markov movement scenarios with variation in the following four basic dimensions for the population dynamics and defined a range of options within each of these dimensions:

1. Fishing mortality rates (low, medium and high);
2. Initial abundance in each area (high proportion in area 1; equal in all areas; 100% in area 1; high proportion in areas 1 and 2);
3. Movement from area 1 to areas 2-4 at the end of season 1 (equal proportions move to each area; only a small proportion move to area 2)
4. Movement from areas 2-4 to area 1 at the end of season 2 (low for area 2 and high for areas 3 and 4; high for all areas).

Even within this range of options, a full cross of the options would yield 48 (3x4x2x2) different scenarios, so for the purpose of presentation we limited the number of options to 16. The set of scenarios for which results are presented are described in general terms in Table 1 and the detailed specification of the parameter values for each scenario are provided in Table 2. In order to limit the range of scenarios to a reasonable level, the same values for natural mortality at age were used in all the scenarios presented here.

For the site-fidelity seasonal migration model, we present results with variation in these same basic dimensions but for a more limited number of scenarios. These are defined in Table 3.

Tagging Design and Data Collection

In the simulations presented, we considered two strategies for the tag releases in each year: (1) tags released in all areas; (2) tags released only in area 1 (i.e., the GAB). In any given year, tagging in area 1 always occurs at the beginning of season 1, and tagging in areas 2 to 4 always occurs at the beginning of season 2. In most cases we kept the total number of tag releases the same (2400 per year) in order to limit the number of simulations to a manageable amount. The effect of differing levels of tag releases was examined in Appendices 5, 7 and 9 in a non-spatial context; similar effects would be expected in the spatial context, so we only present a few limited results in which the total number of releases was varied. In simulations in which the total number of releases was 2400 per year and tags were released in all areas, 1500 of the releases were in area 1 and 300 were in each of the three remaining areas. When we varied the total number of releases, we used the same relative proportions in each area.

We also examined the effect that the quality of data for estimating reporting rates and catch-at-age had on the results. We defined two data quality scenarios, which we refer to as high and low (Table 4). The high quality scenario represents a high, but achievable commitment to the data collection process. For example, it specifies observer coverage of 75%. Although this is a relatively high level, near 100% observer coverage exists for a number of major fisheries. The low quality scenario is

closer to the current data collection situation for SBT. It is in fact overly optimistic with respect to observer coverage for the longline fisheries as it uses a value of 10%. This is the current target level agreed on by the CCSBT (Anon. 2001), but observer coverage for the main longline fisheries has been substantially below this target (Anon. 2004). Moreover, in the simulations, 10% observer coverage is assumed to constitute a random sample of 10% of the catch. In reality, observer sampling is a multi-stage process with observers spending up to several months on a single vessel. As such, effective sample size or coverage would be less than the percent of the catch actually observed, particularly when observer coverage is low.

Simulation Results and Discussion

All of the results presented are based on 100 simulations for each scenario. We have summarized the results in terms of the relative bias and the coefficient of variation (CV). Relative bias is defined as the difference between the mean parameter estimate (over the 100 runs) and the true value divided by the true value, multiplied by 100 to be expressed as a percent.

Is a spatial dimension necessary?

The introduction of a spatial dimension into the experimental design will clearly complicate the analysis and, in most instances, the implementation of tagging experiments. As such a critical question in the design of such experiments is whether a spatial component is necessary. Clearly, if the purpose of the experiment is to get estimates of movement rates or area-specific population sizes and fishing mortality rates, then the spatial component cannot be ignored. However, if the primary purpose of the experiment is to provide overall population-wide estimates of the parameters of primary interest for stock assessment purposes (i.e., natural and fishing mortality and population size), then whether a non-spatial design and analysis is sufficient is worth considering. As such, a focus of the results presented here is to compare the population-wide estimates of mortality and abundance when they are derived from the spatial estimation model versus when they are derived from a non-spatial estimation model applied to data pooled over areas (see Appendix 11 for details of the two estimation approaches).

If the spatial component is ignored, the expectation is that the resulting estimates will be biased if tagged and untagged fish are not completely mixed across the entire spatial range. The extent of bias would be expected to depend on the extent of non-mixing and the relative differential in abundances and fishing mortality rates by area/season. Tables 5-10 provide comparisons of the percent relative bias in the population-wide estimates derived from applying the spatial Markov model to data generated according to the 16 spatial Markov scenarios defined in Table 2 versus the estimates obtained from pooling the data spatially and applying a non-spatial model. Comparisons are provided for the estimates of the total initial population size (at age 1), the natural morality rate at ages 1 and 2, and the overall fishing mortality rates for some selected ages and seasons. Results are presented for the high and low quality data scenarios and for the two different release strategies (tagging in all areas and tagging only in area 1). It should be noted that in scenario 11, all fish are initially in area 1 in season 1, they have equal probability of transitioning into each of areas 2-4 at the end of the season, and 100% return to area 1 at the end of season 2. In essence, this scenario constitutes complete mixing of tagged and untagged fish. While

unrealistic in terms of SBT dynamics, this scenario was included both as a check on the software (i.e., estimates using the pooled data should be unbiased in this case) and to see if there is any loss in precision by using a spatial model when in fact complete mixing exists.

Initial total population size estimate

Focusing first on the initial total population size estimates (Table 5), the results confirm that ignoring the spatial dimension will result in significant biases. Thus, all of the scenarios show significant biases in their pooled estimates of initial total population size except for scenario 11, as expected. The degree of bias is roughly related to the extent of non-mixing, particularly for the high data quality scenarios. For example, although significantly different from zero, the bias in the pooled estimates for scenario 15 in the high data quality scenario is only 3.6%. This scenario has nearly complete mixing (i.e., 85% of the initial population is in area 1 at season 1, fish move equally into areas 2-4 at the end of season 1 at each age, and a high and equal proportion return from areas 2-4 to area 1 at the end of season 2). Thus, the primary difference in the capture probability for tagged and untagged fish is for those fish tagged at the beginning of season 1 in each year. Similarly, scenario 13 has relatively small bias (5.0%). The increased spatial heterogeneity in this case stems from the relatively low proportions of the population that migrate between area 1 and area 2. As the differential in the mixing proportion increases and with lower proportions of the initial population size in area 1, the biases in the pooled estimates increased to as much as 36% even with high quality data. For the low quality data scenarios, the biases increased substantially for the pooled estimates.

It should be noted that the fishing mortality rates within each of areas 2-4 was always the same at any age in the scenarios presented. Having substantially different fishing mortality rates among these areas would also have the potential to introduce substantial differences in recapture probabilities and further increase biases in the pooled estimates, particularly with large differences in the transition rates among areas.

In contrast to the pooled estimates, the spatially-based estimates of total initial population size are unbiased in the high data quality scenarios and remain essentially unbiased even in the low quality data scenarios[1]. This result was somewhat surprising given that substantial biases often exist in the area-specific estimates of initial population size (Table 11). Even more unexpected was that tagging only in area 1 also yielded unbiased estimates of total initial population size, with no real loss in precision as long as the same total number of tags was released. However, as with tagging in all areas, the area-specific population size estimates were often highly biased with tagging in only area 1, and in many scenarios (especially those where the initial distribution of fish was equal in all areas instead of being concentrated in area

---

[1] In a few low quality data scenarios, there is a small apparent bias of the total initial population size for the spatially-based estimates. Most of this is due to the fact the true bias is less precisely estimated for these scenarios because the increased variance in the individual estimate of P results in larger standard errors for the estimates of the bias with the same number of simulations. Thus, in many cases the estimated bias is not significant. Nevertheless, in some cases some estimation bias may be introduced in the poor data quality scenarios because of small sample size effects (e.g., the number of tags returned may be quite low).

1) these biases were significantly larger than when there was tagging in all areas (Table 11).

Overall, these simulation results indicate that if spatial heterogeneity is likely to exist it is important to include a spatial dimension within the analysis phase to achieve unbiased estimates of population sizes. In terms of the precision of the estimates, there appears to be little or no cost for including a spatial dimension when in fact complete mixing exists, as indicated by the fact that the CVs for scenario 11 were essentially the same for the overall spatial estimates and the pooled estimated.

Fishing mortality rates

A comparison of the average seasonal fishing mortality estimates from the spatial model (see section 3.5.3 of Appendix 11) and the corresponding pooled estimates from the non-spatial model provides similar general conclusions as those for the estimates of total initial population size (Tables 6-8[2]). Thus, the estimates of bias in the average spatially-based estimates are in almost all cases insignificant (exceptions are discussed below), while the biases for the pooled estimates are frequently significant and in some cases substantial. Extremely large biases are seen in some of the pooled estimates at age 1 for season 1 (i.e., over 300%) (Table 6). Similarly high biases are not seen in the estimates at ages 2 and 3 in season 1 (not shown), although the bias for scenario 12 is always greater than 100%. The high biases in the age 1 estimates in season 1 are associated with scenarios in which the initial population is distributed equally in all four areas. As such, these scenarios would have the largest heterogeneity in recapture probabilities given that fishing is only taking place in area 1 during season 1, plus all tag releases at this point have been in area 1 and no mixing has occurred into other areas. As with the estimates of total initial population size, the results for the average fishing mortality rates were similar whether tags were released in all areas or only within area 1.

The spatially-based estimates of average fishing mortality appear to be unbiased in all of the high data quality scenarios, as well as in many of the low data quality scenarios but with two general exceptions: (1) a consistent negative bias ranging from around 8 to 19% is seen in the estimates of the average fishing mortality rates for season 1 in the scenarios with a high proportion of the population initially in area 1 (i.e., the odd numbered scenarios) and (2) negative biases on the order of ~10% are seen in all the estimates of fishing mortality rates for scenario 11. It is particularly unclear where this latter bias stems from, recalling that there is no spatial heterogeneity in recapture probabilities in scenario 11 (i.e., there is complete mixing). However, it appears that the bias is related to the low fishing mortality rates. In other simulations we conducted that were similar to scenario 11 but with higher fishing mortality rates, these biases decreased. When the fishing mortality rates are low, the expected number of tag returns is low and the biases may reflect these small sample sizes. In addition, with low quality data for estimating reporting rates, potential confounding exists between estimates of mortality and reporting rates (i.e., it is hard to distinguish whether lack of returns represents low fishing mortality or low reporting rates). The fact that there is relatively substantial bias (23%) in the pooled estimate for the age 1 season 1 fishing mortality rate for this complete mixing scenario would also suggest the potential for

---

[2] Note that results are shown for only 3 of the 6 age and season combinations but the general behaviour of the estimates was similar for those not shown.

confounding when the data for estimating reporting rates are poor and fishing mortality rates this scenario are in fact low. Interestingly, the biases in the pooled estimates of fishing mortality rates for other ages are all small (<3%), with the exception of the estimate for age 2 in season 1 where it was 9%.

Natural mortality rates

The results for the estimates of natural mortality also indicate the spatially-based estimates are generally unbiased or less biased than the pooled estimates (Tables 9-10). As was seen in the non-spatial model results (Appendices 5 and 7), the estimates of natural mortality rates are often poorly estimated and the estimate for ages 2 and older (M_2) can be biased with small number of releases. It is interesting to note that M_2 appears to be the one parameter for which there is a cost in terms of increased biased for including a spatial dimension when in fact complete mixing exists (e.g., compare the results for scenario 11 with low data quality). The high CVs for the estimates for M_2, particularly within the low data quality scenarios (on the order of 55% or greater), suggest that the estimates would have low reliability. However, this appears not so much a matter of whether a spatial estimation framework is used, but a more general problem of estimating M_2 within a Brownie estimation framework (i.e., high CVs exist for the pooled M_2 estimate of scenario 11 in which no spatial heterogeneity exists, and also for the non-spatial M_2 estimates in Appendices 5 and 7). M_2 tends to be the least precise of the main parameters of interests. It is highly correlated with the estimates of fishing mortality for ages 2 and 3. The estimate of M_2 is dependent upon the contrast between the relative return rate (number of return per release) for age 2 and 3 corrected for any age specific differences in reporting rates. As such, in situations with small numbers of returns and large uncertainty in reporting rates (e.g., the poor data quality scenarios and/or low fishing mortalities), it is not surprising that M_2 is poorly estimated. Additional simulations that we conducted along with the results in Appendices 5 and 7 suggest that higher precision in M_2 can be achieved through a combination of a higher number of releases, high quality data for estimating reporting rates, and conducting tagging experiments on multiple cohorts assuming that the age-specific natural mortality rates are constant across years.

Reporting Rate Estimation - Data Quality

Comparing the CVs for the parameter estimates in Tables 5-10 (as well as in Tables 12-17, which will be discussed later) for the high and low data quality scenarios indicates, as expected, that there is a substantial reduction in the precision of the estimates when there is low quality data for estimating reporting rates and the age distribution of the catch. However, given the relatively large differences in the data quality between these two scenarios (e.g., 10% versus 75% observer coverage, and 30% versus 10% CVs for the catch data), it was somewhat surprising that the differences were not even larger. For the estimates of total initial population size and average fishing mortality rates, the estimated CVs were relatively low in many cases, even under the low data quality scenarios (i.e., they had sufficient precision to be informative within a stock assessment/management advice context). This was particularly true for the estimates of the total initial population size, for which the CV did not exceed 21% in the scenarios with tag releases in all areas and were only somewhat higher (i.e., a maximum of 26%) when releases only occurred in area 1.

However, in most actual situations, over dispersion in the tag return and observer data would likely result in higher CVs (see discussion).

Distribution of Tag Releases

Comparison of the results for scenarios in which there were releases in all areas with those in which releases only took place in area 1 indicate that both release scenarios, in general, yielded similar results for the population-wide parameter estimates in terms of bias and precision, with perhaps a small loss in precision in some cases where tags were only released in area 1. This result was surprising as it was not even certain prior to conducting these simulations that all the parameters would be identifiable under such a tag release strategy. These simulations indicate that the distribution of tag releases is not critical and that it is not essential to have releases in all areas to getting meaningful estimates at the population-wide scale. However, we note that with alternative spatial/movement dynamics, simply releasing tags in one area can result in an intractable estimation problem with some parameters becoming unidentifiable (e.g., the generic spatial model – see Appendix 13). As such, it is essential in designing a tagging experiment that the range of potential hypotheses for the possible spatial/movement dynamics be carefully considered to ensure that the distribution of tag releases will in fact allow for the intended parameters to be estimated. Moreover, in such cases, the relative distribution of releases among areas may have an implication both for the precision and bias of the population-wide parameter estimates. We would recommend as a matter of course that simulation testing be conducted prior to conducting a large-scale complex tagging experiment to ensure that the targeted distribution of tag releases can in fact yield estimates across the full range of hypotheses for the spatial dynamics of the stock and fishery. Nevertheless, the results presented here demonstrate that robust estimates of total population size do not require optimal release by area and that it may not be necessary to release tags in all areas. This conclusion increases the potential usefulness of tagging experiments in fishery contexts since in many instances tagging in some areas/fisheries may be logistically problematical and/or highly expensive (e.g., in high seas, off-shore areas in tuna longline fisheries).

Number of Tags Releases

To a rough first approximation, the variance for many of the parameters of interest will be proportional to the number of tags returned. As such, ideally it is important to ensure that sufficient tags are released so that that there are a "reasonable" number of expected returns from each of the release/return strata in the model. While in general, the mortality rates and population size estimates should be asymptotically unbiased[3], with small sample sizes this is not necessarily the case. Some biases were evident in the non-spatial situation (Appendix 5) when the number of releases was small, and would also be expected within a spatial estimation framework.

It is not possible to provide generic guidelines on what constitutes sufficient numbers of releases both to avoid substantial biases and to achieve reasonable levels of

---

[3] Because of the way we have modelled the variance in the catch-at-age data as having a constant CV, the estimates of population sizes from our simulations have a small statistical bias (see Appendix 5). However, this would not be the expected situation for most actual applications because estimates of the catch variance would generally be available.

precision because this will be highly dependent upon both the actual population and fishery dynamics over time and space and the reporting rates. Nevertheless, consideration of the effect of differing sample sizes for the scenarios considered here can provide some indication of what level of releases may be required, at least within an "SBT-like" context. Figures 1-4 provide examples of the effects of three differing release levels on the bias and CV for the total initial population sizes and average mortality rate estimates for three of the population dynamics scenarios from Table 3. (Results are only shown for three scenarios so that the trends for individual scenarios are distinguishable). Note that the different levels of releases were simply multiples of the number of releases used in the results above with tagging in all areas. As would be expected, the biases and CVs decrease with increasing numbers of releases, and trends are similar to those seen in the non-spatial context (Appendices 5 and 7). In some cases, substantial small sample size biases exist with low numbers of releases and there is a marked degradation in performance in the low data quality scenarios relative to the high quality ones, particularly at low sample sizes. For example, the estimates of the total initial population sizes with a total of only 500 tag releases can have substantial biases of over 20% in the low data quality scenarios, with CVs of 35-40% (Figure 1). This compares to biases of 0-4% and CVs of 13-16% for the same population dynamics scenarios but with high quality data. The results suggest that with high quality data for estimating the catch-at-age and reporting rates, releases on the order of 1000 tags per year would be sufficient to obtain estimates of the initial population size and fishing mortality rates with CVs of ~20% or less. In contrast, with low quality data, not even 2400 releases per year can achieve similar levels of precision nor are the biases always small. The estimates of natural mortality rates are more poorly estimated, and to achieve CVs of less than 40% would entail releases on the order of 2400 per year for the high data quality scenarios and substantially greater for the low data quality scenarios. The poor estimation of natural mortality was also seen in the non-spatial results (Appendices 5 and 7). As discussed in Appendices 5 and 7, the CVs on all parameter estimates from these simulations are likely to be underestimates.

How important is having the right spatial/movement dynamics model?

There are three components that to need to be represented when modelling the spatial dynamics of a population: (1) specification of the movement dynamics, (2) specification of the spatial structure, and (3) specification of the time step.

Movement Dynamics

In Appendices 10 and 11, we developed three different models for representing possible movements between areas:
1. A box model representation of unrestricted age-specific movement/transition rates between all areas.
2. A box model representation of an age-specific seasonal Markov (memoryless) migration process with restricted movement between some areas.
3. A box model representation of an age-specific seasonal migration process with restricted movement between some areas and with site fidelity.

Note that model 2 could be considered a subset of model 1, which also assumes Markovian transitions (i.e., that each transition made by a fish is independent of its previous transitions). It is important to emphasise that these are only three of a

potentially much wider class of models that could be used to represent movement dynamics.

Model 1 without any restrictions on the transition probabilities between areas is a truly generic model. As long as the underlying spatial structure and time periods in the model are appropriate and transition probabilities are allowed to vary with age and year, this model is probably capable of reasonably representing most types of movement dynamics. However, the cost in terms of the number of parameters and demands for tagging in many areas is high – particularly as the number of areas and time periods increases. This could easily lead to intractable estimation and design problems. In addition, such a generic structure provides little basis for prediction, as it assumes no consistency in the movement dynamics with time. Thus, imposing some form of spatial and/or temporal structure on the transition probabilities is highly desirable. The structure imposed will depend upon the population being modelled; however it is not unreasonable or unrealistic to expect that there are some consistent underlying biological processes that determine how and when animals move (although it may be useful or necessary to also consider environmental covariates). We note that assuming movement occurs according to a random diffusion process imposes large restrictions on the transition probabilities because the transition probabilities become a function of the distance between and geometry of the boxes used in spatial representation[4].

Models 2 and 3 were constructed to represent an idealized representation of SBT-like spatial dynamics and take advantage of known (or assumed) prior information on the movement dynamics of the stock to reduce the parameterization of the transition probabilities. The two models can have very different implications for the relative abundances in the different areas under differential fishing pressures. For example, under model 3, localized depletion can occur in one of the longline areas if fishing effort is localized in that area, while this would not occur under model 2 as long as the transition probabilities at the end of season 1 from area 1 into that area were non-zero. As such, it is not immediately clear whether misspecification of the underlying movement dynamics would produce unreliable estimates at the population-wide level. The site-fidelity scenarios (Table 3) were used to explore this question.

Tables 12-16 compare the estimates of total initial population size and average fishing and natural mortality rates that were obtained for these scenarios when the correct underlying movement dynamics model was used in the estimation and when the incorrect Markov movement model was used. Results are presented for the two tag release strategies (i.e., releases in all areas and releases only in area 1). When there is tagging in all areas, the results suggest that the consequences of using the wrong movement dynamics model are relatively small except for the estimates of natural mortality rates (particularly M_1), which had substantially greater negative biases when the Markov estimation model was used. There was also some tendency for the fishing mortality rates to be slightly biased, probably as a consequence of the biases in natural

---

[4] The implication of random movement is that if density of animals per unit of space is initially equal in all areas then the density would remain equal with movement over time in the absence of spatial differences in mortality. Thus, in a box model framework, the transition probabilities would need to be set to achieve this. For example, in the case of two boxes, the transition rates between the two boxes would be inversely proportional to their sizes under a random movement assumption and only one movement parameter would need to be estimated.

mortality. The discrepancies between models tended to be greater when tagging was only carried out in area 1.

It is not clear to what extent these results reflect general properties of the underlying movement dynamics model or the particular scenarios chosen for the simulations. Some of the effects appear to be scenario related. For example, scenario 2 had considerably poorer performance than all of the other scenarios in estimating initial population size when tagging was only in area 1, and this was independent of which spatial estimation model was used (i.e., a relative bias of 22% or 37% compared to a maximum of 8% for the other scenarios with high quality data). When tagging occurred in all areas, the initial population size estimate for scenario 2 had low bias (i.e., less than 2%) similar to other scenarios tested. On the other hand, the poor estimation of natural mortality when the wrong movement dynamics model is used (particularly with respect to bias) appears to be a more general feature of the underlying movement dynamics. Thus, substantial biases in M_1 are seen in all scenarios and these biases persist even if sample sizes are doubled (Table 17). Overall, these simulation results indicate that the consequences of having the "correct" movement dynamics model will be parameter and situation dependent. All of this further emphasises the need to undertake simulation testing in the design and analysis of complex large-scale tagging programs.

It should be noted that when there was tagging in all areas the AIC statistic (Akaike 1974) was always greater for the fit with the Markov movement estimation model than the fit with the site-fidelity movement model for each of the 100 simulations within a scenario. Moreover, in nearly all cases the difference was substantive (i.e., in only two simulations was the difference less than 2, and in only 1% was the difference less than 5). In contrast, when tagging only occurred in area 1, there was around an equal probability that the AIC statistic would be greater for the fit with the Markov movement estimation model than the fit with the site-fidelity movement model (Table 18). There was only a slight improvement in the ability to correctly distinguish between the two models when the number of releases was doubled (Table 18). These results indicate that tag releases on the order of 2400 would provide sufficient statistical power to distinguish between these two movement models, but only if tagging occurs in all areas.

Spatial structure

The appropriate number of areas, their size and their location to use in defining the spatial structure will depend upon the biology of the species and the dynamics of the fishery. The aim of defining a spatial structure is to achieve equal recapture probabilities of tagged and untagged fish (i.e., complete mixing) within each area. The smaller the areas, the more likely this will be achieved. However, with increasing numbers of areas, the number of parameters that need to be estimated (i.e., initial population sizes and fishing mortality rates by area, plus transition probabilities between areas) increases rapidly. As such, the choice of the spatial structure will require balancing the achievement of a feasible/tractable estimation problem with the goal of obtaining complete mixing of tagged and untagged fish within each spatial unit. In some cases where there are major concentrations of and/or discontinuities in the spatial distribution of the fishing effort (such as with SBT), these provide obvious areas. Also, existing information on migration and movements (e.g., from archival and previous conventional tagging) can provide insights into appropriate spatial

stratification[5]. However, where fishing effort is more continuously distributed, the definition of areas will be more arbitrary. Evaluating the consequences of having the incorrect spatial structure is not possible in a general context. If the selection of areas is a concern, the sensitivity of the results to alternative spatial structures should be investigated.

We note that spatial effects can be mitigated in the design and implementation of tag releases by randomizing and distributing tagging effort. Over large scales, this is less feasible because actually randomizing tagging effort spatially over large scales is both difficult to implement and inefficient (i.e., potentially large amounts of effort spent in areas of low abundance or zero abundance). However, ensuring that tagging effort is spread out can reduce concerns about spatial effects at smaller scales.

Time step

The question of the appropriate time step to represent the movement dynamics is analogous to that of the appropriate spatial structure. Thus, within a time step, the probability of recapturing a tagged or untagged fish should be equal. As such, the time step needs to reflect both seasonal movements or migrations and temporal patterns of fishing effort. The time step also needs to reflect the distribution of tag releases. It is important to take into account the fact that the information that allows natural and fishing mortality rates to be separated within a Brownie model context comes from the comparison of tag returns from temporally separate releases of the same cohort. As such, in developing the experimental design of the tag releases, there may be advantages in considering tagging at multiple times within a year – particularly if fishing is highly seasonal (see Hearn et al. 2003; Frusher and Hoenig 2001).

General Discussion

The results presented in this appendix indicate that including a spatial component within the estimation model is important to prevent biases in population-wide estimates unless complete mixing of tagged and untagged fish occurs. The amount of bias will be approximately related to the degree of non-mixing. As such, unless there is prior information that indicates that non-mixing is not an issue, testing appropriate spatially based estimators is important to ensure that robust estimates are obtained from Brownie-type tagging experiments. In most cases, such prior information does not exist, but the expectation, at least in populations with a large spatial range, would be that complete mixing is unlikely over the full spatial range. While this conclusion is not surprising given the underlying assumptions in mark-recapture estimators, it does have implications for the design of tagging experiments and the efforts required for their analysis.

In terms of experimental design, the need to incorporate a spatial component in the analysis emphasises the need for spreading releases across the geographic range of the population. Although the results presented for both the Markov and site-fidelity movement scenarios indicate that having releases in all areas and time periods is not essential for a number of population dynamics scenarios, this was only achievable by

---

[5] Note that in interpreting previous conventional tagging results, consideration needs to be given to the distribution of fishing effort and the probability of recovering tags from different areas, particularly if estimates of reporting rates are not available.

imposing some highly informative structure on the movement dynamics (e.g. no direct movement between longline areas). The range of spatial dynamics models that can be tested will depend upon the spatial/temporal pattern of releases. The broader the distribution of tag releases the more flexibility is possible in the analysis phase. When designing tagging experiments, it is critical to consider what is the appropriate range of spatial dynamics hypotheses/models and ensure that the release strategy will allow these to be explored and tested. As emphasised above, simulation testing of the design is important to ensure that robust estimates can be achieved, particularly in situations with complex spatial structures.

Comparison of the results for the site-fidelity movement scenario when the wrong movement dynamics model is used in the estimation with results when the correct one is used provides an example of the enhanced value of tagging in all areas. With tagging in all areas, the results at the population level were relatively robust to which movement model was used and, moreover, it was possible to statistically distinguish which model was the correct underlying one. In contrast, with tagging in only one area, the results were sensitive to which estimation model was used and it was not possible to reliably distinguish which model was correct.

Knowing the form of the underlying movement dynamics can have important management implications. For example, the site-fidelity model can lead to localized depletion while the Markov movement model will not. As discussed above, these two models represent two ends of a continuum for estimating seasonal migration. Furthermore, the Markov movement and site-fidelity models used here are only two of a potentially large number of movement dynamics model (e.g., environmental or localized density driven). In general, knowledge of the dynamics of fish movement/migration mechanisms are poor, as is the development of models for representing alternatives at the population-wide scale. These are questions that could benefit from additional research.

The simulation results suggest that natural mortality rates are particularly difficult to estimate accurately and precisely without large numbers of releases and high quality supporting data for estimating reporting rates and the age distribution of the catch. Natural mortality rates were particularly poorly estimated in the site-fidelity scenarios with 2400 tag releases per year. Only when the numbers were increased to 4800 were reasonably reliable estimates achieved with tagging in all areas. In any actual tagging experiment, more than a single cohort will be tagged in any given year (since cohorts are usually mixed spatially and capture gear catches a range of size/age classes). Although it was beyond the scope of the current project to extend these simulations to tagging of multiple cohorts over extended numbers of years, such multi-cohort tagging would be expected to improve the estimation of natural mortality rates if they are unchanging with time. Such improvements were seen in the non-spatial context (Appendices 5 and 7). As such, a multi-cohort tagging program over a number of years may provide a more viable approach to achieving reliable natural mortality rates than attempting to achieve very large numbers of releases for a single cohort.

Multi-cohort tagging would also offer other advantages. It would allow for additional estimates of initial population sizes and fishing mortality rates for more than a single cohort. It would also allow testing for separability of fishing mortality rates (i.e. the decomposition of fishing mortality into an age and year effect). Such separability

assumptions and their parameterization are an essential component in most catch-at-age stock assessments. If separable assumptions were found to be applicable, this should improve the overall estimation performance for the remaining parameters – particularly given the high covariance among them. Multi-cohort tagging would also provide a more viable approach for improving the area-specific parameters of the model. As noted above, the population-wide estimates were generally well estimated in the simulation results even when the area-specific estimates of population size, fishing mortality rates and transition probabilities were not. The full spatial model for a single cohort has a large number of parameters to estimate relative to the number of data points (i.e., the number of tag returns in each area and time period). Being able to use the information from multi-cohort returns for estimating the transition probabilities within a common time-independent framework would provide an approach for reducing the parameterization and could provide a powerful approach for gaining insights into the actual spatial dynamic processes through the testing of alternative model structures for the movement dynamics (e.g., the Markov versus site-fidelity model). In this context, we note that, regardless of the underlying movement dynamics, there is likely to be variability among years and it may be worth considering developing random effects models as a more appropriate way to represent transition probabilities for multi-cohort tagging data.

Finally, the high quality data scenarios, as expected, always provided more robust estimates than the low data quality scenarios. Somewhat surprising was that the differences in performance were not larger. With poor information on reporting rates we had anticipated that there might be much greater confounding between reporting and mortality rates than was apparent in the results. In general, the parameters in the low quality data scenarios remained unbiased. The CVs, while being substantially greater than in the high data quality scenarios, were often within what might be considered acceptable limits in a stock assessment context (i.e., less than 30%). However, we reiterate that the CVs in the simulations are likely to be more optimistic than what would be realised in practice due to several factors – in particular, over-dispersion in the tag-recapture probabilities (i.e., more variability than assumed by the model; see discussion in Appendix 5). One source of overdispersion in the tag-recapture data would be differences in catchability and selectivity among vessels and areas.

Such differences in catchability and selectivity would also result in the catch-at-age data having more sampling variability than if a simple random sample of the catches were available. In cases where observers perform most or all of the catch sampling, the increased variability in the catch-at-age data is likely to be substantive at low levels of observer coverage (e.g., observer coverage of 10% of the total catch may result in sampling from only a small number of vessels/fishing trips operating in high seas longline fleets where a single trip may be several months or more). Thus, when considering the level of effort in the design phase to put into the estimation of the age distribution of the catch, it is important to ensure that the actual level of sampling variability in the catch-at-age data is being reflected. This will be fishery specific and further work is needed for almost all fisheries on how best to characterize it. In terms of the estimation framework developed here, this might best be done by developing a functional relationship between the actual observer/sampling coverage and the effective coverage in terms of simple random sampling.

Literature Cited

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, **19**: 716-723.

Anon. 2001. CCSBT. Report of the Seventh Annual Meeting of the Commission. 18-21 April 2001. Sydney, Australia.

Anon. 2004. CCSBT. Report of the Extended Scientific Committee for the Ninth Meeting of the Scientific Committee. 13-16 September 2004. Seogwipo City, Jeju, Republic of Korea.

Frusher, S.D. and J.M. Hoenig. 2001. Strategies for improving the precision of fishing and natural mortality estimates from multiyear tagging models: a case study. Mar. Freshwater Res. 52: 1649-1655.

Hearn,W.S., K.H. Pollock and E.N. Brook. 1998. Pre and post season tagging models: estimation of reporting rate and fishing and natural mortality rates. Can. J. Fish. Aquat. Sci. 55: 199-205.

Table 1: Descriptive definition of the 16 population dynamic scenarios for the Markov movement dynamics model simulations – see Table 2 for detailed specification of all parameter values.

Scenario	Fishing mortality	Movement from area 1 into areas 2-4 at the end of season 1	Movement from areas 2-4 into area 1 at the end of season 2	Initial Distribution Among areas
1	Medium	Low to area 2 and equal 3 & 4	Low area 2, high 3 and 4	85% in area 1
2	Medium	Low to area 2 and equal 3 & 4	Low area 2, high 3 and 4	Equal all areas
3	Medium	Equal to areas 2 to 4	Low area 2, high 3 and 4	85% in area 1
4	Medium	Equal to areas 2 to 4	Low area 2, high 3 and 4	Equal all areas
5	High	Equal to areas 2 to 4	Low area 2, high 3 and 4	85% in area 1
6	High	Equal to areas 2 to 4	Low area 2, high 3 and 4	Equal all areas
7	Low	Equal to areas 2 to 4	Low area 2, high 3 and 4	85% in area 1
8	Low	Equal to areas 2 to 4	Low area 2, high 3 and 4	Equal all areas
9	Low	Low to area 2 and equal 3 & 4	Low area 2, high 3 and 4	85% in area 1
10	Low	Low to area 2 and equal 3 & 4	Low area 2, high 3 and 4	Equal all areas
11	Low	Equal to areas 2 to 4	100% all areas	100% in area 1
12	High	Equal to areas 2 to 4	Low area 2, high 3 and 4	High areas 1 &2
13	Medium	Low to area 2 and equal 3 & 4	High all areas	85% in area 1
14	Medium	Low to area 2 and equal 3 & 4	High all areas	Equal all areas
15	Medium	Equal to areas 2 to 4	High all areas	85% in area 1
16	Medium	Equal to areas 2 to 4	High all areas	Equal all areas

Table 2: Definition of the parameter values and input values for the 16 population dynamic scenarios for the Markov movement dynamic model simulations. M = natural mortality; F = fishing mortality; TR = transfer rates; P = population size at age 1.

Parameter				Scenario							
Par	Age	area	Seas	1	2	3	4	5	6	7	8
M	1	-	-	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300
	2+	-	-	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
F	1	1	1	0.100	0.100	0.100	0.100	0.200	0.200	0.020	0.020
		2	2	0.050	0.050	0.050	0.050	0.100	0.100	0.001	0.001
		3	2	0.050	0.050	0.050	0.050	0.100	0.100	0.001	0.001
		4	2	0.050	0.050	0.050	0.050	0.100	0.100	0.001	0.001
	2	1	1	0.075	0.075	0.075	0.075	0.150	0.150	0.080	0.080
		2	2	0.075	0.075	0.075	0.075	0.150	0.150	0.020	0.020
		3	2	0.075	0.075	0.075	0.075	0.150	0.150	0.020	0.020
		4	2	0.075	0.075	0.075	0.075	0.150	0.150	0.020	0.020
	3	1	1	0.050	0.050	0.050	0.050	0.100	0.100	0.200	0.200
		2	2	0.100	0.100	0.100	0.100	0.200	0.200	0.100	0.100
		3	2	0.100	0.100	0.100	0.100	0.200	0.200	0.100	0.100
		4	2	0.100	0.100	0.100	0.100	0.200	0.200	0.100	0.100
TR	1	1 to 2	1	0.050	0.050	0.333	0.333	0.333	0.333	0.333	0.333
		1 to 3	1	0.475	0.475	0.333	0.333	0.333	0.333	0.333	0.333
	2	1 to 2	1	0.100	0.100	0.333	0.333	0.333	0.333	0.333	0.333
		1 to 3	1	0.450	0.450	0.333	0.333	0.333	0.333	0.333	0.333
	3	1 to 2	1	0.200	0.200	0.333	0.333	0.333	0.333	0.333	0.333
		1 to 3	1	0.400	0.400	0.333	0.333	0.333	0.333	0.333	0.333
	1	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
		4 to 1	2	0.900	0.900	0.900	0.900	0.900	0.900	0.900	0.900
	2	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800
		4 to 1	2	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800
	3	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700
		4 to 1	2	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700
P	1	1	1	850000	333000	850000	333000	850000	333000	850000	333000
		2	1	50000	333000	50000	333000	50000	333000	50000	333000
		3	1	50000	333000	50000	333000	50000	333000	50000	333000
		4	1	50000	333000	50000	333000	50000	333000	50000	333000

Table 2 (Continued).

Parameter				Scenario							
Par	Age	Area	Seas	9	10	11	12	13	14	15	16
M	1	-	-	0.300	0.300	0.300	0.300	0.300	0.300	0.300	0.300
	2+	-	-	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200
F	1	1	1	0.020	0.020	0.020	0.200	0.100	0.100	0.100	0.100
		2	2	0.001	0.001	0.001	0.100	0.050	0.050	0.050	0.050
		3	2	0.001	0.001	0.001	0.100	0.050	0.050	0.050	0.050
		4	2	0.001	0.001	0.001	0.100	0.050	0.050	0.050	0.050
	2	1	1	0.080	0.080	0.080	0.150	0.075	0.075	0.075	0.075
		2	2	0.020	0.020	0.020	0.150	0.075	0.075	0.075	0.075
		3	2	0.020	0.020	0.020	0.150	0.075	0.075	0.075	0.075
		4	2	0.020	0.020	0.020	0.150	0.075	0.075	0.075	0.075
	3	1	1	0.200	0.200	0.200	0.100	0.050	0.050	0.050	0.050
		2	2	0.100	0.100	0.100	0.200	0.100	0.100	0.100	0.100
		3	2	0.100	0.100	0.100	0.200	0.100	0.100	0.100	0.100
		4	2	0.100	0.100	0.100	0.200	0.100	0.100	0.100	0.100
TR	1	1 to 2	1	0.050	0.050	0.333	0.333	0.050	0.050	0.333	0.333
		1 to 3	1	0.475	0.475	0.333	0.333	0.475	0.475	0.333	0.333
	2	1 to 2	1	0.100	0.100	0.333	0.333	0.100	0.100	0.333	0.333
		1 to 3	1	0.450	0.450	0.333	0.333	0.450	0.450	0.333	0.333
	3	1 to 2	1	0.200	0.200	0.333	0.333	0.200	0.200	0.333	0.333
		1 to 3	1	0.400	0.400	0.333	0.333	0.400	0.400	0.333	0.333
	1	2 to 1	2	0.100	0.100	1.000	0.100	0.900	0.900	0.900	0.900
		3 to 1	2	0.900	0.900	1.000	0.900	0.900	0.900	0.900	0.900
		4 to 1	2	0.900	0.900	1.000	0.900	0.900	0.900	0.900	0.900
	2	2 to 1	2	0.100	0.100	1.000	0.100	0.800	0.800	0.800	0.800
		3 to 1	2	0.800	0.800	1.000	0.800	0.800	0.800	0.800	0.800
		4 to 1	2	0.800	0.800	1.000	0.800	0.800	0.800	0.800	0.800
	3	2 to 1	2	0.100	0.100	1.000	0.100	0.700	0.700	0.700	0.700
		3 to 1	2	0.700	0.700	1.000	0.700	0.700	0.700	0.700	0.700
		4 to 1	2	0.700	0.700	1.000	0.700	0.700	0.700	0.700	0.700
P	1	1	1	850000	333000	1000000	450000	850000	333000	850000	333000
		2	1	50000	333000	0	450000	50000	333000	50000	333000
		3	1	50000	333000	0	50000	50000	333000	50000	333000
		4	1	50000	333000	0	50000	50000	333000	50000	333000

Table 3: Definition of the parameter values and input values for the 7 population dynamic scenarios for the site-fidelity movement dynamics model simulations. M = natural mortality; F = fishing mortality; TR = transfer rates; P = population size at age 1.

Par	Age	area	Seas	1	2	3	4	5	6	7
M	1	-	-	0.300	0.300	0.300	0.300	0.300	0.300	0.300
	2+	-	-	0.200	0.200	0.200	0.200	0.200	0.200	0.200
F	1	1	1	0.100	0.100	0.100	0.100	0.200	0.200	0.200
		2	2	0.050	0.050	0.050	0.050	0.100	0.100	0.100
		3	2	0.050	0.050	0.050	0.050	0.100	0.100	0.100
		4	2	0.050	0.050	0.050	0.050	0.100	0.100	0.100
	2	1	1	0.075	0.075	0.075	0.075	0.150	0.150	0.150
		2	2	0.075	0.075	0.075	0.075	0.150	0.150	0.150
		3	2	0.075	0.075	0.075	0.075	0.150	0.150	0.150
		4	2	0.075	0.075	0.075	0.075	0.150	0.150	0.150
	3	1	1	0.050	0.050	0.050	0.050	0.100	0.100	0.100
		2	2	0.100	0.100	0.100	0.100	0.200	0.200	0.200
		3	2	0.100	0.100	0.100	0.100	0.200	0.200	0.200
		4	2	0.100	0.100	0.100	0.100	0.200	0.200	0.200
TR	1	1 to 2	-	0.050	0.050	0.333	0.333	0.333	0.333	0.333
		1 to 3	-	0.475	0.475	0.333	0.333	0.333	0.333	0.333
	1	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.900	0.900	0.900	0.900	0.900	0.900	0.900
		4 to 1	2	0.900	0.900	0.900	0.900	0.900	0.900	0.900
	2	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.800	0.800	0.800	0.800	0.800	0.800	0.800
		4 to 1	2	0.800	0.800	0.800	0.800	0.800	0.800	0.800
	3	2 to 1	2	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		3 to 1	2	0.700	0.700	0.700	0.700	0.700	0.700	0.700
		4 to 1	2	0.700	0.700	0.700	0.700	0.700	0.700	0.700
P	1	1	1	850000	333000	850000	333000	850000	333000	450000
		2	1	50000	333000	50000	333000	50000	333000	450000
		3	1	50000	333000	50000	333000	50000	333000	50000
		4	1	50000	333000	50000	333000	50000	333000	50000

Table 4: Definition of the parameter values used in the high and low quality data scenarios for estimating tag reporting rates and catch-at-age.

Parameter	High Quality	Low Quality
Reporting rate (all areas and seasons)	0.75	0.25
Observer cover (in areas 2-4, season 2)	0.75	0.19
Catch CV (all areas/fisheries)	0.10	0.30
Number of seeded tags (area 1)	300	50

Table 5: Comparison of the relative bias and coefficient of variation in the initial population size for the total spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging All areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	10.0	-0.4	6.5	6.2	8.2	-0.6	5.8	6.3	6.9
	2	13.3	-2.3	18.9	17.0	19.2	6.0	19.0	22.8	17.6
	3	10.0	-0.7	10.9	6.4	8.0	-1.3	12.2	5.5	6.6
	4	13.3	-3.9	14.9	6.6	7.1	-2.7	17.7	6.7	7.8
	5	10.0	-0.3	15.1	5.1	6.4	0.1	17.4	5.1	5.9
	6	13.3	-0.7	29.6	5.3	6.1	-0.9	30.9	5.8	7.0
	7	10.0	0.5	32.5	8.1	9.5	-0.1	32.5	7.2	9.2
	8	13.3	1.2	36.8	9.1	10.2	1.7	36.3	8.8	9.6
	9	10.0	0.3	10.5	8.2	10.5	0.3	12.0	9.1	9.8
	10	13.3	-0.8	16.5	7.5	8.8	5.0	17.6	17.1	8.4
	11	10.0	0.7	0.4	7.0	8.6	-1.2	-0.7	6.6	7.2
	12	10.0	-1.1	28.1	5.1	6.4	-1.2	31.2	10.2	6.9
	13	10.0	-0.3	5.0	6.2	7.6	-1.3	4.4	6.1	6.9
	14	13.3	-2.3	15.9	18.8	20.6	-3.3	13.1	16.1	16.9
	15	10.0	-0.8	3.6	6.4	7.9	-1.4	4.4	5.8	6.4
	16	13.3	-2.7	8.8	6.1	7.1	-2.7	10.3	6.2	7.3
Low	1	10.0	-2.9	11.7	15.4	18.4	-0.7	13.3	18.0	22.5
	2	13.3	-1.9	21.6	18.6	19.2	4.2	19.9	24.3	16.9
	3	10.0	1.0	23.8	17.4	23.6	-1.4	22.3	17.6	21.0
	4	13.3	-5.5	22.2	15.9	17.7	-5.4	21.7	17.3	19.4
	5	10.0	-4.3	21.5	15.5	17.8	-6.8	20.6	14.1	16.2
	6	13.3	-6.1	32.7	17.3	19.0	-6.5	35.3	16.1	16.2
	7	10.0	0.5	36.3	21.5	28.5	1.4	41.1	23.0	28.1
	8	13.3	-4.7	33.4	18.8	21.6	-5.5	35.4	24.2	27.9
	9	10.0	3.5	22.8	16.9	20.8	5.8	23.5	20.1	19.9
	10	13.3	-2.1	20.5	20.9	23.8	4.0	23.3	26.2	24.9
	11	10.0	2.0	9.9	19.6	26.8	-3.6	5.8	19.3	23.0
	12	10.0	-7.3	34.2	16.3	17.8	-7.0	39.7	18.6	19.7
	13	10.0	-3.1	9.8	15.1	19.2	-1.6	11.9	16.4	20.6
	14	13.3	-4.2	15.7	18.3	19.3	-3.1	14.7	17.3	15.9
	15	10.0	-1.2	14.0	16.6	22.3	-4.0	12.3	16.6	19.0
	16	13.3	-5.7	15.0	15.3	17.3	-6.7	14.3	15.8	17.7

Table 6: Comparison of the relative bias and coefficient of variation in the fishing mortality rate for age 1 in season 1 for the average spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	0.084	-1.4	20.9	10.1	10.9	-2.8	19.7	10.0	8.4
	2	0.024	-5.0	325.9	35.4	34.2	-7.8	342.8	36.4	33.4
	3	0.084	-1.2	19.9	9.1	10.3	-2.0	21.1	11.4	8.2
	4	0.024	3.2	322.5	10.5	9.7	0.0	326.5	11.8	7.8
	5	0.167	-1.3	20.4	9.1	7.4	-0.4	22.0	9.7	6.4
	6	0.046	0.2	336.9	10.4	8.1	2.2	341.4	11.0	6.9
	7	0.017	-3.2	24.9	11.5	20.9	-1.3	22.5	10.9	18.3
	8	0.005	-2.5	343.2	12.0	21.5	0.5	338.3	12.9	17.5
	9	0.017	-2.9	16.8	12.1	21.3	-2.6	16.9	12.3	18.4
	10	0.005	0.3	310.7	11.9	24.5	-3.3	309.9	19.1	19.1
	11	0.020	-4.0	-1.1	12.3	18.7	-2.5	0.4	11.0	17.6
	12	0.085	0.3	137.6	8.7	8.9	-2.8	19.7	10.0	8.4
	13	0.084	-1.5	21.2	10.2	10.1	-1.7	19.7	9.9	8.0
	14	0.024	-4.9	331.1	35.5	35.0	-0.7	338.1	34.4	32.8
	15	0.084	-1.3	17.0	9.0	11.0	-2.5	18.0	10.9	7.8
	16	0.024	1.9	316.1	10.1	11.3	-0.3	320.2	11.5	8.2
Low	1	0.084	-10.9	26.2	27.5	32.7	0.6	140.4	15.2	7.7
	2	0.024	-2.7	348.5	34.9	39.0	-10.1	28.6	25.6	32.3
	3	0.084	-18.2	30.0	30.7	37.4	-4.1	349.5	39.2	27.7
	4	0.024	0.7	337.1	36.8	34.2	-13.7	27.7	35.7	30.0
	5	0.167	-8.2	29.0	31.3	30.5	-2.5	354.2	32.3	31.9
	6	0.046	-2.1	371.6	34.4	34.4	-5.4	39.8	31.4	35.1
	7	0.017	-8.6	29.3	35.1	55.3	-1.5	371.6	31.9	29.0
	8	0.005	2.5	331.0	33.8	49.8	-11.3	38.2	32.1	53.1
	9	0.017	-13.8	17.2	34.9	53.2	2.3	340.2	39.3	39.9
	10	0.005	-3.5	339.6	35.1	47.9	-19.5	20.8	34.1	40.4
	11	0.020	-18.8	23.6	30.6	52.8	-3.7	336.3	34.6	43.7
	12	0.085	-1.6	156.0	31.6	27.5	-7.8	20.7	37.4	46.0
	13	0.084	-10.6	27.9	27.9	34.9	-9.3	28.0	24.9	33.3
	14	0.024	-0.7	362.2	34.0	46.2	0.1	348.7	35.1	27.8
	15	0.084	-17.1	27.1	29.7	37.6	-11.3	25.5	35.3	27.2
	16	0.024	-1.2	334.6	33.3	36.1	-1.7	345.5	33.3	32.4

Table 7: Comparison of the relative bias and coefficient of variation in fishing mortality rate for age 2 in season 2 for the average spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging All areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	0.075	-0.2	3.6	6.1	6.0	-3.8	5.3	11.9	7.7
	2	0.075	-2.1	8.2	19.0	20.5	-11.4	11.4	28.3	22.2
	3	0.075	-0.6	6.2	7.2	6.6	2.8	12.0	20.2	7.5
	4	0.075	0.2	6.1	7.2	6.6	1.7	11.2	21.1	7.3
	5	0.150	-0.3	4.8	5.9	5.8	4.3	10.0	20.1	7.1
	6	0.150	-0.1	5.4	6.2	5.8	5.4	10.6	19.3	6.0
	7	0.020	-1.5	8.5	9.4	8.9	2.0	18.3	23.7	9.8
	8	0.020	-0.7	9.5	8.8	8.8	4.3	18.9	23.2	9.7
	9	0.020	-0.4	6.5	7.7	9.9	-2.2	10.2	11.3	9.3
	10	0.020	0.0	5.2	7.4	8.4	-6.8	8.9	24.1	10.4
	11	0.020	-2.8	0.4	6.6	8.3	-2.7	1.1	7.8	7.9
	12	0.150	0.0	4.7	7.2	5.6	6.6	8.9	30.8	6.4
	13	0.075	-0.6	2.5	6.5	6.5	-2.1	3.6	8.4	7.9
	14	0.075	-2.3	5.8	20.0	20.1	-3.9	9.0	20.2	21.4
	15	0.075	-0.2	3.3	6.4	6.9	0.2	5.5	8.6	7.7
	16	0.075	-0.4	2.7	6.4	6.6	-1.0	4.5	8.6	7.3
Low	1	0.075	-2.3	6.7	14.9	16.7	-8.2	7.4	20.8	20.0
	2	0.075	-6.9	0.0	19.0	19.1	-14.0	3.2	25.8	19.4
	3	0.075	-4.9	4.4	14.9	14.9	-2.9	9.4	22.9	15.3
	4	0.075	-3.4	4.2	20.8	20.5	3.9	11.9	30.4	22.0
	5	0.150	-2.3	6.5	16.7	16.3	-0.1	10.7	21.5	18.1
	6	0.150	-0.9	6.1	15.4	14.7	-2.4	10.3	24.4	17.5
	7	0.020	-1.6	16.6	23.5	27.4	-4.4	21.5	29.4	33.1
	8	0.020	0.1	14.2	23.7	25.9	-0.1	23.6	28.6	28.0
	9	0.020	-5.4	6.1	26.6	30.6	-10.7	8.6	24.5	28.0
	10	0.020	-3.7	9.7	22.4	26.0	-9.2	13.7	30.7	29.2
	11	0.020	-10.5	0.2	22.3	27.9	-7.5	-1.0	25.2	25.8
	12	0.150	-0.5	6.0	15.8	14.6	8.0	9.9	39.4	14.5
	13	0.075	-2.4	4.4	17.0	18.6	-5.1	4.5	21.9	20.5
	14	0.075	-5.4	1.2	17.2	18.5	-6.8	3.9	21.4	20.1
	15	0.075	-6.8	0.3	16.7	16.1	-6.0	4.8	18.9	18.6
	16	0.075	-4.6	2.4	16.4	19.6	-2.5	6.4	20.0	21.0

Table 8: Comparison of the relative bias and coefficient of variation in the fishing mortality rate for age 3 in season 2 for the average spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging All areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	0.10	-1.7	4.0	10.1	9.4	-5.6	9.2	19.9	12.4
	2	0.10	-4.6	4.5	23.7	21.7	-13.0	13.3	35.4	27.5
	3	0.10	-1.7	7.0	12.4	10.6	6.4	18.3	38.4	12.9
	4	0.10	-0.1	7.3	11.1	9.1	5.7	18.1	41.8	10.4
	5	0.20	0.0	6.1	9.0	6.8	11.6	15.8	37.5	7.7
	6	0.20	-1.7	6.2	8.9	7.4	11.5	17.0	33.6	8.7
	7	0.10	-4.0	22.0	12.8	8.9	6.8	52.2	46.6	10.0
	8	0.10	-2.1	26.6	13.5	9.2	11.7	57.2	42.2	9.9
	9	0.10	0.2	17.7	9.9	8.6	-4.0	31.1	20.9	10.3
	10	0.10	0.0	17.6	10.7	8.9	-8.7	29.5	35.5	11.1
	11	0.10	-3.8	0.1	9.2	8.7	-8.6	-0.2	11.1	8.7
	12	0.20	0.1	5.3	10.1	7.0	15.9	14.9	50.6	8.6
	13	0.10	-3.1	1.7	10.8	9.6	-3.9	5.5	15.9	12.7
	14	0.10	-2.6	4.3	23.7	21.7	-3.0	11.7	32.2	28.3
	15	0.10	-0.5	4.0	11.8	11.1	-1.0	8.5	16.5	13.4
	16	0.10	-1.0	3.8	9.6	9.3	-2.3	8.6	15.3	10.6
Low	1	0.10	-0.4	9.1	21.5	20.6	-6.0	14.9	26.4	24.0
	2	0.10	-7.4	3.1	21.8	19.7	-14.6	10.2	31.6	26.0
	3	0.10	-4.4	4.9	21.7	19.3	3.5	17.3	32.2	23.5
	4	0.10	3.1	13.9	21.2	22.1	17.7	27.3	45.4	24.4
	5	0.20	0.2	9.4	21.4	17.7	2.6	16.3	33.9	19.5
	6	0.20	1.1	7.3	19.6	17.0	1.1	16.2	36.3	20.0
	7	0.10	-0.8	26.6	23.7	18.9	3.6	58.5	40.1	22.9
	8	0.10	-0.2	28.2	24.6	19.9	8.3	68.7	39.7	25.3
	9	0.10	-3.8	16.6	23.3	21.5	-7.5	30.8	26.7	23.4
	10	0.10	-2.1	20.0	24.2	19.3	-7.1	37.1	37.0	22.8
	11	0.10	-12.1	-2.6	20.6	20.5	-14.8	-2.0	24.7	23.0
	12	0.20	0.7	7.2	20.7	16.5	9.3	15.4	45.8	20.1
	13	0.10	-1.8	7.4	20.1	20.1	-1.7	11.6	27.5	23.7
	14	0.10	-6.7	1.5	20.8	20.2	-6.8	7.2	30.2	26.8
	15	0.10	-4.9	2.6	21.6	19.7	-6.5	7.2	25.8	24.6
	16	0.10	0.4	10.0	21.4	21.8	0.0	15.6	27.7	25.1

Table 9: Comparison of the relative bias and coefficient of variation in natural mortality rate for age 1 for the spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	0.30	0.5	12.8	28.4	26.9	-3.4	9.9	30.8	22.1
	2	0.30	-7.7	9.6	55.0	45.9	-7.9	8.7	48.7	35.6
	3	0.30	0.5	31.8	27.5	19.2	1.8	32.5	39.9	15.4
	4	0.30	-8.5	25.0	28.0	19.7	-6.3	24.8	47.4	17.7
	5	0.30	4.2	31.4	22.0	17.4	5.0	29.2	32.7	14.6
	6	0.30	0.4	28.8	20.9	15.0	4.1	26.4	30.2	13.0
	7	0.30	1.7	64.1	33.9	18.2	3.1	58.2	48.5	15.5
	8	0.30	5.4	69.0	29.8	15.8	13.7	62.4	41.1	13.9
	9	0.30	0.4	14.1	31.9	28.9	-0.5	14.2	29.8	22.8
	10	0.30	-3.3	13.6	32.0	27.1	-6.0	15.1	31.7	21.6
	11	0.30	-0.4	1.8	27.8	26.9	-1.0	0.1	24.6	22.8
	12	0.30	0.4	27.2	20.9	14.8	0.6	25.5	29.0	11.7
	13	0.30	1.1	7.9	28.2	27.2	-1.2	6.1	27.7	22.4
	14	0.30	-8.4	4.4	55.5	46.8	-6.5	4.5	49.0	37.5
	15	0.30	0.3	7.1	25.7	22.7	-2.1	7.0	21.7	17.6
	16	0.30	-7.8	0.2	25.4	23.1	-9.2	0.7	25.7	21.3
Low	1	0.30	1.1	16.1	50.9	43.8	-1.0	14.8	52.9	38.9
	2	0.30	-10.4	7.5	53.0	43.1	-11.4	7.4	44.8	35.1
	3	0.30	3.0	34.4	49.4	36.1	0.0	33.2	53.7	30.0
	4	0.30	-9.6	23.3	58.1	42.4	-5.1	25.0	51.3	32.7
	5	0.30	-5.6	26.6	48.0	35.0	-9.8	19.3	50.1	33.1
	6	0.30	-8.6	20.3	57.5	42.1	-7.1	22.4	55.5	35.4
	7	0.30	1.9	58.5	51.8	30.4	4.6	58.6	52.9	27.6
	8	0.30	-7.3	48.3	64.3	39.6	-9.3	45.9	63.3	32.7
	9	0.30	5.9	25.0	47.0	38.2	2.0	24.3	44.7	32.2
	10	0.30	-7.0	14.6	55.3	45.9	-14.5	15.4	52.6	33.6
	11	0.30	3.2	10.9	39.3	37.2	-3.7	5.1	39.1	33.8
	12	0.30	-6.9	24.4	47.9	35.6	-5.2	22.7	49.6	32.6
	13	0.30	1.3	10.5	49.2	43.6	1.7	11.7	46.3	38.6
	14	0.30	-12.4	1.1	53.9	45.5	-9.5	2.2	45.1	37.5
	15	0.30	-3.7	7.5	46.6	41.7	-8.3	6.3	47.7	37.6
	16	0.30	-17.2	-3.5	57.6	50.8	-17.0	-2.4	51.1	40.2

Table 10: Comparison of the relative bias and coefficient of variation in natural mortality rate for ages 2 and older for the spatial (Spat) and pooled (Pool) estimates for the 16 population dynamics scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			Spat	Pool	Spat	Pool	Spat	Pool	Spat	Pool
High	1	0.2	-6.6	9.9	42.7	34.4	-16.9	22.8	62.9	31.8
	2	0.2	-6.6	14.3	69.1	54.7	-12.4	32.8	90.0	53.3
	3	0.2	-4.9	16.0	42.5	32.4	-3.0	44.2	77.1	28.2
	4	0.2	-0.9	16.4	36.2	27.8	-6.9	44.2	79.1	24.2
	5	0.2	-3.1	10.7	29.0	22.6	1.5	34.4	62.2	18.1
	6	0.2	-6.4	10.1	28.4	22.4	4.0	37.4	53.4	19.3
	7	0.2	-13.1	60.0	44.3	19.6	-9.9	119.3	84.7	14.0
	8	0.2	-8.3	70.1	41.9	16.8	-0.9	125.4	71.5	13.0
	9	0.2	1.4	53.7	35.4	20.5	-12.9	76.0	60.0	18.1
	10	0.2	0.1	55.4	35.6	19.7	-21.2	74.9	60.6	18.3
	11	0.2	-12.1	-4.6	35.1	31.4	-25.9	-3.9	45.9	30.4
	12	0.2	-2.2	10.4	26.1	21.1	-3.8	35.2	65.0	20.7
	13	0.2	-10.5	4.1	45.2	35.8	-14.8	14.0	58.8	34.6
	14	0.2	-6.6	8.9	64.8	55.4	-7.4	24.2	78.6	56.8
	15	0.2	-3.0	9.9	41.5	34.2	-4.2	22.7	50.7	32.7
	16	0.2	-1.6	12.1	32.6	29.0	-7.1	24.9	51.4	28.2
Low	1	0.2	-3.1	18.6	62.2	50.6	-14.8	31.8	78.1	49.3
	2	0.2	-10.8	9.6	68.0	54.3	-15.4	26.5	78.9	55.4
	3	0.2	-12.4	7.2	82.2	64.7	1.0	38.4	80.9	56.9
	4	0.2	2.5	22.9	64.7	54.0	16.7	51.8	73.6	46.2
	5	0.2	-1.9	17.0	56.7	43.9	-2.7	37.6	69.5	41.0
	6	0.2	-3.0	8.8	53.7	47.2	-15.2	24.7	78.4	47.4
	7	0.2	0.6	72.3	66.9	39.4	-1.4	122.9	85.2	29.2
	8	0.2	2.9	78.6	63.9	32.9	10.6	135.9	80.5	27.6
	9	0.2	-7.2	45.9	68.3	40.4	-11.1	68.1	77.2	36.4
	10	0.2	-1.0	56.3	61.3	33.8	0.2	78.2	68.8	32.8
	11	0.2	-36.1	-17.1	70.1	52.9	-44.3	-15.0	98.2	61.1
	12	0.2	-4.1	9.6	53.0	42.5	-12.8	28.3	74.6	41.6
	13	0.2	-6.9	14.5	64.2	52.4	-8.6	23.9	71.6	50.8
	14	0.2	-15.3	2.9	69.3	56.8	-12.5	17.9	79.8	58.7
	15	0.2	-17.1	0.2	83.0	67.1	-19.4	16.2	90.6	65.0
	16	0.2	-1.2	17.7	63.6	55.3	-6.3	29.5	76.5	52.2

Table 11: Relative bias in the area-specific estimates of the initial population size for the 16 population dynamic scenarios defined in Table 2 for the Markov movement dynamics model simulations, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	Tagging all areas Bias in estimate for area				Tagging only area 1 Bias in estimate for area			
		1	2	3	4	1	2	3	4
High	1	-4.4	0.9	22.8	42.2	-4.7	89.2	-11.8	-9.9
	2	-2.2	1.4	-0.7	-7.8	-2.4	81.0	-23.9	-30.6
	3	-2.3	9.2	7.1	9.2	-3.9	29.4	22.7	-12.1
	4	0.5	-8.4	-0.3	-7.2	-2.9	20.6	-7.5	-20.9
	5	-1.1	4.5	17.7	-8.8	-1.0	27.7	-3.2	-6.7
	6	0.3	1.3	-2.6	-2.0	0.7	17.7	-18.5	-3.7
	7	-9.0	48.3	71.9	42.1	-5.2	69.4	2.6	14.8
	8	-2.9	4.1	1.1	2.3	-0.2	77.0	-40.5	-29.4
	9	-7.7	12.0	52.2	72.6	-5.8	109.7	19.5	-23.9
	10	4.0	-0.1	-8.2	1.1	1.8	81.0	-29.7	-33.1
	11	-11.7	-	-	-	-12.1			
	12	-0.1	-1.7	-8.7	2.5	-1.7	-1.4	2.5	2.4
	13	-4.9	9.1	22.0	45.8	-4.5	53.8	-11.4	7.9
	14	-2.7	6.2	2.1	-14.7	-2.0	28.6	-8.8	-30.9
	15	-2.1	5.6	2.5	11.1	-3.7	9.3	13.6	12.4
	16	0.6	2.3	-7.5	-6.2	-3.1	-4.8	-2.5	-0.2
Low	1	-21.4	47.6	172.2	84.9	-20.3	201.2	90.0	39.8
	2	-5.4	6.8	-12.3	3.2	-6.8	62.4	-23.5	-15.1
	3	-24.8	195.2	158.9	88.4	-25.3	255.2	84.1	62.7
	4	-2.2	-5.9	-6.1	-7.9	-7.5	28.2	-21.0	-21.3
	5	-20.3	88.8	53.2	117.8	-23.7	143.0	46.9	77.3
	6	-10.7	-3.7	0.1	-10.3	-8.9	23.1	-22.7	-17.4
	7	-22.2	157.7	126.2	102.8	-23.4	248.6	80.2	96.4
	8	10.4	-2.4	-19.5	-7.4	0.1	28.8	-34.8	-16.1
	9	-18.1	62.5	175.0	139.1	-24.8	291.2	123.3	123.4
	10	2.1	9.2	-8.9	-10.9	0.8	91.6	-33.0	-43.3
	11	-35.4	-	-	-	-33.8	-	-	-
	12	-15.0	-6.8	42.3	7.4	-19.3	-1.4	18.9	27.6
	13	-21.0	60.0	141.7	93.7	-19.8	143.3	88.9	72.7
	14	-5.6	3.6	-13.3	-1.6	-6.7	23.1	-10.7	-18.3
	15	-24.1	142.4	127.2	115.5	-24.9	110.4	99.0	133.9
	16	-2.6	3.3	-14.2	-9.2	-6.6	-6.0	-6.4	-7.7

Table 12: Comparison of the relative bias and coefficient of variation in the total initial population size estimates from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	10.0	-0.4	-1.2	6.2	6.1	3.7	8.3	12.3	12.8
	2	13.3	-0.5	-1.7	18.3	17.9	22.4	37.3	53.9	31.3
	3	10.0	-0.9	-1.7	6.6	6.6	-1.4	-1.1	6.1	5.9
	4	13.3	-3.6	-4.4	6.1	6.1	-2.8	-2.7	6.5	6.8
	5	10.0	-0.2	-1.2	5.1	5.1	0.2	0.2	4.9	5.2
	6	13.3	-1.0	-1.8	4.9	4.9	-0.5	-0.4	5.4	5.5
	7	10.0	-1.5	-3.7	5.2	5.2	-1.8	0.3	12.8	16.2
Low	1	10.0	-3.8	-4.5	16.9	16.7	4.3	11.8	32.1	26.8
	2	13.3	-3.1	-4.8	19.4	19.3	16.7	27.1	54.3	31.9
	3	10.0	-1.4	-2.4	17.4	17.2	-3.1	-1.7	18.3	18.6
	4	13.3	-5.8	-6.0	18.2	18.4	-5.6	-4.2	18.7	19.7
	5	10.0	-5.2	-7.1	15.2	14.9	-6.0	-5.6	18.7	17.6
	6	13.3	-7.4	-8.8	16.2	16.1	-8.6	-6.9	15.3	15.3
	7	10.0	-6.3	-9.3	16.9	17.4	-10.0	-3.2	21.5	20.4

Table 13: Comparison of the relative bias and coefficient of variation in the estimates of natural mortality rate at age 1 from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.30	-0.3	-26.1	27.9	37.6	-25.7	-40.0	57.4	80.9
	2	0.30	-8.5	-28.8	56.8	66.1	-10.8	-19.9	52.0	61.9
	3	0.30	-2.7	-23.4	29.7	37.5	-12.4	-32.0	74.0	105.8
	4	0.30	-9.6	-30.9	28.2	37.4	-25.4	-44.4	88.3	124.9
	5	0.30	4.3	-13.6	22.3	26.9	-2.3	-15.3	54.9	79.3
	6	0.30	-1.7	-19.4	20.6	25.6	-8.5	-21.3	60.7	82.1
	7	0.30	-0.6	-21.4	21.3	27.6	-0.6	-9.4	34.2	47.1
Low	1	0.30	-0.2	-20.6	49.9	59.7	-19.0	-36.0	73.0	92.1
	2	0.30	-11.2	-33.8	52.8	68.9	-15.9	-22.5	52.0	52.6
	3	0.30	-1.3	-19.0	52.4	62.6	-18.5	-44.5	79.5	113.8
	4	0.30	-14.3	-30.8	61.7	74.7	-23.4	-43.5	93.7	103.3
	5	0.30	-8.2	-24.4	49.2	59.2	-27.1	-49.7	80.7	112.4
	6	0.30	-11.4	-25.3	58.1	66.4	-34.7	-52.3	93.5	112.8
	7	0.30	-9.3	-27.0	48.9	59.4	5.1	-15.4	45.8	61.0

Table 13: Comparison of the relative bias and coefficient of variation in the estimates of natural mortality rate at ages 2 and older from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.20	-6.4	-15.0	42.1	48.1	-12.2	-22.6	59.5	66.8
	2	0.20	1.2	-16.9	62.6	76.9	0.4	-10.7	78.0	85.2
	3	0.20	1.0	-10.7	40.6	47.0	2.5	-13.3	46.9	58.4
	4	0.20	0.2	-9.6	33.7	38.3	10.2	-4.3	39.7	46.0
	5	0.20	-3.1	-13.1	26.7	30.0	-4.1	-13.4	32.0	40.5
	6	0.20	-1.9	-12.7	25.9	28.7	0.8	-9.8	34.3	38.0
	7	0.20	-1.8	-12.3	24.9	29.4	-1.1	-20.2	33.6	48.5
Low	1	0.20	4.0	-12.2	56.5	67.6	-4.6	-12.0	78.7	79.2
	2	0.20	-4.2	-23.9	62.9	77.9	-8.3	-21.7	76.3	83.7
	3	0.20	-11.0	-27.3	73.0	88.1	-7.4	-16.2	74.8	87.9
	4	0.20	6.2	-12.8	62.9	73.1	-9.5	-12.1	81.3	79.1
	5	0.20	1.2	-18.6	54.5	66.8	-2.9	-12.8	62.5	66.8
	6	0.20	-2.8	-23.9	55.9	69.0	-5.3	-20.7	67.2	75.6
	7	0.20	-3.5	-25.7	48.7	63.8	-13.3	-26.8	67.9	73.8

Table 14: Comparison of the relative bias and coefficient of variation for the estimates of average fishing mortality rates at age 1 in season 1 from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.084	-1.3	-2.5	10.1	10.1	-7.3	-12.3	14.8	15.9
	2	0.024	-5.4	-6.2	37.8	37.3	-9.1	-24.9	47.0	46.7
	3	0.084	-1.1	-1.8	8.8	8.9	-2.7	-4.5	12.5	12.5
	4	0.024	2.9	2.1	10.2	10.2	-1.2	-2.6	13.0	13.2
	5	0.167	-1.4	-1.8	9.0	9.2	-1.0	-1.8	9.8	10.9
	6	0.046	0.3	-0.1	10.5	10.4	0.9	-0.1	11.5	12.0
	7	0.085	0.9	1.5	8.6	8.5	2.0	0.0	17.6	20.5
Low	1	0.084	-9.3	-9.6	27.3	27.9	-12.0	-19.3	28.9	30.8
	2	0.024	-0.9	-0.7	35.9	36.0	-5.2	-19.8	46.6	44.2
	3	0.084	-15.8	-16.6	29.5	29.5	-12.1	-15.9	35.4	35.9
	4	0.024	0.5	-0.4	35.8	35.8	-2.5	-5.6	34.8	34.6
	5	0.167	-7.7	-7.2	29.9	30.5	-5.6	-8.4	30.8	32.1
	6	0.046	-1.4	-0.9	33.2	32.8	-1.1	-4.6	32.2	32.4
	7	0.085	-2.7	-1.1	31.1	31.6	2.0	-8.0	36.8	39.4

Table 15: Comparison of the relative bias and coefficient of variation for the estimates of average fishing mortality rate at age 2 in season 2 from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True Value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.075	-0.3	-7.2	6.4	6.6	-12.1	-20.8	21.8	27.9
	2	0.075	-3.3	-9.6	17.7	17.6	-12.1	-31.3	38.0	44.7
	3	0.075	-0.4	-6.4	6.3	6.4	-0.5	-8.9	21.6	25.0
	4	0.075	0.2	-5.6	6.6	6.5	-1.8	-9.3	21.8	25.1
	5	0.150	0.1	-4.8	5.4	5.5	0.5	-3.8	19.9	26.6
	6	0.150	0.1	-4.8	5.7	5.7	0.2	-4.6	21.0	26.7
	7	0.150	0.2	-3.4	6.5	6.7	7.1	5.1	26.6	37.0
Low	1	0.075	-1.0	-7.4	16.8	17.4	-11.5	-24.3	27.9	33.5
	2	0.075	-4.4	-10.6	17.7	18.1	-13.0	-29.8	37.0	36.2
	3	0.075	-3.4	-9.1	16.1	16.6	-5.1	-16.1	27.0	28.3
	4	0.075	-3.7	-9.8	18.8	19.0	-3.3	-12.0	33.6	33.7
	5	0.150	-1.1	-5.8	14.0	14.3	-4.4	-14.7	28.1	30.3
	6	0.150	-2.5	-7.6	13.1	13.4	-7.1	-17.6	27.1	27.8
	7	0.150	-1.1	-4.6	17.9	17.9	17.9	0.3	35.1	40.8

Table 16: Comparison of the relative bias and coefficient of variation for the estimates of average fishing mortality rate at age 3 in season 2 from applying the site-fidelity (S) and Markov (M) spatial estimation models to the 7 site-fidelity scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	True Value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.100	-2.3	-9.2	10.2	10.8	-14.2	-24.3	26.8	33.5
	2	0.100	-4.4	-11.7	20.4	21.5	-11.1	-33.4	44.5	47.3
	3	0.100	-0.2	-7.2	11.5	11.8	0.9	-10.9	28.4	33.3
	4	0.100	-0.1	-6.7	9.4	9.7	1.3	-9.7	27.2	32.1
	5	0.200	-0.1	-6.1	7.3	7.4	0.9	-5.0	25.8	35.2
	6	0.200	-0.9	-7.0	7.9	7.8	0.7	-6.5	25.6	34.1
	7	0.200	0.4	-6.7	7.9	8.3	10.6	1.1	33.3	49.2
Low	1	0.100	2.1	-5.9	22.3	22.4	-9.2	-22.9	34.3	40.8
	2	0.100	-4.2	-12.2	19.1	19.0	-12.3	-32.5	43.2	40.9
	3	0.100	-4.4	-11.3	21.2	21.3	-3.9	-16.5	31.8	35.3
	4	0.100	3.7	-4.7	22.3	21.4	-0.4	-9.5	42.7	43.4
	5	0.200	-0.8	-7.7	17.7	17.8	-4.6	-17.8	34.7	37.7
	6	0.200	-0.2	-7.6	17.6	17.9	-2.7	-18.2	37.4	39.5
	7	0.200	-2.5	-10.7	21.3	21.0	21.6	-11.4	36.1	44.4

Table 17: Comparison of the relative bias and coefficient of variation for the estimates of natural mortality rate at age 1 for the site-fidelity (S) and Markov (M) spatial estimation model for the 7 population dynamics scenarios defined in Table 3, and for the high and low data quality scenarios defined in Table 4, but with the twice the number of tag releases than in the basic simulations (i.e., same results as Table 12 except using 4800 total tag releases instead of 2400).

Data Quality	Pop. Dynamic Scenario	True value	Tagging all areas				Tagging only area 1			
			Bias		CV		Bias		CV	
			S	M	S	M	S	M	S	M
High	1	0.30	-5.6	-33.4	21.9	32.1	-24.4	-39.8	48.1	64.3
	2	0.30	-6.7	-33.7	37.5	53.6	-8.0	-13.8	40.4	42.6
	3	0.30	0.9	-23.1	21.0	28.3	-11.0	-34.8	73.9	112.1
	4	0.30	-5.3	-29.2	20.1	27.2	-6.5	-32.8	69.4	108.0
	5	0.30	1.2	-18.9	14.7	18.5	-13.3	-24.2	60.9	87.5
	6	0.30	1.0	-19.0	14.8	18.7	-15.4	-25.2	63.2	85.9
	7	0.30	0.1	-23.2	15.6	20.7	3.2	-4.3	27.9	42.9
Low	1	0.30	-1.1	-27.2	35.9	48.0	-31.3	-46.2	68.8	82.0
	2	0.30	-10.5	-36.1	42.1	56.2	-12.9	-18.7	40.6	44.5
	3	0.30	-7.3	-29.4	42.9	54.1	-17.7	-43.9	81.2	102.7
	4	0.30	-8.7	-30.8	42.8	55.4	-20.9	-45.5	85.1	98.0
	5	0.30	-7.6	-27.2	35.3	46.3	-26.0	-47.3	82.5	110.2
	6	0.30	-6.3	-24.0	34.8	42.5	-19.0	-45.2	72.9	111.3
	7	0.30	-4.8	-26.0	36.4	47.7	2.3	-20.2	41.6	54.6

Table 18: Percent of simulations within each scenario for which the value of the AIC statistic for the site-fidelity spatial estimation model exceeded that for the Markov estimation model for the 7 population dynamics scenarios defined in Table 3 with tagging only in area 1, and for the high and low data quality scenarios defined in Table 4.

Data Quality	Pop. Dynamic Scenario	Number of Releases	
		2400	4800
High	1	41	52
	2	42	48
	3	57	63
	4	47	62
	5	62	77
	6	64	74
	7	81	98
Low	1	43	44
	2	47	36
	3	32	48
	4	38	47
	5	41	61
	6	45	59
	7	48	58
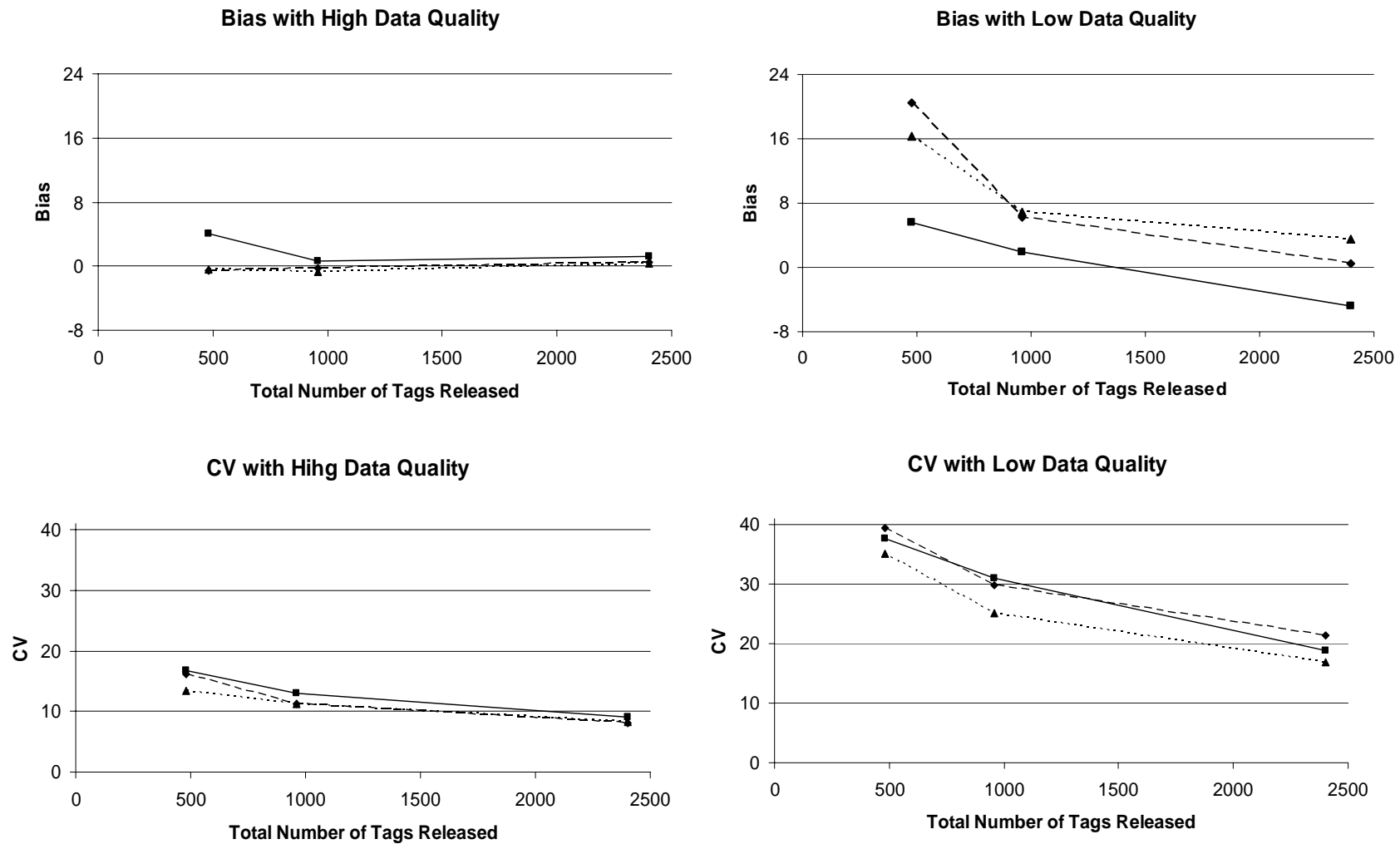
Figure 1: Examples of the relationship between number of tags released and the bias and CV for the estimates of total initial population size for the high and low quality data scenarios. Shown are results for population dynamics scenarios 7 (solid line), 8 (dotted line) and 9 (dashed line) defined in Table 2 (for the Markov movement model).
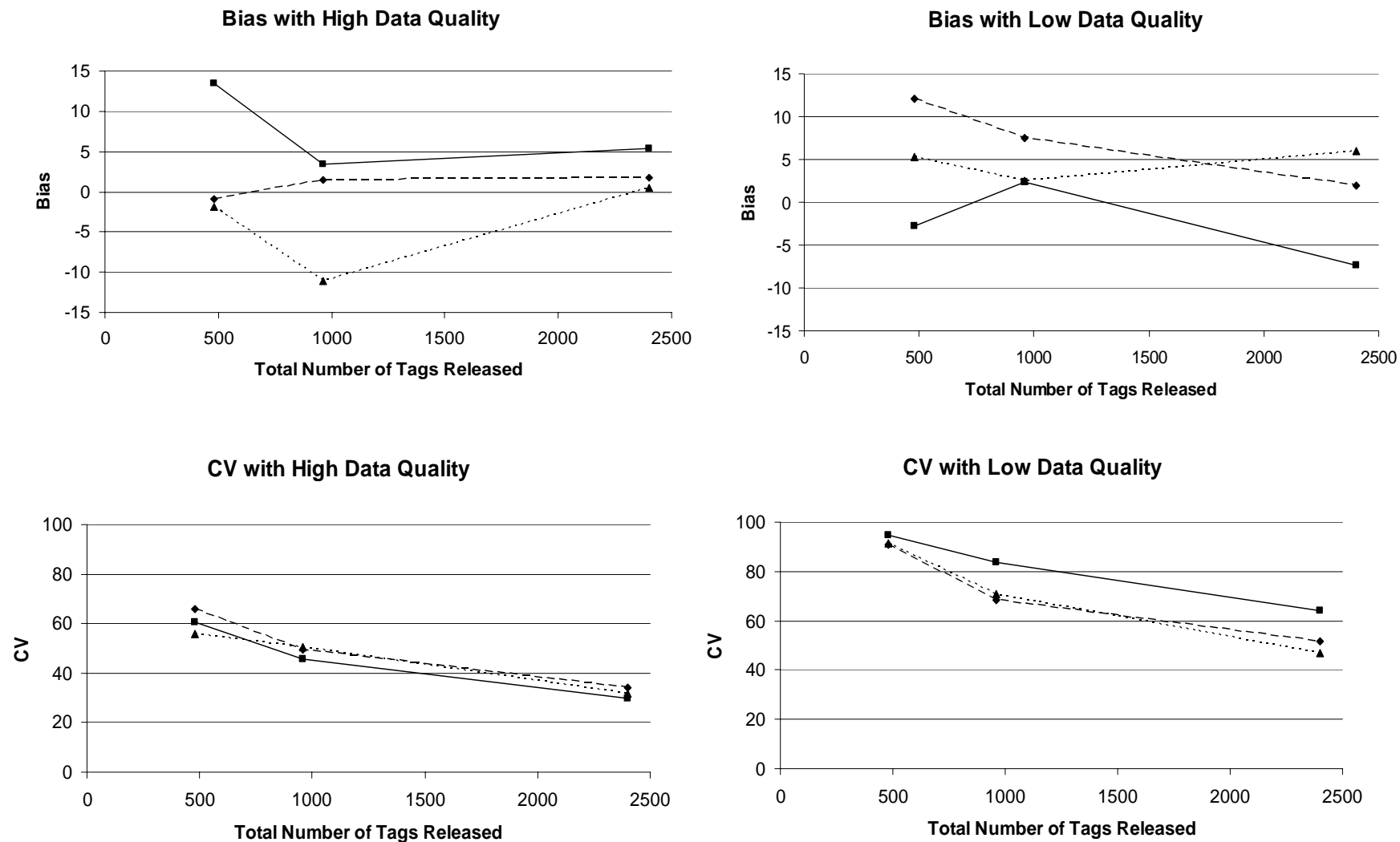
Figure 2: Examples of the relationship between number of tags released and the bias and CV for the estimates of natural mortality at age 1 for the high and low quality data scenarios. Shown are results for population dynamics scenarios 7 (solid line), 8 (dotted line) and 9 (dashed line) from Table 2 (for the Markov movement model).
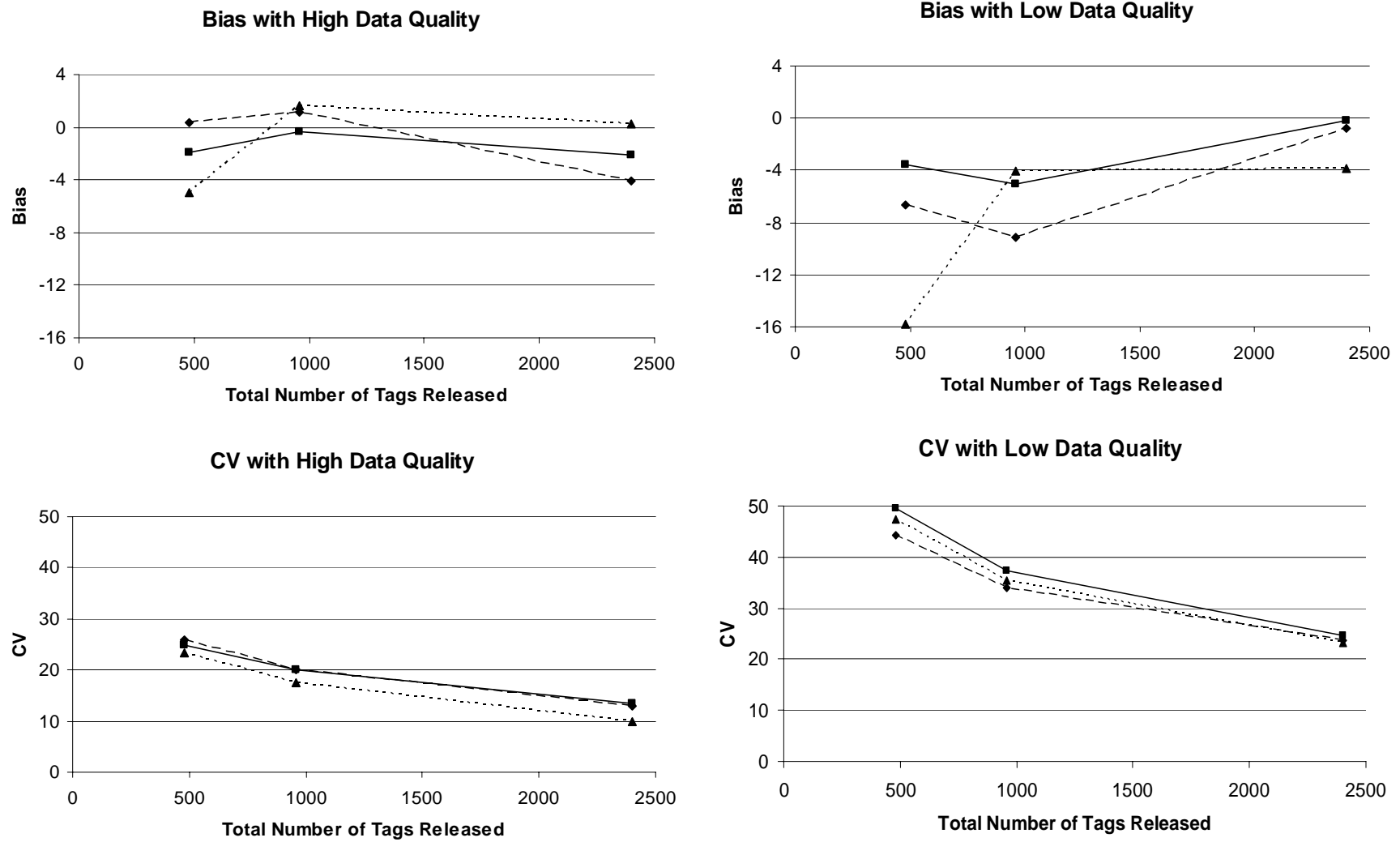
Figure 3: Examples of the relationship between number of tags released and the bias and CV for the estimates of average fishing mortality rate at age 3 in season 2 for the high and low quality data scenarios. Shown are results for population dynamic scenarios 7 (solid line), 8 (dotted line) and 9 (dashed line) from Table 2 (for the Markov movement model).
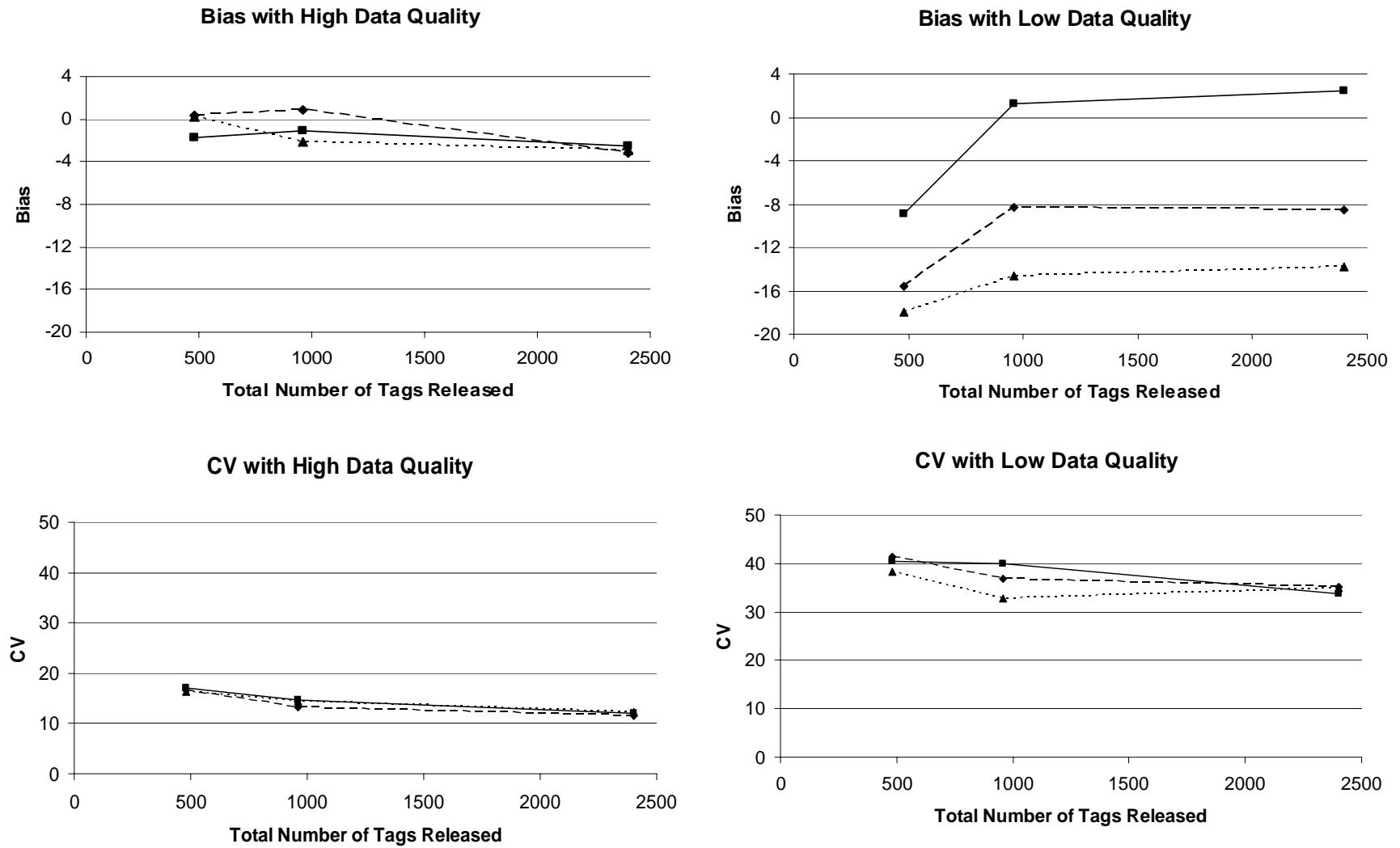
Figure 4: Examples of the relationship between number of tags released and the bias and CV for the estimates of average fishing mortality rate at age 1 in season 1 for the high and low quality data scenarios. Shown are results for population dynamic scenarios 7 (solid line), 8 (dotted line) and 9 (dashed line) from Table 2 (for the Markov movement model).

Appendix 13:

Parameter identifiability in spatial Brownie tag-recapture models

Geoff M. Laslett, J. Paige Eveson and Tom Polacheck

FRDC Project 2002/015

1 Introduction

In this document we examine parameter identifiability in spatial Brownie tagging models that are similar to the tagging component of the model presented in Appendix 10. The general model we consider has the spatial distribution of fish split into distinct regions with unconstrained movement of fish between regions at the end of each time period; in our examples, a time period is equal to a year. Fish from a single cohort are tagged in consecutive time periods (i.e., at consecutive ages when a time period is a year). The parameters to be estimated are natural mortality, which we either assume is constant or allow to vary with year, and fishing mortality, which we allow to vary with year and region. Note that for a single cohort of fish, year and age are redundant so that allowing fishing mortality and natural mortality to vary with year is equivalent to letting them vary with age. We first investigate which parameters are identifiable in models with two regions – we look at the effect of varying the number of recapture years, of varying the number of release years, and of tagging in all regions versus tagging in only one region. We also consider a two-region model in which an additional cohort of fish is tagged. Finally, we consider parameter identifiability in models with R regions.

2 Identifiability in the two-region generic model

2.1 Two release and recapture years

Consider tag-recapture data where a single cohort of fish is tagged in two regions at ages 1 and 2. The natural mortality is a constant M, and the fishing mortality in time period t in region r is F_{tr}. For a fish alive at the beginning of time period t in region r, the conditional probability of survival through time period t is $S_{tr} = \exp(-M - F_{tr})$. A non-surviving fish is caught (rather than dies naturally) with conditional probability $f_{tr} = F_{tr}/(M + F_{tr})$. Movement between regions occurs at the end of each time period. The probability of a fish staying in region 1 at the end of period 1 is π_1, and of staying in region 2 is π_2.

There are four groups of tagged fish: those of age 1 tagged in time period 1 in region 1; age 1 fish in time period 1 in region 2; age 2 fish tagged in time period 2 in region 1; and age 2 fish tagged in time period 2 in time period 2. When a tagged fish is caught, we can identify, from the tag number, which of the four tagging groups it belongs to. Assuming 100% reporting rates, we can calculate the probability of tag returns from each time period and region for fish in each group. This yields a 4×4 matrix P, in which $P[i, j]$ is the probability of a tag return from tagged group i and return group j. Here $i = 1$ corresponds to age 1, region 1; $i = 2$ to age 1, region 2; $i = 3$ to age 2, region 1; and $i = 4$ to age 2, region 2. So for fish tagged at age a in region r, the group index is $i = (a - 1)2 + r$. Similarly for the return groups, $j = 1$ corresponds to time 1, region 1; $j = 2$ to time 1, region 2; $j = 3$ to time 2, region 1; and $j = 4$ to time 2, region 2. In general, for a fish caught in time period t in region r, the return group index is $j = (t - 1)2 + r$.

The matrix of tag return probabilities is presented in Table 1.

Table 1: Probability of tag return from each region and time period
for each tagging group.

Tag group	Return group			
	$t = 1, r = 1$	$t = 1, r = 2$	$t = 2, r = 1$	$t = 2, r = 2$
1 ($a = 1, t = 1, r = 1$)	$(1 - S_{11})f_{11}$	0	$S_{11}\pi_1(1 - S_{21})f_{21}$	$S_{11}(1 - \pi_1)(1 - S_{22})f_{22}$
2 ($a = 1, t = 1, r = 2$)	0	$(1 - S_{12})f_{12}$	$S_{12}(1 - \pi_2)(1 - S_{21})f_{21}$	$S_{12}\pi_2(1 - S_{22})f_{22}$
3 ($a = 2, t = 2, r = 1$)	0	0	$(1 - S_{21})f_{21}$	0
4 ($a = 2, t = 2, r = 2$)	0	0	0	$(1 - S_{22})f_{22}$

There are seven unknown parameters: M, F_{11}, F_{12}, F_{21}, F_{22}, π_1 and π_2. The matrix P has 16 entries, of which 8 are structural zeroes. However, 8 are not structural zeroes, and so identifiability of 7 parameters from 8 equations is at least possible.

To prove identifiability, we adopt a staged approach. The notation '$\Rightarrow p$' means that p is identifiable, where p is a parameter or some combination of parameters.

1. From entries $P(3,3)$ and $P(4,4)$, $\Rightarrow (1 - S_{21})f_{21}$ and $(1 - S_{22})f_{22}$.

2. From entries $P(1,3)$ and $P(1,4)$, $(1 - S_{21})f_{21}$ and $(1 - S_{22})f_{22}$, $\Rightarrow S_{11}\pi_1$ and $S_{11}(1 - \pi_1)$.

3. Summing $S_{11}\pi_1$ and $S_{11}(1 - \pi_1)$, $\Rightarrow S_{11}$.

4. From S_{11} and $S_{11}\pi_1$, $\Rightarrow \pi_1$.

5. From S_{11} and entry (1,1), $\Rightarrow M$ and F_{11}.

6. From entries $P(2,3)$ and $P(2,4)$, $(1 - S_{21})f_{21}$ and $(1 - S_{22})f_{22}$, $\Rightarrow S_{12}(1 - \pi_2)$ and $S_{12}\pi_2$.

7. Summing $S_{12}\pi_2$ and $S_{12}(1 - \pi_2)$, $\Rightarrow S_{12}$.

8. From S_{12} and $S_{12}\pi_2$, $\Rightarrow \pi_2$.

9. From S_{12} and entry $P(2,2)$, $\Rightarrow M$ and F_{12}.

10. From M, $(1 - S_{21})f_{21}$ and $(1 - S_{22})f_{22}$, $\Rightarrow F_{21}$ and F_{22}.

Thus all parameters are identifiable.

We cannot generalise the model (without additional data) so that M depends on time: that is, $S_{tr} = \exp(-M_t - F_{tr})$, rather than $\exp(-M - F_{tr})$. Even though there are 8 equations for 8 unknowns, there are only 7 independent equations. Following the same steps as previously, the parameters M_1, F_{11}, F_{12}, π_1 and π_2 are identifiable. The quantities $(1 - S_{21})f_{21}$ and $(1 - S_{22})f_{22}$ are identifiable, but these two quantities are not enough to extract the three parameters M_2, F_{21} and F_{22}. Of course, if we know, for example, that $M_2 = M_1/2$, or more generally that $M_2 = f(M_1)$ for some known function $f(.)$, then all parameters are identifiable.

2.2 Tag returns from additional time periods

2.2.1 Tag returns from three time periods

Suppose tags are returned from three time periods. Then we have two additional columns in the return matrix.

Table 2: Probability of tag return from each region and time period
for each tagging group.

Tag group		Return group	
		$t = 3, r = 1$	$t = 3, r = 2$
1 ($a = 1, t = 1, r = 1$)	$\ldots$	$S_{11}\pi_{11\mid1}S_{21}\pi_{11\mid2}(1 - S_{31})f_{31}$	$S_{11}\pi_{11\mid1}S_{21}\pi_{12\mid2}(1 - S_{32})f_{32}$
		$+ S_{11}\pi_{12\mid1}S_{22}\pi_{21\mid2}(1 - S_{31})f_{31}$	$+ S_{11}\pi_{12\mid1}S_{22}\pi_{22\mid2}(1 - S_{32})f_{32}$
2 ($a = 1, t = 1, r = 2$)	$\ldots$	$S_{12}\pi_{21\mid1}S_{21}\pi_{11\mid2}(1 - S_{31})f_{31}$	$S_{12}\pi_{21\mid1}S_{21}\pi_{12\mid2}(1 - S_{32})f_{32}$
		$+ S_{12}\pi_{22\mid1}S_{22}\pi_{21\mid2}(1 - S_{31})f_{31}$	$+ S_{12}\pi_{22\mid1}S_{22}\pi_{22\mid2}(1 - S_{32})f_{32}$
3 ($a = 2, t = 2, r = 1$)	$\ldots$	$S_{21}\pi_{11\mid2}(1 - S_{31})f_{31}$	$S_{21}\pi_{12\mid2}(1 - S_{32})f_{32}$
4 ($a = 2, t = 2, r = 2$)	$\ldots$	$S_{22}\pi_{21\mid2}(1 - S_{31})f_{31}$	$S_{22}\pi_{22\mid2}(1 - S_{32})f_{32}$

Note that we have slightly changed the notation. The transition probabilities now have an addition to their subscripts ($\mid t$) to indicate after which time period the transition took place, and we also show the initial and final regions for each transition. Thus π_1 is replaced by $\pi_{11\mid1}$ and π_2 by $\pi_{22\mid2}$. In this formulation (tagging a single cohort), time and age are completely confounded, so that the transition probabilities could be interpreted as depending on age ($\pi_{rr'\mid a}$) or time period ($\pi_{rr'\mid t}$).

We can identify the 7 parameters $M, F_{11}, F_{12}, F_{21}, F_{22}, \pi_{11\mid1}$ and $\pi_{22\mid1}$ as before. We now have four additional parameters to identify: $F_{31}, F_{32}, \pi_{11\mid2}$ and $\pi_{22\mid2}$. The proof that these can be identified is not as immediate as before.

Since S_{21} and S_{22} are already known, we can eliminate them from $P(3, 5)$ and $P(4, 5)$. Hence $\pi_{11\mid2}(1 - S_{31})f_{31}$ and $\pi_{21\mid2}(1 - S_{31})f_{31}$ are both identifiable. Taking their ratio, we can identify $\alpha = \pi_{11\mid2}/\pi_{21\mid2}$. Applying similar reasoning to $P(3, 6)$ and $P(4, 6)$, we can identify $\beta = \pi_{12\mid2}/\pi_{22\mid2}$. It is immediate that $\pi_{12\mid2} = (1 - \beta)/(\alpha - \beta)$, and that $\pi_{11\mid2} = (1 - 1/\beta)/(1/\alpha - 1/\beta)$. Hence both are identifiable. It is immediate, from $P(3, 5)$ and $P(3, 6)$ or $P(4, 5)$ and $P(4, 6)$, that F_{31} and F_{32} are identifiable.

It is noteworthy that $P(1,5)$, $P(1,6)$, $P(2,5)$ and $P(2,6)$ do not feature in this argument. Also, note that having data from an additional return year still does not allow us to estimate a time-dependent M.

2.2.2 Tag returns from four time periods

Now suppose that we have four years of returns. We look only at cells $P(3,7)$, $P(3,8)$, $P(4,7)$ and $P(4,8)$, as given in Table 3, from the return probability matrix.

Table 3: Probability of tag return from each region and time period
for each tagging group.

Tag group		Return group	
		$t = 4, r = 1$	$t = 4, r = 2$
3 $(a = 2, t = 2, r = 1)$	$\ldots$	$S_{21}\pi_{11|2}S_{31}\pi_{11|3}(1 - S_{41})f_{41}$	$S_{21}\pi_{11|2}S_{31}\pi_{12|3}(1 - S_{42})f_{42}$
	$\ldots$	$+ S_{21}\pi_{12|2}S_{32}\pi_{21|3}(1 - S_{41})f_{41}$	$+ S_{21}\pi_{12|2}S_{32}\pi_{22|3}(1 - S_{42})f_{42}$
4 $(a = 2, t = 2, r = 2)$	$\ldots$	$S_{22}\pi_{21|2}S_{31}\pi_{11|3}(1 - S_{41})f_{41}$	$S_{22}\pi_{21|2}S_{31}\pi_{12|3}(1 - S_{42})f_{42}$
	$\ldots$	$+ S_{22}\pi_{22|2}S_{32}\pi_{21|3}(1 - S_{41})f_{41}$	$+ S_{22}\pi_{22|2}S_{32}\pi_{22|3}(1 - S_{42})f_{42}$

Remember that all parameters except $\pi_{11|3}$, $\pi_{12|3}$, F_{41} and F_{42} have already been identified. The ratio $P(3,7)/P(4,7)$ is a linear combination of $\pi_{11|3}$ and $\pi_{12|3}$ with known coefficients. The ratio $P(3,8)/P(4,8)$ is also linear in these parameters. These two equations can be solved for $\pi_{11|3}$ and $\pi_{12|3}$, so that these parameters are identifiable. The identifiability of F_{41} and F_{42} follows immediately from $P(3,7)$ and $P(3,8)$ respectively.

2.2.3 Tag returns from T time periods

The proof of identifiability for tag returns from a general T time periods follows the same pattern. The proof is to show directly that the parameters are identifiable for $T = 2$, then use induction for larger T. If we can assume that all parameters are identifiable for T, then for $T + 1$ we need only show that $\pi_{1r|T}$ and $F_{T+1,r}$ are identifiable for $r = 1, 2$. The ratios $P(3, 2(T + 1) - 1)/P(4, 2(T + 1) - 1)$ and $P(3, 2(T + 1))/P(4, 2(T + 1))$ are linear functions of $\pi_{11|T}$ and $\pi_{12|T}$ with known

coefficients, so we can identify these two parameters. The identifiability of $F_{T+1,1}$ and $F_{T+1,2}$ follows immediately.

We once again note that only a single M parameter can be estimated in an experiment with two release years, regardless of the number of recapture years. The situation with more release years is explored in the next section.

2.3   Releases from an additional year

Suppose we tag a single cohort in three consecutive years, up to age 3 instead of age 2, and assume we have three recapture years. Then the matrix of return probabilities are given in Table 4.

Table 4: Probability of tag return from each region and time period for each tagging group.

Tag group	Return group $t = 1, r = 1$	$t = 1, r = 2$
1 $(a = 1, t = 1, r = 1)$	$(1 - S_{11})f_{11}$	0
2 $(a = 1, t = 1, r = 2)$	0	$(1 - S_{12})f_{12}$
3 $(a = 2, t = 2, r = 1)$	0	0
4 $(a = 2, t = 2, r = 2)$	0	0
5 $(a = 3, t = 3, r = 1)$	0	0
6 $(a = 3, t = 3, r = 2)$	0	0
	$t = 2, r = 1$	$t = 2, r = 2$
1 $(a = 1, t = 1, r = 1)$	$S_{11}\pi_{11|1}(1 - S_{21})f_{21}$	$S_{11}\pi_{12|1}(1 - S_{22})f_{22}$
2 $(a = 1, t = 1, r = 2)$	$S_{12}\pi_{21|1}(1 - S_{21})f_{21}$	$S_{12}\pi_{22|1}(1 - S_{22})f_{22}$
3 $(a = 2, t = 2, r = 1)$	$(1 - S_{21})f_{21}$	0
4 $(a = 2, t = 2, r = 2)$	0	$(1 - S_{22})f_{22}$
5 $(a = 3, t = 3, r = 1)$	0	0
6 $(a = 3, t = 3, r = 2)$	0	0
	$t = 3, r = 1$	$t = 3, r = 2$
1 $(a = 1, t = 1, r = 1)$	$S_{11}\pi_{11|1}S_{21}\pi_{11|2}(1 - S_{31})f_{31}$ $+ \; S_{11}\pi_{12|1}S_{22}\pi_{21|2}(1 - S_{31})f_{31}$	$S_{11}\pi_{11|1}S_{21}\pi_{12|2}(1 - S_{32})f_{32}$ $+ \; S_{11}\pi_{12|1}S_{22}\pi_{22|2}(1 - S_{32})f_{32}$
2 $(a = 1, t = 1, r = 2)$	$S_{12}\pi_{21|1}S_{21}\pi_{11|2}(1 - S_{31})f_{31}$ $+ \; S_{12}\pi_{22|1}S_{22}\pi_{21|2}(1 - S_{31})f_{31}$	$S_{12}\pi_{21|1}S_{21}\pi_{12|2}(1 - S_{32})f_{32}$ $+ \; S_{12}\pi_{22|1}S_{22}\pi_{22|2}(1 - S_{32})f_{32}$
3 $(a = 2, t = 2, r = 1)$	$S_{21}\pi_{11|2}(1 - S_{31})f_{31}$	$S_{21}\pi_{12|2}(1 - S_{32})f_{32}$
4 $(a = 2, t = 2, r = 2)$	$S_{22}\pi_{21|2}(1 - S_{31})f_{31}$	$S_{22}\pi_{22|2}(1 - S_{32})f_{32}$
5 $(a = 3, t = 3, r = 1)$	$(1 - S_{31})f_{31}$	0
6 $(a = 3, t = 3, r = 2)$	0	$(1 - S_{32})f_{32}$

Here we assume that the natural mortality of age 1 fish is M_1 and of age 2 and older fish is M_2. Using the now familiar arguments, we can identify all the parameters. Note that we cannot estimate a separate M parameter at age 3. The reason follows an analogous argument to the case for two release years, namely that the quantities $(1 - S_{31})f_{31}$ and $(1 - S_{32})f_{32}$ are identifiable, but that these two quantities are not enough to extract the three parameters M_3, F_{31} and F_{32}. For a tagging experiment in which a single cohort is tagged in n consecutive release years, the fact that only $n - 1$ natural mortality rate parameters can be estimated is a well-known feature of non-spatial Brownie models; we see here that this feature extends to the spatial case as well.

2.4   Tagging only in region 1

For a tagging experiment involving a single cohort of fish with two release years, two recapture years and two regions, assume that fish are only tagged in region 1, but there are returns from both regions. The matrix of tag return probabilities is presented in Table 5.

Table 5: Probability of tag return from each region and time period
for each tagging group.

Tag group	$t=1, r=1$	$t=1, r=2$	Return group $t=2, r=1$	$t=2, r=2$
1 $(a=1, t=1, r=1)$	$(1-S_{11})f_{11}$	0	$S_{11}\pi_1(1-S_{21})f_{21}$	$S_{11}(1-\pi_1)(1-S_{22})f_{22}$
3 $(a=2, t=2, r=1)$	0	0	$(1-S_{21})f_{21}$	0

For consistency with Table 1, we label the second row with the number 3, and refer to entries from this row as $P(3, j)$ for return cell j. Clearly it is impossible to identify the seven unknown parameters, M, F_{11}, F_{12}, F_{21}, F_{22}, π_1 and π_2, because there are only 4 non-zero cells in Table 5. Two of the parameters, π_2 and F_{12}, are completely missing from the table, so there is no hope of identifying them. Even so, we have

only four equations for the remaining five parameters.

What can be done? If we regard F_{11} as known, it is immediately clear that the other four parameters can be identified directly. Similar results apply if we fix M, F_{21} or π_1. It is less clear if we fix F_{22}. However,

$$
\begin{aligned}
P(1,1) & = (1 - S_{11})f_{11} \\
& = (1 - S_{11})\left(1 + \frac{M}{\log(1 - S_{11})}\right),
\end{aligned}
$$

so that

$$
M = \log(1 - S_{11})\left(\frac{P(1,1)}{1 - S_{11}} - 1\right). \tag{1}
$$

Also $S_{11}\pi_1 = P(1,3)/P(3,3)$. Hence

$$
P(1,4) = \left(S_{11} - \frac{P(1,3)}{P(3,3)}\right)(1 - \exp(-M - F_{22}))\frac{F_{22}}{M + F_{22}}.
$$

Substituting in the expression for M in equation (1), we obtain a single equation for S_{11}, showing that this quantity is identifiable. It follows from (1) that M is identifiable, and hence all the other parameters are identifiable.

A moment's reflection suggests that when we have the same number of equations as parameters, with no redundant equations (as is likely when they are complex non-linear equations), the parameters will be identifiable.

Is there any other way out of the non-identifiability quandary? We could make the five parameters a function of a smaller number of latent parameters. Clearly M and π_1 have to be left as free parameters, so that F_{ij} are the only candidates. We could set $F_{11} = q_1 E_{11}$, $F_{21} = q_2 E_{21}$ and $F_{22} = q_2 E_{22}$, where q_j are unknown catchability coefficients, and E_{ij} are known measures of effort. Then there are four parameters to be estimated from four equations, so all parameters are identifiable. One of the advantages of this approach is that we can now estimate F_{12}, say, as $q_1 E_{12}$, where q_1 has been estimated from the available data on tag returns.

An alternative solution may be to treat the F parameters as random effects; however, we leave this as an idea for future investigation.

2.5 Tagging an additional cohort

Suppose that in addition to the tagging in section 2.1, we also tag age 1 fish in time period 2 (i.e., we tag two cohorts of fish in time period 2). Note that year and age can no longer be used interchangeably. We assume that natural mortality varies only with age, and that fishing mortality varies with time period and region but not age. Then we define $S_{atr} = \exp(-M_a - F_{tr})$ and $f_{atr} = F_{tr}/(M_a + F_{tr})$. The matrix of return probabilities is presented in Table 6.

Table 6: Probability of tag return from each region and time period for each tagging group.

Tag group	$t=1, r=1$	$t=1, r=2$	Return group $t=2, r=1$	$t=2, r=2$
1 $(a=1, t=1, r=1)$	$(1-S_{111})f_{111}$	0	$S_{111}\pi_1(1-S_{221})f_{221}$	$S_{111}(1-\pi_1)(1-S_{222})f_{222}$
2 $(a=1, t=1, r=2)$	0	$(1-S_{112})f_{112}$	$S_{112}(1-\pi_2)(1-S_{221})f_{221}$	$S_{112}\pi_2(1-S_{222})f_{222}$
3 $(a=2, t=2, r=1)$	0	0	$(1-S_{221})f_{221}$	0
4 $(a=2, t=2, r=2)$	0	0	0	$(1-S_{222})f_{222}$
5 $(a=1, t=2, r=1)$	0	0	$(1-S_{121})f_{121}$	0
6 $(a=1, t=2, r=2)$	0	0	0	$(1-S_{122})f_{122}$

By following parallel steps to 1 to 9 from section 2.1, we can identify M_1, F_{11}, F_{12}, π_1 and π_2. From $P(5,3)$ and $P(6,4)$ we can identify F_{21} and F_{22}. From $P(3,3)$ we can identify M_2. Thus tagging the age 1 fish in period 2 has allowed us to identify all 8 parameters. Note, however, that this argument breaks down if natural mortality varies with age *and* cohort, or if fishing mortality depends in a completely general way on age, time period and region.

Oftentimes, however, fishing mortality will vary with age as well as time period and region. One solution in such cases is to assume that M is constant. Then, $S_{atr} = \exp(-M - F_{atr})$ and $f_{atr} = F_{atr}/(M + F_{atr})$, with the return probabilities the same as those in Table 6. We can use similar arguments as above to show that all parameters are identifiable in this situation as well.

3 Identifiability in the R-region generic model

For a single cohort, we assume that we have tagging in all R regions up to age A for times t from 1 to A. Thus there are $R * A$ tagging groups. Consider the tagging group in which fish of age $a - 1$ (where $a > 1$) are tagged in region r'. The return probability from region r at time a is set out in the first line of Table 7.

Table 7: Probability of tag return from each region and time period
for each tagging group.

Tag group	Return group $t = a, r$
$(a-1, t = a-1, r')$	$S_{a-1,r'} \pi_{r'r|a-1}(1 - S_{ar})f_{ar}$
$(a, t = a, r)$	$(1 - S_{ar})f_{ar}$

Note that $t = a$ for a single cohort, so we can use F_{ar} and F_{tr} interchangeably. The return probability for fish tagged at age a in region r caught during that season is given in the second line of Table 7. Comparing these two probabilities, we can identify $S_{a-1,r'} \pi_{r'r|a-1}$, for any r. Summing across all regions r, we can identify $S_{a-1,r'}$ for all r', and hence all $\pi_{r'r|a-1}$ for all r' and r.

To summarise, we have identified all the transition probabilities $\pi_{r'r|a}$ for $a = 1, \ldots, A-1$ and all r' and r. We have also identified S_{ar} for all $a = 1, \ldots, A-1$ and all r. From line 2 of Table 7, we can identify $(1 - S_{ar})f_{ar}$. But $S_{ar} = \exp(-M_a - F_{1r})$ is already identified, so we can identify $f_{ar} = M_a/(M_a + F_{ar})$. From these we can identify the two components M_a and F_{tr} for all t from 1 to $A-1$ and all r. To identify F_{Ar} as well, we have to assume that M_A is a known function of $M_1, \ldots, M_{A-1}$.

It is important to note that this proof of identifiability implicitly assumes that the probabilities $(1 - S_{ar})f_{ar} > 0$ for all a and r. This translates to the assumption that $M_a \geq 0$ for all a and that $F_{ar} > 0$ for all a and r. Suppose that there is no fishing in region r at time $t = a$, so that $F_{ar} = 0$. We make the plausible assumption that $M_a > 0$, so that $0 < S_{ar} < 1$. We still have no direct information on $\pi_{r'r|a-1}$, so we cannot immediately identify it, and hence we cannot identify $S_{a-1,r'}$.

Appendix 14:

Estimating tag shedding parameters from 1990s southern bluefin tuna tagging data

William S. Hearn, Tom Polacheck and J. Paige Eveson

FRDC Project 2002/015

Abstract

Tags are often shed from tagged fish, so that when a fish is recaptured the fisher cannot know that it has been tagged and cannot report it, which biases the results from analyses of the data. Tag shedding rates are mainly assessed from double-tagging experiments using the fraction of recaptured tagged fish that have shed one tag. We review the literature on methods for estimating tag shedding parameters from such experiments. We outline two methods for using shedding parameter estimates subsequently in tagging models for estimating mortality rates: (1) each fish caught with one or more tags is multiplied by a weighting factor that allows for the number of fish that have shed both tags, and this scaled up data can be input as data into a mortality model, or (2) the estimated shedding parameters may be directly incorporated into the structure of a mortality model. The second method has theoretical and practical advantages over the first. In an alternative approach, we develop an integrated model that is an extension of the instantaneous mortality rates model (i.e., Brownie model) that simultaneously estimates mortality and shedding parameters.

Independent of mortality considerations, we analyse southern bluefin tuna tag-return data from double-tagging experiments conducted in the 1990s to estimate individual shedding rates associated with each tagger (17 in total), as well as common shedding rates from data pooled over all taggers. The shedding model allows for shedding immediately after tagging and also continuous shedding over time. The model fitted to the data from individual taggers gave a significantly better fit than the model fitted to data pooled over all taggers. Furthermore, the fraction of tagged fish estimated to have lost both tags ranged among taggers from about 1% to 14%, indicating that differences among taggers' need to be taken into account when incorporating shedding into mortality models. In order to reduce the complexity of having to incorporate shedding rates from 17 individual taggers into the mortality model, we used model selection to pool data from taggers into 6 groups with similar shedding rates that gave a more parsimonious fit than the individual tagger model.

Introduction

Tagging experiments depend on accurate estimates of the number of tagged fish that are recaptured in order to be able derive reliable estimates of mortality rates and/or abundance. However, tags may shed from fish due to inadequately attached tags, tag failure, allergic rejection, accidental snagging, active removal by the fish, predators biting off the tag, or the attachment site becoming infected. Fish that have shed all their tags prior to recapture will generally be indistinguishable from fish that were never tagged; as such, the actual number of tagged fish recaptured will be underestimated. The number of recaptured fish that have shed all their tags needs to be estimated and accounted for when analysing tag-return data or else the mortality rates estimates will be biased. Accounting for tag shedding should be considered an essential component of any complete analysis of tag-return data.

To estimate the tag shedding rate, Beverton and Holt (1957, pp. 202-208) first suggested the technique of analysing the proportion of recaptured fish that have shed one tag in a double-tagging experiment. Double-tagging experiments are the primary approach used for estimating shedding rates and extensive work has been done on the problem of modelling and estimating shedding rates from such experiments (Kirkwood and Walker 1984; Barrowman and Myers 1996; Cadigan and Brattey 2003). In the current appendix, several models and methods for estimating tag shedding rates from double-tagging experiments are briefly reviewed and the double-tagging data for southern bluefin tuna (SBT) from the tagging experiments conducted in the 1990s are analysed. The results provide an indication of the level of tag shedding that might be expected in future tagging experiments and that should be allowed for in the design of such experiments. The tag shedding estimates are also used in the analysis of the 1990s SBT tag-return data presented in Appendix 15. It should be emphasized that the main reason for estimating shedding rates is generally so that unbiased estimates of mortality rates and/or abundance can be obtained from analyses of tag-return data (or at least estimates with acceptably small biases).

Methods

The use of double tagging to estimate shedding rates requires assumptions additional to those for a normal tag-return experiment (Hearn et al. 1991), namely:

1. Tags attached to a fish are a random sample from the tags used in an experiment.
2. Natural mortality, migration, catchability, and reporting of tags by finders are independent of the number of tags attached to a fish.

A third assumption is also needed in most models used to estimate tag shedding rates:

3. The shedding of any tag occurs independently of the shedding of other tags, including whether or not the other tag on a fish is still attached.

Tag shedding can be considered to be comprised of two components: (1) immediate tag loss (e.g. from improperly attached tags) and (2) shedding that occurs over time.

Immediate Tag Loss

Gulland (1963) and Russell (1980) considered the case where two tags, of types A and B, were attached to each fish, which are assumed to have independent probabilities of shedding. We follow their derivations, but initially we assume that all shedding occurs immediately after release. Let the retention probabilities of type A and B tags be Q_A and Q_B, respectively. A type A tag could be a tag placed on a fish's right side and a type B tag could be an identical tag placed on its left, as is the case with SBT tagging experiments. Although the tags may be identical, the shedding rates of these tags may not be the same because of differing inclinations/access of the two sides of the fish to the tagger at the time of tagging. Then the probabilities of a tagged fish retaining both A and B tags, A tag only, B tag only, and neither tag are $Q_A Q_B$, $Q_A(1-Q_B)$, $(1-Q_A)Q_B$, and $(1-Q_A)(1-Q_B)$, respectively. If r^* previously tagged fish are recaptured, then the expected number of these fish in each tag category are $r^* Q_A Q_B$, $r^* Q_A(1-Q_B)$, $r^*(1-Q_A)Q_B$ and $r^*(1-Q_A)(1-Q_B)$. The corresponding numbers of fish observed in each category are r_{AB}, r_A, and r_B (the number of fish that have lost both tags is not observed). Equating the expected and observed values yields estimates of the proportions shed, viz,

$$\hat{Q}_A = \frac{r_{AB}}{\left(r_{AB} + r_B\right)} = \frac{1}{\left(1 + \frac{r_B}{r_{AB}}\right)}. \tag{1a}$$

This makes intuitive sense – if we consider only recaptures that have a B tag (i.e. those from either category r_{AB} or r_B), then $\hat{Q}_A$ is the proportion of fish that have retained their A tag. Similarly,

$$\hat{Q}_B = \frac{r_{AB}}{(r_{AB} + r_A)} = \frac{1}{\left(1 + \frac{r_A}{r_{AB}}\right)}, \tag{1b}$$

and if $Q_A = Q_B = Q$ then

$$\hat{Q} = \frac{r_{AB}}{(r_{AB} + 0.5(r_A + r_B))} = \frac{1}{\left(1 + \frac{r_A + r_b}{2r_{AB}}\right)}. \tag{2}$$

The estimated number of tagged fish that were recaptured, including those that have lost both tags, is derived from

$$E(r_{AB}) = E(r^*)Q_A Q_B,$$

i.e.,
$$\hat{r}^* = \frac{r_{AB}}{\hat{Q}_A \hat{Q}_B}. \tag{3}$$

Define the weighting factor, W, as the number that needs to be multiplied by the total number of tagged fish recaptured with tags (i.e. $r_{AB} + r_A + r_B$) to estimate the total number of recaptured tagged fish r^*, i.e.

$$\hat{r}^* = \hat{W}(r_{AB} + r_A + r_B). \tag{4}$$

From (4), the estimated number of recaptured tagged fish that lost both tags is $(r_{AB} + r_A + r_B)(\hat{W} - 1)$, and the estimated fraction of tagged fish that lost both tags is $1 - 1/\hat{W}$.

Equating (3) and (4) and simplifying gives

$$\hat{W} = \frac{1}{\hat{Q}_A \hat{Q}_B \left(1 + \frac{r_A}{r_{AB}} + \frac{r_B}{r_{AB}}\right)}. \tag{5}$$

Substituting from (1a) and (1b) and simplifying yields

$$\hat{W} = \frac{1}{\hat{Q}_A + \hat{Q}_B - \hat{Q}_A \hat{Q}_B},$$

and if $Q_A = Q_B = Q$, then

$$\hat{W} = \frac{1}{2\hat{Q} - \hat{Q}^2} = \frac{1}{\hat{Q}(2 - \hat{Q})}. \tag{6}$$

This weighting factor can be used in equation (4) to estimate r^*, the actual number of recaptured tagged fish including those that shed both tags prior to recapture.

Time-Dependent Tag Loss

We now consider the case where tags continue to shed over time; i.e., the retention probability (Q) is a function of time. Beverton and Holt (1957) investigated the situation where tags are assumed to have an immediate component of shedding and a long-term constant proportional rate of shedding. In this case, the proportion of tags retained as a function of time since release, τ, is

$$Q(\tau) = \xi e^{-\Omega\tau} \tag{7}$$

where parameter ξ is the fraction of tags immediately retained (i.e. proportion $1 - \xi$ are immediately shed) and parameter Ω is the continuous shedding rate. Other functional forms for the retention probability can be used (e.g., Kirkwood 1981); however, they will not be considered in this appendix.

Kirkwood and Walker (1984) developed a method of estimating shedding rates from a double-tagging experiment with exact times of recapture, which they applied to data from Australian salmon. Here we describe their method in a slightly more general way, allowing for an arbitrary time-dependent shedding rate and for the two tags to shed at different rates. Each tag on a fish is assumed to shed independently of the presence or absence of the other tag.

Given that a fish is still alive, let the probabilities of tag A and tag B being retained at time τ after release (time-at-liberty) be $Q_A(\tau)$ and $Q_B(\tau)$, respectively. These functions will incorporate parameters to be estimated from the double-tagging data. For example, $Q_A(\tau)$ and $Q_B(\tau)$ may have the form of equation (7), in which case there will be four shedding parameters to be estimated: ξ_A, Ω_A, ξ_B, and Ω_B. Consider the ith caught tagged fish which is recaptured after time-at-liberty τ_i. The probabilities that it has both tags, tag A only, tag B only, or no tags are $Q_A(\tau_i)Q_B(\tau_i)$, $Q_A(\tau_i)(1-Q_B(\tau_i))$, $Q_B(\tau_i)(1-Q_A(\tau_i))$, and $(1-Q_A(\tau_i))(1-Q_B(\tau_i))$, respectively. However, a fish shedding two tags cannot normally be identified, so the probabilities conditional on a fish retaining at least one tag are

$$p_i(\tau_i) = \begin{cases} \dfrac{Q_A(\tau_i)Q_B(\tau_i)}{Q_A(\tau_i)+Q_B(\tau_i)-Q_A(\tau_i)Q_B(\tau_i)} & \text{if both tags are still attached} \\[3mm] \dfrac{Q_A(\tau_i)(1-Q_B(\tau_i))}{Q_A(\tau_i)+Q_B(\tau_i)-Q_A(\tau_i)Q_B(\tau_i)} & \text{if tag A only is still attached} \\[3mm] \dfrac{Q_B(\tau_i)(1-Q_A(\tau_i))}{Q_A(\tau_i)+Q_B(\tau_i)-Q_A(\tau_i)Q_B(\tau_i)} & \text{if tag B only is still attached} \end{cases}$$

A maximum likelihood approach is used to estimate the shedding parameters. The combined likelihood for all returns (those with both tags, tag A only and tag B only) is

$$L = \prod_{i=1}^{r} p_i(\tau_i),$$

where r is the number of fish for which one or more tags were returned. Note that provided the release procedures are kept uniform, the likelihood is a function of time-at-liberties only, and is independent of release times, even though recapture rates between release times may change. The shedding parameters are estimated by maximizing the likelihood (L), and statistical inferences are made in the normal way. These matters are examined further in Kirkwood and Walker (1984), Xiao (1995) and Barrowman and Myers (1996).

The estimated shedding parameters can then be used to estimate the weighting factor for a tagged fish recaptured after time at liberty τ_i. In particular,

$$\hat{W}(\tau_i) = \frac{1}{\hat{Q}_A(\tau_i)+\hat{Q}_B(\tau_i)-\hat{Q}_A(\tau_i)\hat{Q}_B(\tau_i)},$$

where $\hat{Q}_A$ and $\hat{Q}_B$ are the estimated retention functions (using the estimated shedding parameters). Also, if we assume that $Q_A(\tau)=Q_B(\tau)=Q(\tau)$, then

$$\hat{W}(\tau_i) = \frac{1}{\hat{Q}(\tau_i)(2-\hat{Q}(\tau_i))}, \tag{8}$$

which is analogous to equation (6). So for each tagged fish recaptured after time at liberty τ_i, $\hat{W}(\tau_i)$ tagged fish are presumed to have been recaptured instead of one. This uses assumption (2), that tagged fish have the same probability of being recaptured and their tags returned regardless of how many tags they still have

attached. The sum of all the $\hat{W}(\tau_i)$'s will then give an estimate of the total number of tagged fish that were recaptured, including those that lost both tags (analogous to r^* in the previous section on immediate tag loss).

Incorporating Shedding Rates into the Mortality Models

Shedding rates are generally of interest because they are needed in order to get unbiased estimates of mortality rates from analyses of tag-return data. There are two basic approaches for incorporating shedding rate estimates into mortality models (Wetherall and Yong 1981; Wetherall 1982): (1) estimate the tag shedding parameters independently (using a method such as one of those described above) and then use these estimates in subsequent analyses that estimate mortality rates (i.e., two-stage modelling), or (2) incorporate tag shedding parameter estimation into an integrated model that simultaneously estimates mortality rates and shedding rates (i.e., integrated modelling).

Two-Stage Modelling Approach

We first examine the two-stage modelling approach. This approach depends on the assumption that shedding of tags from fish are independent of mortality rates, which means the latter parameters do not need to be considered when estimating shedding rates (note that this assumption is generally made in the integrated modelling approach as well). We describe two approaches for how shedding parameter estimates can be used in subsequent mortality rate analyses.

The first two-stage approach is to adjust (scale up) the observed recapture data for the number of recaptured fish that had lost both tags, and then input these adjusted recapture data into the model for estimating mortality rates. The adjusted recapture data are calculated using the estimated weighting factor(s), as described in the previous sections (e.g. equations (6) and (8)). We call this the two-stage weighting option. For example, it is used in the exact-time model of Leigh et al. (2005).

The second two-stage approach is to adjust the mortality model rather than the data; that is, to incorporate shedding parameters into the mortality model equations and then substitute the estimated shedding parameters into the model. We call this the two-stage mortality and shedding model. For example, consider an instantaneous mortality rates model for estimating mortality rates in which the instantaneous fishing

mortality rate, F, is assumed to differ with year, and the instantaneous natural mortality rate, M, and the shedding parameters are assumed to remain constant over years. For simplicity, we also assume that tag reporting rates are one (incorporating reporting rate estimation is discussed briefly in the Discussion). At the start of year i, suppose that N_i fish are double tagged. If the retention function, $Q(\tau)$, is the same for both tags and has the form of equation (7), then the proportion of fish with at least one tag after time τ, is

$$2Q(\tau) - Q^2(\tau) = 2\xi e^{-\Omega\tau} - \xi^2 e^{-2\Omega\tau}.$$

Then, the expected number of fish recaptured with at least one tag (i.e., that could be identified as being tagged fish) for release year i and recapture year j is

$$
\begin{cases}
N_i\left(2\xi u'_j - \xi^2 u''_j\right) & \text{if } i = j \\
N_i\left(2\xi u'_j \prod_{k=i}^{j-1} S'_k - \xi^2 u''_j \prod_{k=i}^{j-1} S''_k\right) & \text{if } i < j
\end{cases}
\tag{9}
$$

where

$$S'_j = e^{-F_j - M - \Omega}$$

$$S''_j = e^{-F_j - M - 2\Omega}$$

$$u'_j = \frac{F_j}{F_j + M + \Omega}\left(1 - S'_j\right)$$

and

$$u''_j = \frac{F_j}{F_j + M + 2\Omega}\left(1 - S''_j\right).$$

Note that S'_j and S''_j cannot strictly be interpreted in the usual way as survivorship functions, nor can u'_j and u''_j be strictly interpreted as exploitation functions. The estimates of the shedding parameters ξ and Ω can be inputted into these equations and the analysis can then proceed as usual (e.g., a multinomial likelihood can be set up for the recapture data and maximized to obtain mortality rate estimates).

Both two-stage approaches (adjusting the data or adjusting the model) as described so far use the shedding parameter estimates as if they contain no uncertainty, and therefore the uncertainty in the mortality rate estimates will be underestimated. In the two-stage mortality and shedding model, the variance and co-

variance estimates of the estimated shedding parameters can be carried into the model through an additional likelihood term (see Appendix 15), and then the standard approach of using the inverse Hessian can be used to provide realistic uncertainty estimates for all parameters. There is not an equivalent approach for the two-stage weighting option; however, various methods can be used to account for the uncertainty in the shedding parameter estimates, such as using bootstrapping or Bayesian methods (these methods can also be used with the two-stage mortality and shedding model).

Integrated Modelling Approach

Integrating the estimation of shedding parameters directly into the mortality rate model provides an alternative to the two-stage modelling approach. Beverton and Holt (1957) used the method of moments with a model that incorporated the estimation of both shedding and mortality rates (assumed to be constant over years) to analyse data from a double-tagging experiment with a single tagging event. Wetherall (1982) extended this approach to use maximum likelihood estimation. We now show how to incorporate the estimation of tag shedding rates into the instantaneous mortality rates model, which we call the integrated mortality and shedding model. The two-stage mortality and shedding rates model described in the previous section only considered the total number of observed recaptures (i.e., it did not distinguish recaptures with one or two tags); this model formulation does not provide enough information to estimate shedding parameters directly from the model. Thus, for the integrated model we need to bin the recapture numbers not only by year of release and year of recapture but also by whether one or two tags were retained. We again assume that fishing mortality is year-dependent, that natural mortality and shedding rates are constant over years, and that reporting rates are equal to one. The tag retention function is assumed to be the same for both tags and to have the form of equation (7). Thus, after time τ, the proportion of tagged fish with one tag is

$$2Q(\tau)(1-Q(\tau)) = 2\xi e^{-\Omega\tau} - 2\xi^2 e^{-2\Omega\tau}$$

and the proportion with two tags is

$$Q^2(\tau) = \xi^2 e^{-2\Omega\tau}.$$

Suppose that N_i fish are double tagged at the start of year i, then the expected number of fish recaptured in year j with one tag is

$$
\begin{cases}
2N_i\left(\xi u'_j - \xi^2 u''_j\right) & \text{if } i = j \\
2N_i\left(\xi u'_j \displaystyle\prod_{k=i}^{j-1} S'_k - \xi^2 u''_j \prod_{k=i}^{j-1} S''_k\right) & \text{if } i < j
\end{cases}
\tag{10}
$$

and the expected number of fish recaptured with two tags is

$$
\begin{cases}
N_i \xi^2 u''_j & \text{if } i = j \\
N_i \xi^2 u''_j \displaystyle\prod_{k-i}^{j-1} S''_k & \text{if } i < j
\end{cases}
\tag{11}
$$

where S'_j, S''_j, u'_j, and u''_j are defined the same as in equation (9) for the two-stage mortality and shedding model.

For clarity, Table 1 contains the expressions for the expected number of recaptures for an experiment with three release years and three recapture years. The observed recaptures are binned over years of release and recapture and also tag retention category (i.e., whether one or two tags were retained), so that they correspond with the expected recapture expressions given above. The recapture numbers corresponding to a given release year (i.e., those from all recapture years with one tag or two tags) are multinomial so that the likelihood for all recaptures from all release years is the product of multinomials, similar to any Brownie-type model except expanded to include more categories for tag retention type. The likelihood can be maximized to estimate the shedding and mortality parameters and their standard deviations, using, for example, the maximum likelihood SURVIV software (White 1983). Note that the estimates of the shedding parameters obtained from the integrated model are not expected be the same as those obtained from the independent shedding analysis because the integrated model uses recovery data binned over years whereas independent shedding models generally use exact times of recovery (e.g., the method of Kirkwood and Walker (1984)).

In our presentations of both the two-stage mortality and shedding model and the integrated mortality and shedding model, we assumed reporting rates were 100%.

If this is not the case and we have an estimate of the reporting rate, λ_j, in each recapture year j from a prior independent analysis, then reporting rates can be incorporated into the mortality model by multiplying all terms for the expected number of recoveries in year j by λ_j. As with the two-stage shedding and mortality model, the variance and co-variance estimates of the estimated reporting rates can be carried into the model through an additional likelihood term (see Appendices 9 and 15). Alternatively, reporting rate estimation can be integrated into the mortality model. The way in which this is done depends on the method used to gather additional information for estimating reporting rates. For some methods (e.g., planted tags), the data can be considered independent of the tag-return data being used to estimate mortality rates and a likelihood can be developed for the reporting rate data and simply multiplied to the likelihood for the tag-return data (see Appendix 12). For other methods (e.g., based observers), the data is not independent of the tag-return data and the likelihood for the tag-return data needs to be modified to incorporate the reporting rate data (see Appendix 7).

Variation in Shedding Rates Between Tagging Operators

Some tagging operators (taggers) may not be as efficient in attaching tags as others, which can lead to biases if data pooled over taggers are analysed (Hearn et al. 1991). For example, suppose tagger x has 81 of his tagged fish returned with two tags and 18 returned with one, whereas tagger y has 36 of his tagged fish returned with two tags and 48 returned with one. If there is no long-term shedding and all tags have the same probability of being shed, then from equation (2), we estimate $Q_x = 0.90$ and $Q_y = 0.60$ (i.e., 10% of tagger x's tags have been shed and 40% of tagger y's). Then, using equations (4) and (6), the estimated numbers of fish that have shed both tags are 1 (i.e., $(81+18)(1/(0.9(2-0.9))-1)$) and 16 for x and y respectively, which total 17 altogether. However, if the data are pooled (117 fish returned with two tags and 66 with one tag), then we estimate $Q = 0.78$ (i.e., 22% of tags being shed), and the estimated numbers of fish that have shed both tags are 5 and 4.3 for x and y respectively, which total 9.3 altogether. Hearn et al. (1991) concluded that such biases may be markedly reduced by having taggers as similar as possible (e.g. by implementing a standard protocol) and by improving the efficiency of all taggers. New taggers should be properly trained and their work monitored. Results from

taggers should be compared with each other, and especially with those from experienced taggers; however, to be able to perform statistical tests comparing taggers' data, each tagger needs to have tagged a goodly number of fish.

One way to analyse return data from fish tagged by multiple taggers is to split it into tagger-associated subsets and estimate the shedding parameters for each subset by the exact-time method of Kirkwood and Walker (1984). A subset could be an individual tagger or a group of taggers assumed or shown to be similar in terms of shedding rates. Once shedding parameters have been estimated for each tagger subset, they can be incorporated into mortality models using either of the two-stage methods previously discussed. For the two-stage weighting option, a weighting factor $W_k(\tau_i)$ can be calculated for each subset k using the estimated shedding parameters for that subset. Each recaptured fish is then multiplied by its tagger-associated weighting factor, $W_k(\tau_i)$, to allow for the number of fish that have lost both tags ($W_k(\tau_i) \geq 1.0$). These adjusted recapture numbers are then used as data in the mortality model. For the two-stage mortality and shedding model, the expected number of recaptures given in equation (9) need to be duplicated for each tagger subset (i.e., we need to replace N_i with N_{ki}, ξ with ξ_k, and Ω with Ω_k) and, correspondingly, the data need to be broken down by tagger subset. The estimates of the shedding parameters for each tagger subset can then be inputted into these equations before proceeding with the mortality rate analysis.

Alternatively, the estimation of shedding rates for different taggers can be directly incorporated into the mortality model, which we call the integrated mortality and tagger-specific shedding model. For example, suppose there were two taggers (or tagger subsets), 1 and 2, operating for three years. Then duplicate the cells in Table 1 for each tagger, but replacing N_i with N_{ki}, ξ with ξ_k, and Ω with Ω_k for taggers $k = 1$, 2. This can become very complex if there are many taggers and/or the number of recaptures from taggers is small. As such, it is beneficial to pool data from taggers with similar shedding rates to the greatest extent possible.

Analysis of Southern Bluefin Tuna 1990s Tag-Recapture Data

In this appendix, we use the 1990s SBT tagging data to estimate shedding rates only (i.e., we do not incorporate the estimates into mortality models). Our

purpose is to gain an understanding of shedding rates and determine the best way to incorporate the estimates into a mortality model.

We selected data from the double-tagged fish that were deemed to be suitable for estimating shedding rates. We excluded data from fish that were found dead, tags that were reported but not recovered, tags found on beaches, tagged with an archival or planted tag, or badly tagged. We also excluded data with recovery year uncertain or recovery month uncertain within a year. Where the recovery day was uncertain within a month, we excluded the datum if the fish was at liberty for less than 270 days. If tagged fish were captured and re-released (e.g. by CSIRO taggers) we only analyse data from the last capture (which sometimes may be the recapture for which the fish was re-released). Juvenile SBT are commonly caught and placed into farm cages for fattening up; these fish are not inspected for tags prior to being placed into the cages, so we use the data associated with the time that the fish are removed from the cages and inspected for tags. We assume that fish shed tags at the same rate before and after they are placed into cages. The purpose is to estimate the probability of a fish losing both tags before it is recaptured and found. One may plant double-tagged fish in cages to independently evaluate the shedding of tags from fish in cages, but such an experiment needs to be carefully designed and its cost effectiveness evaluated. Furthermore, tag shedding in a cage may be higher for freshly tagged fish than for tagged fish that had their tags attached for a considerable period of time because of the crowded conditions in the cage and contact with cage netting.

Data sets associated with a tagger were only analysed if there were 30 or more acceptable recaptures in the set; this was true for 16 taggers. The data associated with the remaining taggers were pooled into a set we called "tagger" Z. Table 2 gives the total numbers of recaptures and the numbers of those with one or two tags attached for each tagger, as well as the fraction of recaptures with only one tag for all data and for times at liberty more than 9 months. To each tagger's data we applied the method of Kirkwood and Walker (1984) to estimate shedding parameters for that tagger. We used equation (7) for the retention function and we assumed the function was the same for both tags. Seber and Felton (1981) indicate that this assumption is likely to make little difference to the estimation of the number of unseen fish that shed both tags; however, we later investigate this issue. We tried three variations on the retention function: first, we assumed that shedding is immediate and has no

continuous component, so that Ω from equation (7) is set to zero and ξ only is estimated; second, we assumed that only continuous shedding occurs, so that ξ is set equal to one and only Ω is estimated; third, we assumed that both parameters are free. In all cases, we applied the logical limitations that $\xi \leq 1$ and $\Omega \geq 0$ (i.e. shed tags cannot spontaneously re-attach themselves to fish). We refer to these as shedding models 1, 2 and 3, respectively. We then used Akaike's (1973) information criterion (AIC) to determine which of the three shedding models provided the best fit to each tagger's data.

The results from the model that best fitted the data for each tagger are presented in Table 3 (standard deviations were estimated for the parameter estimates using the inverse Hessian matrix). Shedding model 1 fitted best to data of one tagger, shedding model 2 fitted best to data of nine taggers and model 3 fitted best to data of seven taggers. We used the estimated shedding parameters for each tagger to estimate a weighting factor, $W(\tau_i)$, for each recapture associated with that tagger, where τ_i is the time at liberty for the recapture. We then calculated an average weighting factor, W_a, for each tagger as the average over all $W(\tau_i)$'s for that tagger. The estimates of W_a ranged considerably between taggers from 1.011 to 1.156 (Table 3). In other words, the fraction of tagged fish estimated to have lost both tags (calculated using the approximate formula $1 - 1/W_a$) ranged between taggers from 1.1% to 13.6%.

We also fit the three shedding models to the data pooled over all taggers and the parameter estimates (referred to as the pooled parameter estimates) from the best fitting model are given in the last row of Table 3. The AIC for the pooled model fit is much larger than the sum of the AIC values from the individual tagger model fits (Table 3), rejecting the hypothesis that shedding rates are the same among all taggers. To evaluate the consequences of pooling data from all taggers, we calculated the number of recaptured fish estimated to have lost both tags for each tagger using the pooled parameter estimates and using the tagger's individual parameter estimates, then took the difference, which we define as the bias. These biases are presented in Table 3 as Bias 1, and we see that the bias can be substantial for some taggers.

Although evidence suggests it is not reasonable to assume common shedding rates for all taggers, it may be possible to identify groups of taggers with similar shedding rates. We fitted models to data pooled over groups of taggers and used AIC to determine groups with statistically similar shedding rates. Details are given in

Annex A to this appendix. We identified 6 groups as follows: Group A ≡ (5, 46), group B ≡ (7, 8, 53, 76), group C ≡ (3), group D ≡ (1, 2, 4), E (52, 77, 71) and group F ≡ (47, 82, 99, Z). For each group, the estimated shedding parameters and their variance-covariances are listed in Table 4. The probability of a tagged fish losing both its tags as a function of its time at liberty is compared for the 6 tagger groups (Figure 1). The biases in the number of recaptured fish estimated to have lost both tags for each tagger using the group-specific shedding parameter estimates versus using individual parameter estimates were calculated (Bias 2 in Table 3); the biases are greatly reduced compared to when common shedding parameters over all taggers were used. In Appendix 15, we input the group-specific shedding parameter estimates and their estimated variances and covariances to a mortality model for SBT; i.e., the two-stage mortality and shedding model is applied.

Finally, we analysed the taggers' data to determine if estimates of the shedding differed between the two tags. For SBT we label the two tags on a fish as "primary" and "companion", because the primary tag may not be associated with a particular side of a fish as some of the taggers are left-handed. For each tag type the shedding function used was equation (7), where parameters ξ_1 and Ω_1 refer to the primary tags and ξ_2 and Ω_2 refer to the companion tags. The best fit model to each data set was determined by the least AIC. We found differences in shedding rates between primary and companion tags for 10 taggers. In no case did the four parameter model give the best fit. In six cases it was found that $\xi_1 = \xi_2$ gave the least AIC, and in four cases $\Omega_1 = \Omega_2$ gave the least AIC. The resultant shedding parameter estimates, negative log-likelihood and AIC are given in Table 5. Also listed in Table 5 is the bias in the estimated number of recaptured tagged SBT that have lost both tags using the shedding parameters from Table 3 instead of Table 5; the largest absolute bias for any tagger is 2.

Discussion

Taking account of shedding rate differences between primary and companion tags made little difference to the estimated numbers of SBT that lost both tags (Table 5). This is in agreement with the finding of Seber and Felton (1981). As such, we conclude that it is sufficient when analysing the SBT tagging data to only consider tag

shedding models that assume shedding rates are the same for primary and companion tags.

There are substantial and statistically significant differences in shedding rate estimates between SBT tagged by various taggers. This is not surprising given the range of values observed among taggers for the fractions of recaptured tunas with one tag – from 15.7% for tagger 46 to 51.1% for tagger 47 when all recaptures are considered, and similar when only recaptures with times at liberty more than 9 months are considered (Table 2). The average weighting factors ranged from 1.011 for tagger 46 to 1.156 for tagger 47 (Table 3), which strongly suggests that shedding rates of individual taggers need to be taken into account when estimating mortality rates for SBT from the tag-return data. Incorporating individual shedding parameters for the 17 taggers into the mortality model would get very complicated since the data would need broken down into a large number of cells and the number of parameters to be estimated would be large. To reduce this complexity, we pooled data from taggers with similar shedding rates into 6 groups using AIC (see Annex A). Any bias due to this pooling of data into groups is small (Bias 2, Table 3) compared with the numbers of fish involved.

The probability of both tags being shed from a fish increases with time (see Figure 1), and therefore the number of recaptured tagged fish expected to have lost both tags will increase with time. This means that potential biases from ignoring individual tagger differences in shedding rates will likely progressively increase as the times of liberty used in the analysis increase (e.g. parameter estimates that are based on only the first year of returns would be less affected since the actual number of fish that had shed both tags would be more similar among taggers than for long terms at liberty). If shedding rates are high, it may be valid to truncate the data analysed on the basis of time at liberty and age, depending on the specifics of the analysis (e.g. "the chop option" of Latour et al. 2001, page 737). Such truncation is usually easily implemented and the results can be compared with analyses of the full data set.

It should be emphasized that estimating shedding parameters is generally not an end in itself, but to facilitate the estimation of mortality rates. The ideal is to use the integrated mortality and shedding model, but this can be complex and impractical if there are many taggers with different shedding rates. The two-stage weighting option is probably the simplest but has theoretical problems; e.g., the weighted data may not yield a maximum likelihood solution for the mortality estimates and the

variance-covariance matrix cannot be used to derive standard errors. The two-stage mortality and tagger-specific shedding model, for which a tagger dimension is added into the mortality model but the shedding parameters are estimated independently, seems like a more parsimonious approach and is theoretically and practically satisfactory, especially if the data are pooled into tagger groups of like data. Additionally, shedding information collected from sources other than the tagging experiment designed for mortality estimation can be readily incorporated into the independent shedding model; whereas it is more complex, and perhaps not even not possible, to incorporate this extra shedding information into the integrated model.

Tag shedding can substantially increase the uncertainty associated with estimates of mortality and abundance derived from tagging programs, particularly if shedding is high and/or poorly estimated. As such, it is critical that tag shedding be taken into account in the design phase. This includes ensuring that appropriate tagging techniques are used which will minimize the level of tag shedding and that sufficient data are collected to allow for reasonable estimation of the shedding rate. In most situations we would recommend that all fish be double tagged to allow for robust modelling of tag shedding (particularly when there are large numbers of taggers). It is also important to account for tag shedding when considering the number of tags to be released and the trade-off between resources devoted to the number of releases and other components of the tagging experiment (e.g. recovery activity, observer effort, catch estimation, etc). If this is not done, the expectation about the precision of parameter estimates that can be achieved will be over-optimistic.

References

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principal. *In* Second International Symposium on Information Theory. *Edited by* B.N. Petrov and F, Csaki. Akademiai Kiado, Budapest, Hungary.

Barrowman, N.J., and Myers, R.A. 1996. Estimating tag-shedding rates for experiments with multiple tag types. Biometrics, 52: 1410-1416.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. 2, Mar. Fish. GB Minist. Agric. Fish. Food, 19, 533 pp. pages 202-8, 217-218, 457-458.

Cadigan, N.G., and Brattey, J. 2003. Semiparametric estimation of tag loss and reporting rates for tag-recovery experiments using exact time-at-liberty data. *Biometrics* **59**, 869-876.

Gulland, J.A. 1963. On the analysis of double-tagging experiments. Int. Comm. Northwest Atl. Fish. Spec. Publ., 4: 228-229.

Hearn, W.S., Leigh, G.M., and Beverton, R.J.H. 1991. An examination of a tag-shedding assumption, with application to southern bluefin tuna. ICES J. Mar. Sci., 48: 41-51.

Kirkwood, G.P. 1981. Generalized models for the estimation of tag shedding by southern bluefin tuna (*Thunnus maccoyii*), J. du Cons. 39:256-260.

Kirkwood, G.P., and Walker, M.H. 1984. A new method for estimating tag shedding rates, with application to data for Australian Salmon, *Arripes trutta esper* Whitley. Aust. J. Mar. Freshw. Res., 35, 601-606.

Latour, R.J., Pollock, K.H., Wenner, C.A., and Hoenig, J.M. 2001. Estimates of fishing and natural mortality for subadult Red Drum in South Carolina waters. N. Am. J. Fish. Manag. 21: 733-744.

Leigh, G.M., Hearn, W.S., and Pollock, K.H. 2005 (in press). Time-dependent instantaneous mortality rates from multiple tagging experiments with exact times of release and recovery, J. Envir. Ecol. Stats.

Russell, H.J. 1980. Analysis of double-tagging experiments: an update. Can. J. Fish. Aquat. Sci. 37: 114-116.

Seber, G.A.F., and Felton, R. 1981. Tag loss and the Petersen mark-recapture experiment. Biometrica, 68, 211-219.

Wetherall, J.A. 1982. Analysis of double-tagging experiments. Fisheries Bulletin 80:687-701.

Wetherall, J.A., and Yong, M.Y.Y. 1981. Planning double-tagging experiments. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-SWFC-13, 13 p.

White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data J. Wildl. Manag. 47: 716-728.

Xiao, Y. 1995. A general model for estimating tag-specific shedding rates and tag interactions from exact or pooled times at liberty for a double tagging experiment. Can. J. Fish. Aqua. Sci. 53: 1852-1861.

Figure 1. Comparison among the 6 tagger groups of the probability of a tagged fish having lost both its tags as a function of its time at liberty.
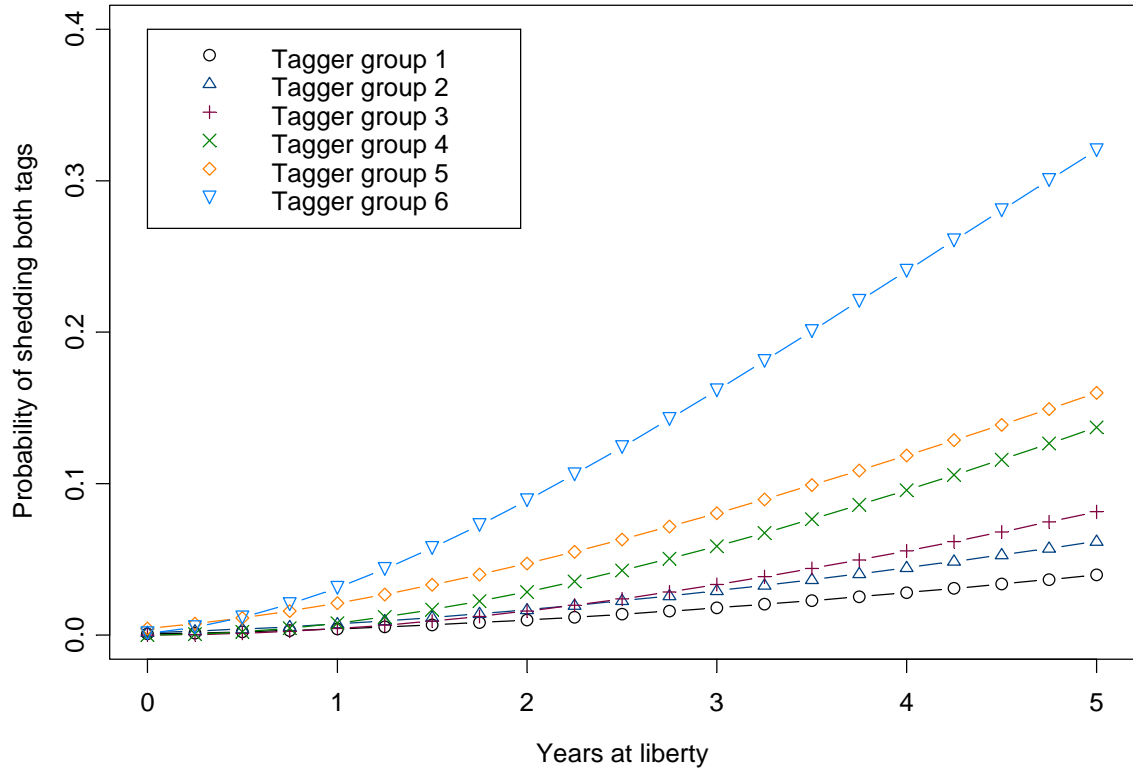
Table 1. Recapture probabilities for a multi-year double-tagging experiment in which N_i double-tagged fish are released at the start of each year i (i = 1,2,3). Probabilities of recapture are given for fish with one tag or two tags for years j (j=1,2,3), and are expressed in terms of year-dependent fishing mortality rates (F_j), constant natural mortality rate (M), the proportion of tags initially retained (ξ), and the continuous tag shedding rate (Ω).

Year of tagging	Number tagged	Number tags on recaptured fish	Year of recapture		
			1	2	3
1	N_1	1	$2\xi\left(u_1' - \xi u_1''\right)$	$2\xi\left(u_2' S_1' - \xi u_2'' S_1''\right)$	$2\xi\left(u_3' S_1' S_2' - \xi u_3'' S_1'' S_2''\right)$
		2	$\xi^2 u_1''$	$\xi^2 u_2'' S_1''$	$\xi^2 u_3'' S_1'' S_2''$
2	N_2	1	—	$2\xi\left(u_2' - \xi u_2''\right)$	$2\xi\left(u_3' S_2' - \xi u_3'' S_2''\right)$
		2	—	$\xi^2 u_2''$	$\xi^2 u_3'' S_2''$
3	N_3	1	—	—	$2\xi\left(u_3' - \xi u_3''\right)$
		2	—	—	$\xi^2 u_3''$

$$S_j' = e^{-F_j - M - \Omega}, \quad S_j'' = e^{-F_j - M - 2\Omega}, \quad u_j' = \frac{F_j}{F_j + M + \Omega}\left(1 - S_j'\right) \text{ and } u_j'' = \frac{F_j}{F_j + M + 2\Omega}\left(1 - S_j''\right)$$

Table 2. Summary of recaptures by tagger for the 1990s southern bluefin tuna double-
tagging data. For each tagger, the numbers of acceptable recaptured fish with primary tag
only attached, companion tag only attached, and the total (i.e. with primary, companion or
both tags attached) are given. The fraction of fish with one tag only attached are also given
for each tagger using: i) all recaptures, and ii) only recaptures after 9 months at liberty. Data
are pooled for taggers with less than 30 acceptable returns ("tagger" Z).

Tagger	Prim.	Comp.		Fraction with 1 tag	
Code	Only	Only	Total	All data	>9 mth
1	69	41	411	0.268	0.302
2	41	19	182	0.330	0.337
3	192	145	1368	0.246	0.257
4	55	78	479	0.278	0.309
5	208	210	2465	0.170	0.180
7	76	77	602	0.254	0.274
8	15	16	161	0.193	0.292
46	6	2	51	0.157	0.190
47	128	95	436	0.511	0.515
52	29	30	167	0.353	0.362
53	88	68	642	0.243	0.244
71	6	6	39	0.308	0.333
76	32	10	160	0.263	0.272
77	26	19	112	0.402	0.405
82	15	3	64	0.281	0.438
99	18	14	64	0.500	0.500
Z	25	14	113	0.345	0.405
Total	1029	847	7516	0.250	0.272

Table 3. Results from the best fitting model to data from each tagger. Shown are estimates of the tag shedding parameters (ξ and Ω), their estimated standard deviations (SD), negative log-likelihoods (*-LL*), AIC values, and average weighting factors (W_a). Also listed are the biases in the estimated number of fish that lost both tags from using shedding parameters estimated from (1) data pooled over all taggers and (2) data pooled into six tagger groups (see Table 4) instead of using tagger-specific shedding parameters. Data for all taggers with less than 30 acceptable returns have been pooled as tagger Z.

Tagger Code	Immediately retained fraction ξ	SD(ξ)	Continuous shedding rate Ω	SD(Ω)	*-LL*	AIC	W_a	Bias 1	2
1	0.9452	0.031	0.0589	0.019	308.1	620.3	1.033	-2	5
2	1.0000	-	0.0834	0.012	145.7	293.5	1.053	-2	2
3	1.0000	-	0.0672	0.004	951.3	1904.7	1.026	6	0
4	1.0000	-	0.0987	0.010	344.2	690.5	1.044	-9	-2
5	0.9727	0.007	0.0387	0.004	1364.7	2733.4	1.012	34	0
7	0.9760	0.019	0.0550	0.010	430.8	865.7	1.030	6	-1
8	0.9634	0.018	0.0512	0.016	94.4	192.8	1.021	1	0
46	1.0000	-	0.0435	0.016	24.7	51.3	1.011	1	0
47	0.9236	0.050	0.1436	0.026	447.1	898.2	1.156	-54	0
52	0.9310	0.049	0.0911	0.032	144.1	292.1	1.066	-6	-1
53	0.9532	0.025	0.0439	0.012	457.7	919.4	1.022	6	1
71	1.0000	-	0.1293	0.042	30.3	62.6	1.052	-1	0
76	0.8489	0.023	0.0000	-	121.2	244.4	1.023	0	-1
77	1.0000	-	0.1111	0.019	106.6	215.2	1.081	-5	0
82	1.0000	-	0.1876	0.050	44.4	90.8	1.061	-3	0
99	1.0000	-	0.1892	0.039	65.9	133.8	1.132	-7	-1
Z	1.0000	-	0.1654	0.030	92.7	187.3	1.096	-8	1
Total					5173.9	10395.8		-43	3
Pooled	0.9788	0.005	0.0657	0.003	5300.6	10605.2	1.029		

Table 4. Estimates of the fraction of tags immediately retained (ξ) and the continuous shedding rate (Ω), their variances and co-variances for data pooled into tagger groups. Group A $\equiv$ (5, 46), group B $\equiv$ (7, 8, 53, 76), group C $\equiv$ (3), group D $\equiv$ (1, 2, 4), E $\equiv$ (52, 77, 71) and group F $\equiv$ (47, 82, 99, Z).

Tagger group	No. tags returned	Immediately retained fraction ξ	Var(ξ) ($\times 10^{-4}$)	Continuous shedding rate Ω	Var(Ω) ($\times 10^{-4}$)	Cov(ξ, Ω) ($\times 10^{-4}$)
A	2516	0.9737	0.442	0.0391	0.188	0.224
B	1565	0.9608	1.403	0.0492	0.390	0.597
C	1368	1.0000	-	0.0672	0.157	-
D	1072	1.0000	-	0.0925	0.345	-
E	318	0.9342	16.050	0.0885	5.330	7.750
F	677	0.9666	4.950	0.1601	2.560	2.430
Pooled	7516	0.9788	0.220	0.0657	0.088	0.106

Table 5. Results when differences between primary and companion tags are accounted for; results only shown for taggers where the AIC is less than that in Table 3. The estimates of the fraction of tags immediately retained (ξ) and the continuous shedding rates (Ω), negative log-likelihoods (-*LL*), and AIC are listed. Also listed is the bias in the estimates of the number of fish that lose both tags when differences between primary and companion tags are not accounted for (i.e., as in Table 3).

Tagger code	Prim. tag only	Comp. tag only	Total tags	Immediately retained fraction ξ	Continuous shedding rates Ω_1	Ω_2	-*LL*	AIC	Bias
2	41	19	182	1.0000	0.0552	0.1095	141.49	286.99	1
3	192	145	1368	1.0000	0.0585	0.0757	947.97	1899.94	1
4	55	78	479	1.0000	0.1130	0.0837	342.32	688.65	1
46	6	2	51	1.0000	0.0222	0.0644	23.54	51.08	0
82	15	3	64	1.0000	0.0700	0.2865	40.00	83.99	2
xx	25	14	113	1.0000	0.1259	0.2010	91.12	186.23	0

Tagger code	Prim. tag only	Comp. tag only	Total tags	Immediately retained fractions ξ_1	ξ_2	Continuous shedding rate Ω	-*LL*	AIC	Bias
1	69	41	411	0.9765	0.9051	0.0556	304.38	614.75	1
47	128	95	436	0.9577	0.8675	0.1374	444.51	895.02	2
53	88	68	642	1.0000	0.9604	0.0561	457.04	918.08	-1
76	32	10	160	0.9219	0.7867	0.0000	115.16	234.32	1

Annex A

       To pool similar tagger data sets, we first ranked the tagger data sets 1 to 17 according to increasing continuous shedding rate (Ω) estimates from the model where no immediate shedding was assumed (i.e. $\xi = 1.0$). This ranking was used in preference to ranking according to immediate shedding $(1-\xi)$ because the immediate shedding rate did not differ greatly among taggers and was often estimated to be zero. When pooling data sets from taggers we assumed that taggers with data sets near each other in the ranking were more likely to be similar. This strategy was used to reduce the number of possible groupings, which is impractically large for 17 taggers. We started by comparing the lowest two ranked taggers, 46 and 5. The AIC of the model fitted to the pooled data from these two taggers was smaller than the sum of the AICs from the model fitted to the two individual tagger data sets, therefore pooling was accepted. The resultant pooled set was then compared with the next ranked set, 7, and pooling was rejected, as were the next four sets when we stopped the search. The pooled set of 46 and 5 is called group A. We then compared the next two ranked sets, 7 and 53, which were accepted. The next ranked set, 3, was rejected. Then in order, 8 was accepted, 2 was rejected, 76 was accepted, and four more were rejected. Thus the pooled set of 7, 8, 53 and 76 is called group B. We continued in this fashion to form four more groups C to F. This procedure was used to form initial groups. We then tried variations on these groups and found in a few cases that a tagger bettered belonged to a different group.

       We finally decided on the following six tagger groups: Group A $\equiv$ (5, 46), group B $\equiv$ (7, 8, 53, 76), group C $\equiv$ (3), group D $\equiv$ (1, 2, 4), E (52, 77, 71) and group F $\equiv$ (47, 82, 99, Z). For each group the shedding parameters estimates and their estimated variances and covariance are listed in Table 4 of the main appendix.

       To establish that this grouping is adequate, we present four arguments. Firstly, for a generic tagger group G, the corresponding negative log-likelihood ($-LL_G$) and AIC (AIC_G) are listed in Table A1. Also listed for each group is the sum of the AICs from Table 3 for all individual taggers belonging to the group (i.e., $\sum_{i \in G} AIC_i$). The pooled data is deemed to be a better fit if $AIC_G - \sum_{i \in G} AIC_i \leq 0.0$; i.e., if the pooling leads to an overall AIC reduction. This condition is met for all groups (Table A1). Note that group C is not comprised of pooled data as it consists of data from only one

tagger.

Secondly, using AIC we tested for significant differences in shedding rates between any two taggers within a group. In only one case (for taggers 47 and 82 of group F) was the AIC from the pooled model greater than sum of the two AICs from the individual tagger models; the difference in the AIC was only 0.148, and the statistical significance from a likelihood ratio test was estimated as $p = 0.143$ ($\chi^2 = 2.148$, df $= 1$).

Thirdly, we checked if AIC could identify any sub-groups within a group. In only one case was the AIC difference less than zero, (i.e. -0.228) for sub-groups (47, 99) and (82, Z) of group F, which is weak evidence for breaking F into two groups. In this case the statistical significance using a likelihood ratio test was estimated as $p = 0.136$ ($\chi^2 = 2.228$, df $= 1$). We decided to accept the grouping of sets in F, despite the slight infringement of the AIC. We investigated possible biases due to its acceptance and found it to be negligible.

Fourthly, we investigated all cases where AIC indicated that a tagger could possibly belong to another group. For example, we might accept adding tagger 7 to group C since the AIC from pooling group C and tagger 7 is smaller than the AIC for group C plus the AIC from tagger 7. However, in all such cases, the increase in AIC from removing the tagger from its current group was larger than the decrease in AIC from the new grouping. For example, adding tagger 7 to group C means removing him from group B, which results in two new groups that have a larger total AIC than the original two groups.

Table A1. Model fitting results from pooling data into six tagger groups. Shown are the negative log-likelihood (-LL_G) and AIC$_G$ for the model fitted to the pooled data for each group, the sum of the AICs from the models fitted to data from individual taggers within a group ($\sum_{i \in G}$ AIC$_i$ from Table 3), and the change in AIC due to pooling a group's data. Note that group C comprises data from only one tagger (3).

Tagger group	-LL_G	AIC$_G$	$\sum_{i \in G}$ AIC$_i$	AIC$_G - \sum_{i \in G}$ AIC$_i$
A	1390.050	2784.100	2784.728	-0.628
B	1105.180	2214.360	2222.244	-7.884
C	951.329	1904.658	1904.658	0.000
D	800.379	1602.758	1604.198	-1.440
E	280.894	565.788	569.962	-4.174
F	651.272	1306.544	1310.052	-3.508
Total	5179.104	10378.208	10395.842	-17.634

Appendix 15:

Estimation of mortality rates and abundance for southern bluefin tuna (*Thunnus maccoyii*) using tag-return and catch data from 1991 to 1997

J. Paige Eveson, Tom Polacheck and Geoff M. Laslett

FRDC Project 2002/015

Introduction

Data from conventional tagging experiments have become increasingly important in assessing the stock status of southern bluefin tuna (SBT) because they provide one of the few viable alternatives to catch-per-unit-effort data for estimating mortality rates and abundance. Extensive tagging programs have been carried out on juvenile SBT during various periods from the 1960s to present. Although some analyses of the tag-return data from these experiments have been conducted, especially for the 1990s data, a comprehensive analysis of the data taking into account all of the major potential sources of heterogeneity has not been completed because of the lack of a comprehensive modelling framework.

Polacheck et al. (1996, 1997, 1998) analysed the 1990s SBT tag-return data using Brownie models to provide estimates of fishing and natural mortality rates; however, in all of these analyses, tag shedding was assumed to be minimal enough that it could be ignored and reporting rates were assumed to be known without error. Recent analyses of the double-tagging data for SBT suggest that tag shedding can be substantial for some taggers (Appendix 14), so estimates of mortality rates and abundance obtained ignoring shedding are likely to be biased. Also, reporting rates are one of the more uncertain inputs in the tag-return models, so assuming they are known without error gives overly optimistic estimates of the variance of the parameter estimates.

The 1990s tag-return data have also been included in many of the integrated stock assessments for SBT (e.g., Kolody and Polacheck 2001; Polacheck et al. 2001). Like the analyses mentioned above, these stock assessments also assume that tag shedding is negligible and that reporting rates are known without error. Furthermore, the multi-year nature of the tagging data has not been fully exploited in the assessment models, as none of them incorporate a Brownie-type estimator for the tagging data. Instead, they tend to use attrition models that only allow for total mortality to be estimated and not the breakdown into fishing and natural mortality.

In the current appendix, we draw upon the non-spatial methods developed and presented throughout this report to construct a rigorous model for analysing the 1990s

SBT tag-return data. We use as a base model the combined Brownie and Petersen model described in Appendix 5, which integrates catch at age data with tag-return data in order to provide joint estimates of mortality rates (both fishing and natural) and abundance. More specifically, we use the modified version of the model described in the 'Application to southern bluefin tuna' section of Appendix 5 that allows for an initial period of non-mixing in the tag-return probabilities. This model assumes reporting rates are known without error, so we add an additional component to the model to take into account uncertainty in the reporting rate estimates. We also modify the tag-return probabilities as outlined in Appendix 14 to allow for instantaneous and continuous tag shedding. In a similar manner to the reporting rate estimates, we add an additional component to the model to take into account uncertainty in the estimates of the tag shedding parameters.   The model is applied to SBT data collected from 1991 to 1997 to provide estimates of fishing mortality rates, natural mortality rates, and initial abundance for a number of cohorts.

Materials and Methods

Southern bluefin tuna data

Four sets of data on southern bluefin tuna are used as input to the model: tag-return data from tagging experiments conducted from 1991 to 1997; estimates of tag reporting rates for years 1991 to 1997; estimates of tag shedding rates for six groups of taggers; and catch at age data from the commercial fisheries from 1991 to 1997. Further details about each data set are given below.

Tagging operations were carried out by CSIRO Marine Research from 1991 to 1997 in which juvenile SBT were caught, tagged, and released primarily in the coastal waters south of Western Australia and South Australia. Tagged fish were subsequently recaptured in the commercial fisheries and tags returned to CSIRO along with the date and location of recapture. A complete description of the tag-return data, including the tagging protocol, sampling procedures, method of age determination, and data screening processes, can be found in Appendix 4. Table 1a provides a summary of the 1990s tag-return data by cohort, age of release and year of recapture. Note that we have constrained our analysis to releases for ages 1 through 3. While

there were a small number of age 4 and 5 releases (<300 in total), these were not included in our analyses (and are not included in Table 1) because the number for any cohort was too small to provide meaningful estimates. The data are presented in terms of cohorts of fish to be consistent with the way that the model is developed and presented.

Estimating reporting rates for SBT is complicated because of limited data and because the complex nature of the SBT fishery, which comprises multiple components with varying reporting rates. Some limited observer data and tag seeding data exist, and these data have been used along with a number of alternative assumptions to provide year- and age-specific estimates of reporting rates for SBT from 1991 to 1997 (see Appendix 19 for details). A large number of alternatives have been provided but the reporting rate estimates used in our primary analysis (given in Table 1b) correspond to the reporting rates presented in Table 5a, option 8, of Appendix 19[1]. This is considered the 'most plausible' option because it is the most highly information-based (Anon. 2005). These reporting rate values differ from those presented in Appendix 5 because they correspond to a different option and because the reporting rate estimates were updated in 2005 after the analysis in Appendix 5 had been completed.

Essentially all SBT tagged in the 1990s were double-tagged. Data on the number of tagged fish that were recaptured with only one tag still attached versus both tags still attached were used to estimate shedding rates for SBT (Appendix 14). Tags were assumed to have an immediate component of shedding and a long-term constant proportional rate of shedding. Specifically, the proportion of tags retained as a function of time since release, τ, was assumed to be $Q(\tau) = \xi e^{-\Omega\tau}$, where ξ is the fraction of tags immediately retained (i.e. proportion 1 - ξ are immediately shed) and Ω is the continuous shedding rate. The shedding parameters were assumed to be

---

[1] The reporting rate estimates in Appendix 19 were prepared in 2005 for the CCSBT, for which it was decided that the discarded catch of small SBT recorded by the Japanese longline fishery in 1995 and 1996 should not be included in the catch at age data. For our current analysis, we prefer to include the estimated non-surviving portion of the discarded catches (Preece et al. 2001) in the catch data. As such, the reporting rate estimates presented in Table 1b of the current appendix differ very slightly in these two years from those presented in Table 5a of Appendix 19.

independent of age or year; however, they were found to differ significantly between taggers. Six groups of taggers with similar shedding rates were identified. Estimates of the instantaneous and continuous shedding parameters for the six tagger groups are given in Table 1c, along with standard errors and correlations for the estimates (these are all taken from Table 7a of Appendix 14). Note that because shedding rates were found to differ between groups of taggers, the probability of a tag being returned depends in part on the tagger group. As such, the model requires the release and recapture data in Table 1a to be broken down by tagger group as well as by cohort and release age; for brevity, we have not presented the data at this level of detail.

SBT are caught by a number of different fishing fleets and countries and the catch information available for each component differs considerably; thus, compiling total catch numbers by age is a complicated process. The catch at age data used in our analysis are taken from the catch at age data used in the 2004 stock assessments for SBT conducted by CSIRO. The only differences are:

1) Significant numbers of small SBT were caught and released by Japanese longline vessels in 1995 and 1996, and we have chosen to include the estimated non-surviving portion of the discarded catches in our catch data, whereas the data used in the assessments did not (Preece et al. 2001; Preece et al. 2004).

2) The catch at age data for the assessments were compiled by calendar year (starting January 1) whereas for our analysis we compiled the data by 'fishing' year, defined as starting November 1, to be more consistent with the major fishing seasons for SBT.

More information about the fishery components and the processing and compiling of the catch data can be found in Appendix 4. Table 1d summarizes the total SBT catch data by cohort and year for 1991 to 1997. These numbers differ slightly from those presented in Appendix 5 because the catch data were updated for the 2004 stock assessment, after the analysis in Appendix 5 had been completed.

The model

The model consists of four independent likelihood components, one for each of the tag-return data, the reporting rate estimates, the tag shedding estimates, and the catch data. Each of these likelihood components is described in detail below; however, before proceeding we introduce the notation that is used throughout the components.

A15-5

Data to be inputted into the model:

K = number of tagged cohorts

A_k = minimum age of tagging (and also minimum age of returns) for cohort k

B_k = maximum age of tagging for cohort k

I_k = maximum age of returns for cohort k

T = number of tagger groups

$N_{k,t,a}$ = number of age a fish from cohort k tagged and released by tagger group t

$R_{k,t,a,i}$ = number of tags returned from age i fish from cohort k that were tagged at age a by tagger group t

$C_{k,i}$ = number of age i fish caught from cohort k

υ_C = coefficient of variation of the catch data (common across ages and cohorts)

$\hat{\lambda}_{k,i}$ = estimated reporting rate for tagged fish caught at age i from cohort k

σ_λ = standard error of the estimated reporting rates (common across ages and cohorts)

$\hat{\xi}_t$ = estimated immediate tag shedding rate for tagger group t

σ_{ξ_t} = standard error of $\hat{\xi}_t$

$\hat{\Omega}_t$ = estimated continuous tag shedding rate for tagger group t

σ_{Ω_t} = standard error of $\hat{\Omega}_t$

ρ_t = correlation between $\hat{\xi}_t$ and $\hat{\Omega}_t$

Parameters to be estimated in the model:

M_i = instantaneous natural mortality rate for age i fish

$F_{k,i}$ = instantaneous fishing mortality rate for age i fish from cohort k (excluding fish tagged at age i)

$F^*_{k,t,i}$ = instantaneous fishing mortality rate for age i fish from cohort k tagged by tagger group t at age i (i.e. for newly tagged fish in their first year of tagging)

P_{k,A_k} = population size of cohort k at age A_k (the minimum age of tagging for cohort k)

$\lambda_{k,i}$ = reporting rate for tagged fish captured at age i from cohort k

ξ_t = immediate tag shedding rate for tagger group t

Ω_t = continuous tag shedding rate for tagger group t

Note that we allow fishing mortality (F) to differ between both ages and cohorts, whereas we only allow natural mortality (M) to differ between ages.

Underlying the tag-return and catch likelihoods are the general population dynamics equations commonly used in fisheries, expressed in terms of exponential and competing natural and fishing mortality rates. In particular, for a cohort of animals of a given age, the expected number of animals that survive to the next age and the expected number caught are expressed by

$$P_{k,i+1} = P_{k,i} \exp\left(-F_{k,i} - M_i\right) \tag{1}$$

and

$$C_{k,i} = \frac{F_{k,i}}{F_{k,i} + M_i} P_{k,i} \left(1 - \exp\left(-F_{k,i} - M_i\right)\right) \tag{2}$$

$P_{k,i}$ is the population size of cohort k at age i, and all other notation is as defined above.

First consider the tag-return component of the model. Analogous to the application to SBT in Appendix 5, we modify the likelihood for a standard Brownie model to allow for fishing mortality to differ between tagged fish in the year of tagging and untagged fish in that same year (following the model presented in Hoenig et al. 1998). This is to allow for the fact that newly tagged fish will not be fully mixed with the untagged population immediately after tagging, and for the fact that tagging generally occurs during the fishing season so tagged fish are only vulnerable for part of the season. We assume that tagged and untagged fish are fully mixed by the year following release (all tagging of SBT occurred between November and April so this allows several months for mixing to occur). In addition, we modify the tag-return likelihood to allow for group-specific estimates of tag shedding parameters. In particular, we

revise the return probabilities in the same manner as outlined in Appendix 14 (equation 9).

Taking into consideration tag shedding, tag reporting rates, and different return rates for newly tagged fish, the probability of a tag being returned from an age i fish from cohort k that was tagged at age a by tagger group t and has retained at least one tag is

$$
p_{k,t,a,i} = \begin{cases}
\left(2\xi_t u'^*_{k,t,i} - \xi_t^2 u''^*_{k,t,i}\right)\lambda_{k,i} & i = a \\[2mm]
\left(2\xi_t u'_{k,t,i} S'^*_{k,t,a} - \xi_t^2 u''_{k,t,i} S''^*_{k,t,a}\right)\lambda_{k,i} & i = a+1 \\[2mm]
\left(2\xi_t u'_{k,t,i} S'^*_{k,t,a} \prod_{m=a+1}^{i-1} S'_{k,t,m} - \xi_t^2 u''_{k,\tau,i} S''^*_{k,t,a} \prod_{m=a+1}^{i-1} S''_{k,t,m}\right)\lambda_{k,i} & i > a+1
\end{cases}
$$

where

$$
S'_{k,t,i} = \exp\left(-F_{k,i} - M_i - \Omega_t\right)
$$

$$
S''_{k,t,i} = \exp\left(-F_{k,i} - M_i - 2\Omega_t\right)
$$

$$
u'_{k,t,i} = \frac{F_{k,i}}{F_{k,i} + M_i + \Omega_t}\left(1 - S'_{k,t,i}\right)
$$

$$
u''_{k,t,i} = \frac{F_{k,i}}{F_{k,i} + M_i + 2\Omega_t}\left(1 - S''_{k,t,i}\right)
$$

$$
S'^*_{k,t,i} = \exp\left(-F^*_{k,t,i} - M_i - \Omega_t\right)
$$

$$
S''^*_{k,t,i} = \exp\left(-F^*_{k,t,i} - M_i - 2\Omega_t\right)
$$

$$
u'^*_{k,t,i} = \frac{F^*_{k,t,i}}{F^*_{k,t,i} + M_i + \Omega_t}\left(1 - S'^*_{k,t,i}\right)
$$

$$
u''^*_{k,t,i} = \frac{F^*_{k,t,i}}{F^*_{k,t,i} + M_i + 2\Omega_t}\left(1 - S''^*_{k,t,i}\right).
$$

Note that we allow fishing mortality for newly tagged fish (F^*) to differ not only between ages and cohorts, but also between tagging groups. This is necessary because different tagging groups will tag fish at different locations and different times during the season (and some tagging groups may not tag any fish in a particular year); thus, the probability of fish tagged by a particular tagging group being caught in the

same year it was tagged will depend in part on the tagging group. Note that these F^* parameters are nuisance parameters and are of little interest relative to the overall dynamics of the stock.

If tag returns are assumed to be independent, then the number of returns at age (including those not returned) corresponding to releases from a particular cohort at a particular age by a particular tagger group will be multinomial with probabilities as given above. Thus, the likelihood for all the returns at age data, over all cohorts, tagger groups and ages of release, is

$$L_R = \prod_{k=1}^{K} \prod_{t=1}^{T} \prod_{a=A_k}^{B_k} \left\{ \mathrm{K}_{k,t,a} \left(\prod_{i=a}^{I_k} \left(p_{k,t,a,i} \right)^{R_{k,t,a,i}} \right) \left(1 - \sum_{i=a}^{I_k} p_{k,t,a,i} \right)^{N_{k,t,a} - \sum_{i=a}^{I_k} R_{k,t,a,i}} \right\} \tag{3}$$

where

$$\mathrm{K}_{k,t,a} = \frac{N_{k,t,a}!}{\displaystyle\prod_{i=a}^{I_k} R_{k,t,a,i}! \left(N_{k,t,a} - \sum_{i=1}^{I_k} R_{k,t,a,i} \right)!} .$$

Note that $\mathrm{K}_{k,t,a}$ is a constant that can be left out when maximizing the likelihood.

Next consider the reporting rate estimates. The procedure used to produce these estimates did not provide associated standard error estimates, so we assume that the reporting rate estimates have a common and known standard error and explore the effect of varying this value. The reporting rate estimates ($\hat{\lambda}_{k,i}$'s) and their assumed standard error (σ_λ) are brought into the model as data through an independent likelihood term. We assume that $x_{k,i} = n_{k,i} \hat{\lambda}_{k,i}$ is the number of tags reported at age i from cohort k, and that $x_{k,i}$ is binomial with probability $\lambda_{k,i}$ and sample size $n_{k,i}$. Note $n_{k,i}$ can be thought of as the sample size required to achieve the level of precision in the reporting rate estimate $\hat{\lambda}_{k,i}$ specified by σ_λ. Using the variance formula for a binomial distribution, we know that

$$Var\left(\hat{\lambda}_{k,i} \right) = \frac{\lambda_{k,i} \left(1 - \lambda_{k,i} \right)}{n_{k,i}} .$$

We also know that

$$Var\left(\hat{\lambda}_{k,i}\right) = \sigma_\lambda^2,$$

so setting these equal we can solve for the sample size as

$$n_{k,i} = \frac{\lambda_{k,i}\left(1-\lambda_{k,i}\right)}{\sigma_\lambda^2} \approx \frac{\hat{\lambda}_{k,i}\left(1-\hat{\lambda}_{k,i}\right)}{\sigma_\lambda^2}.$$

Thus, the likelihood for the reporting rates can be specified as

$$L_\lambda = \prod_{k=1}^{K}\prod_{i=A_k}^{I_k} \frac{n_{k,i}!}{x_{k,i}!\left(n_{k,i}-x_{k,i}\right)!} \lambda_{k,i}^{x_{k,i}}\left(1-\lambda_{k,i}\right)^{n_{k,i}-x_{k,i}} \tag{4}$$

where $n_{k,i} = \dfrac{\hat{\lambda}_{k,i}\left(1-\hat{\lambda}_{k,i}\right)}{\sigma_\lambda^2}$ and $x_{k,i} = n_{k,i}\hat{\lambda}_{k,i}$.

For the tag shedding data, we take a similar approach as for the reporting rates and bring the group-specific tag shedding parameter estimates and their estimated standard errors and correlations into the model as data through an independent likelihood term. We assume that the two estimates for a given tagger group, $\hat{\xi}_t$ and $\hat{\Omega}_t$, have a bivariate normal distribution. Thus, the likelihood for the tag shedding data over all tagger groups is

$$L_{\xi,\Omega} = \prod_{t=1}^{T} \frac{1}{2\pi\sigma_{\xi_t}\sigma_{\Omega_t}\sqrt{1-\rho_t^2}} \exp\left(-\frac{1}{2\left(1-\rho_t^2\right)}\left\{\left(\xi_t'\right)^2 - 2\rho_t\xi_t'\Omega_t' + \left(\Omega_t'\right)^2\right\}\right) \tag{5}$$

where $\xi_t' = \dfrac{\hat{\xi}_t - \xi_t}{\sigma_{\xi_t}}$ and $\Omega_t' = \dfrac{\hat{\Omega}_t - \Omega_t}{\sigma_{\Omega_t}}$.

Following the arguments presented in Appendix 5, we have chosen to model the error in the catch data as Gaussian with a constant coefficient of variation across ages and years/cohort. The coefficient of variation, denoted by υ_C, is intended to encompass both process error, which results from fishing being a random process, and sampling error, which results from the age distribution of the catch being determined by taking a sample, estimating the ages of fish in the sample (either from lengths or from direct aging of hard parts), and using the estimated age frequencies of the sample to represent the total catch. We assume that υ_C is known because, as discussed in Appendix 5, it cannot be estimated reliably. Assuming the catch data are independent between cohorts and ages, the likelihood for the catch data is

$$L_C = \prod_{k=1}^{K} \prod_{i=A_k}^{I_k} \frac{1}{\sqrt{2\pi}\,\upsilon_C\,E\left(C_{k,i}\right)} \exp\left(-\frac{1}{2}\left(\frac{C_{k,i} - E\left(C_{k,i}\right)}{\upsilon_C\,E\left(C_{k,i}\right)}\right)^2\right) \tag{6}$$

where

$$E\left(C_{k,i}\right) = \begin{cases} P_{k,A_k}\,u_{k,A_k} & i = A_k \\ P_{k,A_k}\,u_{k,A_k} \prod_{m=A_k}^{i-1} S_{k,m} & i > A_k \end{cases}$$

$$S_{k,i} = \exp\left(-F_{k,i} - M_i\right)$$

$$u_{k,i} = \frac{F_{k,i}}{F_{k,i} + M_i}\left(1 - S_{k,i}\right).$$

The overall likelihood is given by the product of the four likelihood terms, namely

$$L_{Total} = L_R \times L_\lambda \times L_{\xi,\Omega} \times L_C \tag{7}$$

Multiplying the likelihood terms together is only a valid procedure if the likelihoods are independent of each other. The assumption of independence between the tag-return and the catch likelihoods may not be met if during the process of sampling the catch data, tags are found and returned; however, we assume that tags are removed at the time of catch, prior to catch sampling. Moreover, if the catch sample is relatively

small, then the expected number of tags in the sample will be so small that the independence assumption will not be seriously violated. Independence between the tag shedding and tag-return likelihoods should be a valid assumption – even though the tag shedding estimates were obtained using the tagging data, the shedding estimates only use information on the number of recaptures with one tag versus two, and this information should have no bearing on the mortality rate estimates obtained from the tag-return likelihood. Independence between the reporting rate likelihood and the tag-return likelihood would be true if the reporting rate estimates were based on independent tag seeding data. Alternatively, if reporting rates were estimated using observer data, then their estimation would be incorporated directly into the tag-return and catch likelihoods (see Appendix 7). Unfortunately insufficient data from any one source meant that we had to use a complex method for constructing reporting rate estimates for SBT, combining estimates from tag seeding data and from observer data and using some rather ad hoc assumptions. As such, it is unlikely that the reporting rate likelihood is independent of the other likelihoods (since dependence between the observer data and other data sets has not been accounted for). Nevertheless, we do not expect the violation to be serious, nor do we expect the results, namely the mortality rate and abundance estimates, to be appreciably affected.

The overall likelihood L_{Total} can be maximized (or, more commonly done in practice, the negative log of the likelihood can be minimized) to give estimates of the unknown model parameters (listed at the start of the section). However, it must be noted that not all of the natural mortality parameters can be estimated. Information for estimating M_i comes from tagging a single cohort at consecutive ages; in particular, from the differential between the expected returns at age $i+1$ of fish from the cohort released at age i and those released the next year at age $i+1$. Thus, in an experiment in which n consecutive ages of fish from a particular cohort are tagged, estimates can only be obtained for $n-1$ natural mortality rate parameters (regardless of the number of recapture years). For the SBT data being considered, we have a maximum of three consecutive release ages, so we can only estimate two age-specific natural morality rates. In the application to SBT presented in Appendix 5, we addressed this issue by assuming that $M_i = M_2$ for $i \geq 2$, but this is probably not the most realistic assumption for SBT. Natural mortality is generally assumed to decrease with age for

SBT, at least over the young ages we are considering, so in the current analysis we chose to model natural mortality as a linear function of age, and we parameterized the function in terms of natural mortality at the youngest age and the oldest age of returns. For example, if the youngest and oldest returns being included in the model are ages 1 and 5 respectively, then we let

$$M_i = M_1 + \frac{M_5 - M_1}{5 - 1}(i - 1)$$

where M_1 and M_5 are the two parameters to be estimated.

The model fitting was performed using the commercially available software AD Model Builder (Otter Research Ltd., P.O. Box 2040, Sidney BC, V8L 3S3, Canada). The software provides point estimates of the parameters as well as variance estimates calculated using the inverse negative Hessian matrix.

To evaluate model fits, we computed 'standardized' residuals for the return data and catch data. Ordinary residuals are difficult to interpret because the variance differs so much between observations within each data set. For the Gaussian catch data, we defined a standardized residual as

$$\frac{C_{k,i} - \hat{C}_{k,i}}{\upsilon_C \hat{C}_{k,i}}$$

where $\hat{C}_{k,i}$ is the fitted catch value. If the assumption that the catch data are independent Gaussian with coefficient of variation as specified is reasonable, then we expect the standardized residuals have a standard normal distribution (so approximately 95% should fall within the range -2 to 2). With regard to the tag return data, there does not appear to be a conventional way to compute standardized residuals for multinomial data; therefore, we defined a standardized residual as

$$\frac{R_{k,t,a,i} - N_{k,t,a} \hat{p}_{k,t,a,i}}{\sqrt{N_{k,t,a} \hat{p}_{k,t,a,i} \left(1 - \hat{p}_{k,t,a,i}\right)}}$$

where $\hat{p}_{k,t,a,i}$ is the fitted tag return probability. Interpretation of these residuals is not straightforward because they are not independent and their distribution is not evident. If the expected return counts were adequately large, then it would seem reasonable to assume the standardized residuals should follow a standard normal distribution;

however, for the SBT data, many of the expected return counts are close to zero so this is not likely a good approximation. Nevertheless, the standardized residuals provide a rough diagnostic to check for extreme outliers and patterns that indicate a violation of the model. Also, the sign of the standardized residuals is the same as that of the ordinary residuals, so they can be used without reservation in identifying tendencies for under- or over-estimation.

Results

We present results from analyses that included data from cohorts 1989 through 1994. These are the only cohorts with sufficient release and recapture data to warrant inclusion. Additionally, we only included recaptures up to a maximum of age 5 because the numbers of recaptures beyond age 5 are relatively small and because the assumption that natural mortality is a linear function of age is less likely to hold true at older ages. Thus, in the notation presented in the model section, the number of cohorts being modelled is $K = 6$, which we will index by $k = 1989, 1990, \ldots, 1994$ for ease of reference; the minimum age of release/return is $A_{1989} = 2$ and $A_k = 1$ for all other k; the maximum age of release is $B_k = 3$ for all k; and the maximum age of return is $I_k = 5$ for $k = 1989, \ldots, 1992$, $I_{1993} = 4$ and $I_{1994} = 3$ (because fish from cohorts 1993 and 1994 are ages 4 and 3 respectively in 1997, which is the last year being considered).

Before applying the model, we needed to specify a coefficient of variation for the catch at age data (υ_C) and a standard error for the reporting rate estimates (σ_λ). In our initial analyses, we set υ_C to be 0.3 based on results from the SBT analysis presented in Appendix 5, and we set σ_λ to be 0.1 since the uncertainty in the reporting rate estimates is expected to be quite high; however, later we will look at the effect of varying these values.

First, we fit the model with all parameters free except for the constraints already discussed (i.e., natural mortality linear with age). We will refer to this as model 1. Second, we fit the model with the constraint that fishing mortality can be separated

into a multiplicative age and year effect[2]; i.e., we assumed that $F_{k,j} = F_Y(k+j)F_A(j)$, where $F_Y(k+j)$ is the year-specific component of fishing mortality in year $k + j$ and $F_A(j)$ is the age-specific component of fishing mortality at age j (commonly referred to as selectivity). We refer to this as model 2. Note that F_Y and F_A are only unique up to a multiplicative constant because $F_Y F_A = (gF_Y)(F_A/g)$ for any constant g. Therefore, to get a unique solution, we fixed F_A at age 5 to be 1.0.

The parameter estimates from the two models are compared in Figures 1 to 3 (for completeness, the estimates and their standard deviations for the two models are also tabulated in Annex A). The fishing mortality rate estimates by age and cohort shown in Figure 2 for model 2 can be calculated by multiplying the estimated age effects and year effects obtained from the model together; the standard deviations were outputted from the estimation software (but in theory could be calculated explicitly using statistical methods for calculating the variance of the product of two random variables). Both models provide an estimate of the population size at the minimum age of tagging for each cohort. For the 1989 cohort the minimum age of tagging was age 2, whereas for all other cohorts in the model it was age 1. In order to make the abundance estimates comparable between cohorts, we back-calculated an estimate of age 1 abundance for the 1989 cohort. To do so, we used the estimates of age 2 abundance for the 1989 cohort and age 1 natural mortality obtained from the model (which we will denote by $\hat{P}_{1989,1}$ and $\hat{M}_1$ respectively), and brought in external information on the catch of age 1 fish for the 1989 cohort, then solved equations (1) and (2) for both the fishing mortality rate and the population size at age 1 (we denote these by $\tilde{F}_{1989,1}$ and $\tilde{P}_{1989,1}$ to indicate that they are estimates, but not maximum likelihood estimates from the model). We calculated an approximate variance for $\tilde{P}_{1989,1}$ using the formula

---

[2] This constraint was only applied to the general fishing mortality parameters, not to the fishing mortality parameters for newly tagged fish in their first year of tagging (i.e. the F^* parameters).

$$Var\left(\tilde{P}_{1989,1}\right) \approx \left(\exp\left(\tilde{F}_{1989,1} + \hat{M}_1\right)\right)^2 Var\left(\hat{P}_{1989,2}\right),$$

where the variance of $\hat{P}_{1989,2}$ is obtained from the model output. This formula assumes that $\tilde{F}_{1989,1}$ and $\hat{M}_1$ are known without error; although this is not true, it provides a reasonable approximation for our purposes.

The parameter estimates obtained from the two models are very similar; the only parameters for which the error bars on the estimates (defined as plus or minus one standard deviation) do not overlap are $F_{1989,2}$ and $F_{1993,3}$. Natural mortality at age 1 is quite high (~0.4) and decreases to about 0.2 by age 5, but the uncertainty in the age 5 estimate is very high (Figure 1). Fishing mortality is generally close to zero for ages 1 and 2, and for ages 3 to 5 it appears to have increased with cohorts (or years) (Figure 2). These patterns are more apparent if we look at the separate age- and year-effect estimates from model 2; the results from model 2 suggest that selectivity at young ages is dome-shaped with the peak at age 3 (Figure 4, top), and that fishing mortality was a fairly smooth U-shaped function of time over the years of the analysis (Figure 4, bottom). Note that the estimate for 1991 is very high and uncertain (0.54±0.38), the reasons for which are discussed below when we examine the residuals, so it has been omitted from the graph. Cohort abundance appears to have decreased over time, from about 2.5-3 million age 1 fish for the 1989 cohort to just over 1 million age 1 fish for cohorts 1993 and 1994 (Figure 3).

The negative log-likelihood value for model 1 is 25295.0, and for model 2, which has 15 fewer fishing mortality parameters to be estimated than model 1, it is 25326.7 (Table 2). According to Akaike's information criterion (Akaike 1974), which takes the extra number of parameters in model 1 into account, model 1 provides a statistically better fit to the data than model 2. However, the breakdown of the likelihood into its components shows that the difference in the likelihoods is mainly due to model 1 fitting the catch at age data better; the other data sets, in particular the tag-return data, are fitted almost equally well by model 2 (Table 2).

Standardized residuals, as defined in the model section, were computed for the tag-return and catch data for both models. Not surprisingly given the comparison of the likelihood components between the two models, the residuals are very similar between the models for the tag-return data, but are worse for model 2 for the catch data (compare Tables 3 and 4). With regard to the tag-return residuals, an obvious feature is that the residuals for the returns at the same age as release are always very close to zero. This is due to having a unique F^* parameter for every observation corresponding to these residuals. While this is clearly a case of over-fitting, we showed in Appendix 5 and confirmed in our current analysis that the inclusion of F^*'s is necessary to get a good fit. More importantly, because tags tended to be released in the area of the surface fishery either near the beginning or end of the fishing season, the fishing mortality for fish tagged in the year of release would be expected to differ from that for the population as a whole and would also be expected to vary greatly for different releases depending upon the exact release time and location.

The standardized residuals pooled over tagging groups (Tables 3a and 4a) show no obvious outliers or patterns, except perhaps for a tendency for returns at age from the same cohort and release age to all be overestimated or underestimated (indicated by rows of mostly negative values or mostly positive values). Boxplots of the (unpooled) standardized return residuals broken down by a number of factors suggest that the returns for tagger groups 5 and, especially, 6 may be overestimated (Figure 5); otherwise, there is nothing to cause alarm (note that only model 1 results have been plotted because the model 2 results are so similar). The standardized catch residuals for model 1 suggest a very good fit (Table 3b); model 2 does not fit the catch data as well, with age 1 catches for cohorts 1992 to 1994 and age 2 catch for cohort 1989 being notably overestimated. Looking at the catch data (Table 1d), the age 1 catch was much higher in 1991 than in subsequent years of the analysis (almost 50 000 fish in 1991 versus <8000 fish in all other years and <500 in 1992 to 1996).[3] In order for model 2 to estimate an age 1 fishing mortality effect that fits both the large 1991 value and the small values in later years, it must reach a compromise and, thus, ends up

---

[3] Historically, large numbers of age 1 SBT were caught off of Western Australia, but changes in the fishery resulted in 1991 being the last year of any substantive catches in this area (see Discussion).

underestimating the 1991 value and overestimating the others. However, having said this, the age 1 catch in 1991 is not underestimated to the degree we might expect; this is because, in response to the age 1 effect being estimated so low, the year effect for 1991 is estimated very high. There are only two catch observations contributing to the 1991 year effect – age 1 catch from the 1990 cohort and age 2 catch from the 1989 cohort – so the best model fit is achieved with a very high year effect for 1991, which gives a reasonable (although still somewhat underestimated) fit to the large age 1 catch value, but decidedly overestimates the age 2 value. Keep in mind that the model uses not only the catch data but also the tag-return data in estimating age- and year-specific fishing mortality effects; however, for 1991, only the catch data influence the fishing mortality estimates because the two tag-return observations for 1991 are from fish tagged in that same year so the F^* parameters for newly tagged fish apply instead of the fishing mortality parameters for the general population.

We re-fit both models only including data from cohorts 1991 to 1994 (which excludes any data from 1991 or prior). In this case, model 2 provides a significantly better fit than model 1 according to AIC.

To test the sensitivity of the results to the coefficient of variation assumed for the catch data and the standard error assumed for the reporting rate estimates we re-fit model 1, first, keeping σ_λ at 0.1 and varying υ_C and, second, keeping υ_C at 0.3 and varying σ_λ. In both situations, the point estimates of the parameters did not change significantly (all were within one standard deviation of each other), and the uncertainty in the estimates tended to increase as the variability in the data increased. For illustrative purposes, we have shown the results of varying υ_C (Figure 6) and varying σ_λ (Figure 7) on the fishing mortality rate estimates for the 1990 cohort and the age 1 population size estimates. We may have expected the increase in the standard deviation of the parameter estimates to be greater in response to increased uncertainty in the catch data or reporting rate estimates; however, the variance of the parameter estimates is determined by the variability of all data inputs and will tend to be dominated by the data set that is most variable, so changing one component may not necessarily have a large effect.

We also wanted to test the sensitivity of the results to the reporting rate estimates chosen as input to the model. As discussed in the data section, a number of reporting rate options are proposed in Appendix 19 based on a range of assumptions. Although we chose to use the option that is most highly information based, it is also yields the lowest reporting rates; thus, we re-fit model 1 using the option with the highest reporting rates (option 1 of Table 5a, Appendix 19) to evaluate the effect. Not surprisingly, the fishing mortality rates decreased and the population size estimates increased; however, the effect was greatest for the population size estimates (Figure 8). The changes were fairly uniform in that all fishing mortality estimates shifted down by relatively equal amounts, and all population size parameters shifted up by relatively equal amounts. Thus, if relative indices and trends in fishing mortality and abundance are of greater interest than actual magnitude, then the reporting rate option chosen does not matter as much. Note that the natural mortality rate estimates were largely unaffected by the reporting rate option used.

The model estimates of the reporting rates and the tag shedding parameters (see Annex A) have not been discussed. There is little information in the tag-return or catch data to draw these estimates away from their previously estimated values. As such, the model estimates of these parameters are quite similar to the estimates that are inputted. In fact, for the tag shedding parameters they are virtually identical because the standard errors being used for the shedding estimates (σ_ξ and σ_Ω) are so small that there is almost no flexibility in their estimation. The primary reason for including likelihoods for the reporting rate and shedding parameters is to acknowledge their uncertainty and thereby get more realistic variance estimates on the mortality rate and abundance estimates, not to improve the estimates of the reporting rate and shedding parameters themselves.

Estimates of the fishing mortality rate parameters for newly tagged fish (F^*'s) have not been presented. These parameters are not of general interest because they do not represent fishing mortality on the population as a whole. They are simply necessary in order to get realistic estimates of the parameters that are of interest.

Conclusions and Discussion

A comprehensive model for estimating mortality rates and abundance for southern bluefin tuna using tag-return data and catch data has been presented. Two versions of the model were fitted – one in which the age- and year- specific fishing mortality rates were unconstrained (model 1) and one in which they were constrained to have separable, multiplicative age and year effects (model 2). Both models led to similar parameter estimates and the same general conclusions. The results suggest that natural mortality at age 1 is quite high (~0.4) and decreases to about 0.2 by age 5; however, the uncertainty in the age 5 estimate is very high and we found that the estimate is sensitive to changes in either the model or the data inputs. Tagging cohorts at age 4 (in sufficient numbers) in addition to ages 1 to 3 could provide valuable information for better estimating natural mortality at older ages. Fishing mortality is generally close to zero for ages 1 and 2, is greatest at ages 3 and 4, and declines at age 5. The results also suggest that juvenile fishing mortality decreased in the first couple of years of the 1990s then increased fairly steadily from 1994 to 1997. Cohort abundance appears to have decreased from about 2.5-3 million age 1 fish for the 1989 cohort to just over 1 million age 1 fish for cohorts 1993 and 1994.

When fit to the data from cohorts 1989 to 1994, model 1 provided a better fit from a statistical point of view; however, from a practical point of view, model 2 may still be preferred given the fact that it led to very similar parameter estimates using substantially fewer parameters and it also provided better insight into trends in fishing mortality with age and years. However, the lack of fit of model 2 to some of the catch observations highlighted potential problems with separable models when fishing practices (e.g., selectivity) have changed over time. For example, we saw that the SBT fishery caught large numbers of age 1 fish in 1991 then dramatically decreased its catch of age 1 fish in subsequent years. Looking back at catch data prior to 1991 shows that age 1 fish were caught in even larger numbers over the history of the fishery, with over a million age 1 fish being caught in 1983. A large portion of the historic age 1 catches occurred off of Western Australia, but when joint venture fishing opportunities began in 1992 fishing off of Western Australia decreased substantially. It is important that such changes in selectivity over time are recognized and accounted for when estimating separate age and year effects. Because selectivity

was relatively constant over most years included in our analysis, model 2 was still able to provide a reasonable fit. When we re-ran the analysis only including data from cohorts 1991 to 1994 in order to exclude any data from 1991 or prior, we found that model 2 provided a statistically better fit than model 1. These results suggest that a separable model is appropriate for SBT, and illustrate the importance of incorporating changes in selectivity.

Reporting rates and tag shedding rates were estimated from independent analyses for input to the model. We accounted for their uncertainty by including additional likelihood terms for the estimates and their standard errors (similar to putting a prior on the parameters in a Bayesian framework). Ideally, we would estimate reporting rates directly within the model (for example, using observer data as described in Appendix 7), but unfortunately the observer data for SBT are insufficient to let us do so. In the case of the shedding parameters, it is possible to directly incorporate their estimation into the model, as described in Appendix 14 (equations 10 and 11); however, this would require breaking the returns down not only by cohort, release age, return age and tagging group, but also according to whether one tag or both tags were returned. Many of the return counts will be very small when broken down to this level and will likely introduce estimation problems. Furthermore, the estimates of the shedding parameters will almost certainly be similar to those obtained from the independent analysis since the only information available for estimating them (i.e., comparing numbers of returns with one tag versus two) is the same in both cases.

Instead of modelling the reporting rate estimates directly, we modelled a variable representing the number of reported tag returns using a binomial distribution and an estimated effective sample size. It would have been more straightforward to model the reporting rate estimates themselves as having, say, a beta distribution (this was the approach taken in Appendix 9 for modelling the reporting rates in the surface fishery component of the 2-fishery model). However, the maximum likelihood estimates of the parameters of a beta distribution correspond to the mode, not the mean, so that the reporting rate estimates obtained using a beta likelihood are not equal to the reporting rate estimates inputted as data to the likelihood. In fact, they can be substantially different when the distribution is highly skewed. This result is undesirable because, in the absence of any other information, we do not want the reporting rate estimates to

change. To resolve this problem, we used the binomial approach. Using a normal distribution for the reporting rates would also have resolved this problem but it would not have been a realistic choice because it does not constrain the estimates to be between 0 and 1. This is not necessarily a problem if the estimates are sufficiently away from these bounds and/or have small variances, but this is not true for many of the reporting rate estimates. We were able to use a bivariate normal distribution for the shedding parameters, even though they too should be constrained between 0 and 1, because their estimated variances were so small that the normal approximation was adequate.

The model assumes that catch numbers between ages within a year are independent. Although this is not true conditional on the total catch in a year (since catching more fish at one age means catching less fish at another age to achieve the same total), it is a reasonable assumption unconditionally. For example, consider the following argument: if the total catch within a year is random and follows a Poisson distribution, and the distribution of the age counts conditional on the total is multinomial, then the unconditional age counts are independent Poisson. At large catch sizes, these can be approximated as independent Gaussian, which is what we have done.

Even after taking measures to incorporate uncertainty in the reporting rates and shedding rates into the model, the standard errors of the mortality rate and abundance estimates may still to be underestimated. This is because the variance in the number of returns is likely to be greater than predicted by a multinomial distribution due to incomplete mixing and heterogeneity in the capture probabilities of fish. One way of accounting for overdispersion in the tag-return data is to model the data as Dirichlet-multinomial, as described in Appendix 9. To do so requires an assumption be made about the level of overdispersion, either assuming it is known or keeping it constant since it cannot be estimated otherwise. The necessity for, and potential gain from, incorporating overdispersion in the tag-return model for southern bluefin tuna is an area for further investigation.

References

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, **19**: 716-723.

Anon. 2005. Report of the special management procedure technical meeting. Commission for the Conservation of Southern Bluefin Tuna. 15-18 February. Seattle, USA.

Hoenig, J.M., Barrowman, N.J., Pollock, K.H., Brooks, E.N., Hearn, W.S., and Polacheck, T. 1998. Models for tagging data that allow for incomplete mixing of newly tagged animals. Can. J. Fish. Aquat. Sci. 55: 1477-1483.

Kolody, D., and Polacheck, T. 2001. Application of a statistical catch-at-age and length integrated analysis model for the assessment of southern bluefin tuna stock dynamics 1951-2000. Commission for the Conservation of Southern Bluefin Tuna. CCSBT-SC/0108/13.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1996. Estimation of natural and fishing mortality rates for juvenile southern bluefin tuna based on multi-year tagging of cohorts. Commission for the Conservation of Southern Bluefin Tuna. CCSBT/SC/96/11.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1997. Updated estimates of mortality rates for juvenile southern bluefin tuna from multi-year tagging cohorts 1997. Council for the Conservation of Southern Bluefin Tuna. CCSBT/SC/9707/26.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1998. Updated estimates of mortality rates for juvenile southern bluefin tuna from multi-year tagging cohorts 1998. Council for the Conservation of Southern Bluefin Tuna. CCSBT/SC/9807/20.

Polacheck, T., Preece, A. and Ricard, D. 2001. Assessment of the status of the southern bluefin tuna stock using virtual population analyses – 2001. Council for the Conservation of Southern Bluefin Tuna. CCSBT-SC/0108/20.

Preece, A., Polacheck, T., Kolody, D., Eveson, P., Ricard, D., Jumppanen, P., Farley, J., and Davis, T. 2001. Summary of the primary data inputs to CSIRO's 2001

stock assessment models. Commission for the Conservation of Southern Bluefin Tuna. CCSBT-SC/0108/21.

Preece, A., Cooper, S., and Hartog, J. 2004. Data post-processing for input to the 2004 stock assessments and comparisons of 2001 and 2004 assessment datasets. Commission for the Conservation of Southern Bluefin Tuna. CCSBT-ESC/0409/27.

Table 1. Summary of the four southern bluefin tuna data sets used as input to the model.

a) Tag-return data

Cohort	Release year	Release age	Number releases	Number returns by year						
				1991	1992	1993	1994	1995	1996	1997
1988	1991	3	810	63	8	16	7	1	5	1
1989	1991	2	3127	103	148	59	34	20	7	5
	1992	3	1097		57	18	11	9	3	2
1990	1991	1	3299	20	40	46	23	13	5	4
	1992	2	4646		88	159	101	33	12	8
	1993	3	2777			66	78	32	17	15
1991	1992	1	2144		1	21	56	37	11	7
	1993	2	2937			60	68	67	21	11
	1994	3	3640				77	145	30	40
1992	1993	1	4898			2	41	201	91	58
	1994	2	3158				29	167	76	52
	1995	3	2629					55	103	74
1993	1994	1	9003				4	110	401	364
	1995	2	5899					83	395	363
	1996	3	1511						115	201
1994	1995	1	8585					0	87	622
	1996	2	2518						77	339
	1997	3	526							91
1995	1996	1	82						0	3
	1997	2	592							15
1996	1997	1	884							1

b) Reporting rate estimates

Cohort	Reporting rate estimate, $\hat{\lambda}$						
	1991	1992	1993	1994	1995	1996	1997
1988	0.597	0.327	0.402	0.390	0.179	0.254	n/a
1989	0.654	0.543	0.471	0.457	0.192	0.275	0.596
1990	0.933	0.577	0.625	0.559	0.258	0.267	0.568
1991		0.887	0.750	0.600	0.388	0.250	0.537
1992			0.926	0.498	0.622	0.411	0.597
1993				0.522	0.592	0.474	0.639
1994					0.725	0.388	0.727
1995						0.321	0.775
1996							0.805

c) Tag shedding data (parameter estimates, standard errors and correlations)

Tagger Group	$\hat{\xi}$	σ_ξ	$\hat{\Omega}$	σ_Ω	ρ
1	0.974	0.007	0.039	0.004	0.005
2	0.961	0.012	0.049	0.006	0.008
3	1.000	0.000	0.067	0.004	0.000
4	1.000	0.000	0.093	0.006	0.000
5	0.934	0.040	0.089	0.023	0.028
6	0.967	0.022	0.160	0.016	0.016

d) Catch data

Cohort	Number fish caught						
	1991	1992	1993	1994	1995	1996	1997
1988	176057	77731	48640	24928	20560	15357	11443
1989	76744	150758	65802	32144	27442	18972	17492
1990	48450	33638	120232	72806	39073	24743	21673
1991		7624	38414	119166	61080	38646	27398
1992			404	10398	133300	76136	43001
1993				187	30789	171859	72177
1994					416	26276	203883
1995						422	32025
1996							1965

Table 2. Negative log-likelihood values for models 1 and 2. The total as well as breakdown into likelihood components is given.

Component	Model 1	Model 2
Tag-return	24691.0	24695.1
Reporting rates	360.0	361.5
Shedding rates	2.2	2.2
Catch	241.8	267.8
Total	25295.0	25326.7

Table 3. Standardized tag-return and catch residuals for the model with unconstrained fishing mortality rates (model 1).

a) standardized tag-return residuals, pooled over tagging groups[4]

| | Release | Recapture age | | | | |
Cohort	age	1	2	3	4	5
1989	2		0.1	0.1	1.4	1.1
1989	3			-0.1	-1.9	-1.2
1990	1	0.0	0.4	-2.4	-2.5	-0.4
1990	2		0.1	1.3	1.8	0.2
1990	3			0.0	-0.3	0.3
1991	1	0.0	-0.1	2.4	0.5	0.9
1991	2		0.0	-2.0	-0.5	0.7
1991	3			0.0	0.4	-0.8
1992	1	0.0	0.7	1.3	1.5	0.6
1992	2		0.0	-1.3	-0.4	-0.4
1992	3			0.0	-0.9	-0.3
1993	1	0.0	-0.1	0.5	0.4	
1993	2		0.0	-1.1	-1.4	
1993	3			0.6	1.5	
1994	1	0.0	0.1	-1.7		
1994	2		0.3	2.4		
1994	3			0.0		

b) standardized catch residuals

Cohort	Age 1	Age 2	Age 3	Age 4	Age 5
1989		0.3	0.3	0.5	0.0
1990	0.3	-1.0	0.9	0.7	-0.1
1991	0.3	0.4	1.0	-1.0	0.1
1992	0.3	-1.8	0.8	0.3	0.6
1993	0.3	0.7	0.1	-0.1	
1994	0.3	-0.4	0.8		

---

[4] Pooled standardized residuals over tagging groups were calculated as

$$\sum_{t=1}^{6}\left(R_{k,t,a,i} - N_{k,t,a}\,\hat{p}_{k,t,a,i}\right)\Big/ \sqrt{\sum_{t=1}^{6} N_{k,t,a}\,\hat{p}_{k,t,a,i}\left(1 - \hat{p}_{k,t,a,i}\right)} \ .$$

Table 4. Standardized recapture and catch residuals for the model with fishing mortality rates constrained to have separable age and year effects (model 2).
a) standardized tag-return residuals, pooled over tagger groups (see footnote to Table 3)

Cohort	Release age	Recapture age 1	2	3	4	5
1989	2		0.1	0	1.6	1.9
1989	3			-0.1	-1.9	-0.9
1990	1	0.0	0.5	-2.5	-2.5	-0.4
1990	2		0.1	1.2	1.8	0.3
1990	3			0.0	-0.5	0.2
1991	1	0.0	0.4	2.2	0.6	0.8
1991	2		0.0	-2.1	-0.5	0.7
1991	3			0.0	0.2	-1.0
1992	1	0.0	0.9	1.3	1.5	0.5
1992	2		-0.1	-1.4	-0.3	-0.4
1992	3			0.0	-1.1	-0.6
1993	1	0.0	0.2	0.7	0.1	
1993	2		0.0	-0.8	-1.6	
1993	3			0.4	2.6	
1994	1	0.0	-0.4	-1.8		
1994	2		0.4	2.4		
1994	3			0.0		

b) standardized catch residuals

Cohort	Age1	Age2	Age3	Age4	Age5
1989		-2.5	-0.1	0.6	0.9
1990	1.0	-1.0	0.5	0.4	-0.2
1991	1.0	1.0	0.3	-1.2	-0.8
1992	-2.7	-1.5	1.4	0.3	0.5
1993	-3.1	0.3	1.2	-0.7	
1994	-2.9	-0.7	1.3		

Figure 1. Comparison of natural mortality rate (*M*) estimates and their standard deviations (SD) by age for the model with unconstrained fishing mortality rates (model 1) and the model with fishing mortality rates constrained to have separable age and year effects (model 2). Black circle = model 1 results; blue triangle = model 2 results.
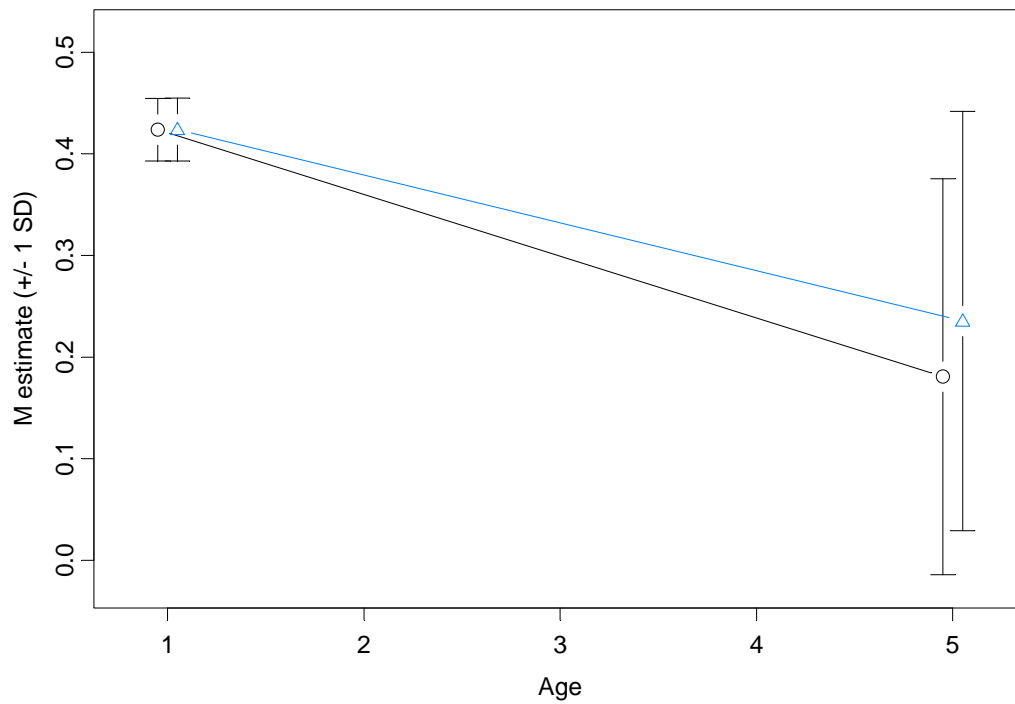
Figure 2. Comparison of fishing mortality rate (F) estimates and their standard deviations (SD) by cohort and age for the model with unconstrained fishing mortality rates (model 1) and the model with fishing mortality rates constrained to have separable age and year effects (model 2). Black circle = model 1 results; blue triangle = model 2 results.

Figure 3. Comparison of population size (P) at age 1 estimates and their standard deviations (SD) by cohort for the model with unconstrained fishing mortality rates (model 1) and the model with fishing mortality rates constrained to have separable age and year effects (model 2). Black circle = model 1 results; blue triangle = model 2 results. For the 1989 cohort, only a direct estimate of P at age 2 is obtained from the models, so the age 1 estimates shown are post-calculated (see text).
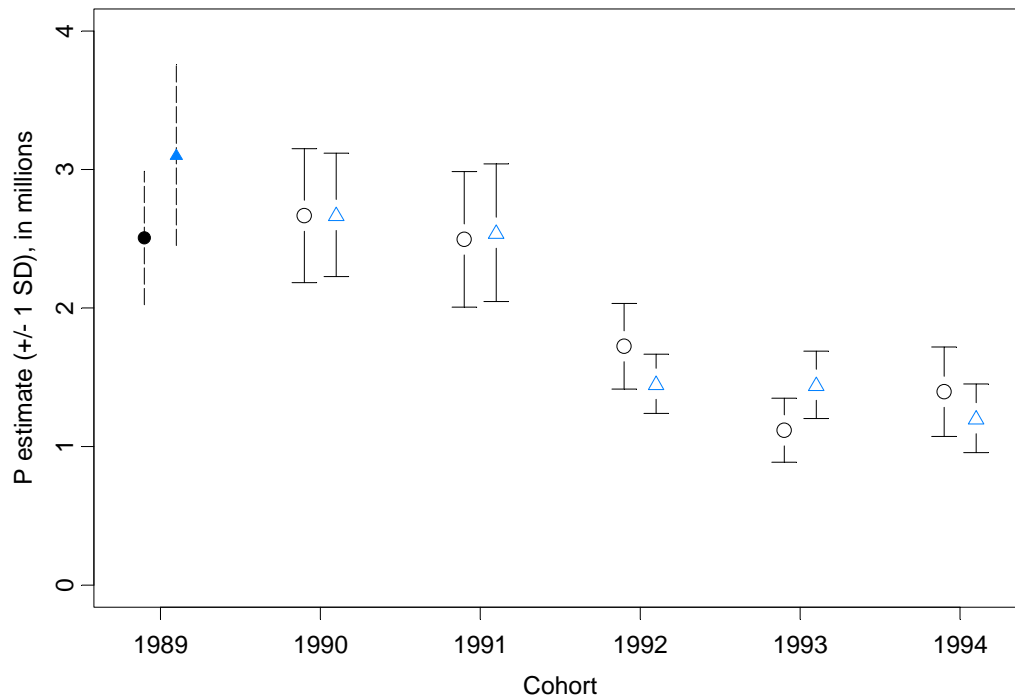
Figure 4. Estimates (± 1 standard deviation) of the age-specific fishing mortality rate effect (upper panel) and the year-specific fishing mortality effect (lower panel) for the model with separable fishing mortality rates (model 2). Note that the estimates should be interpreted as relative indices; the age effect at age 5 has been fixed at 1.0 (see text). The estimate of the year-specific component for 1991 has been omitted because it is based on very little data and has large uncertainty associated with it (refer to text).
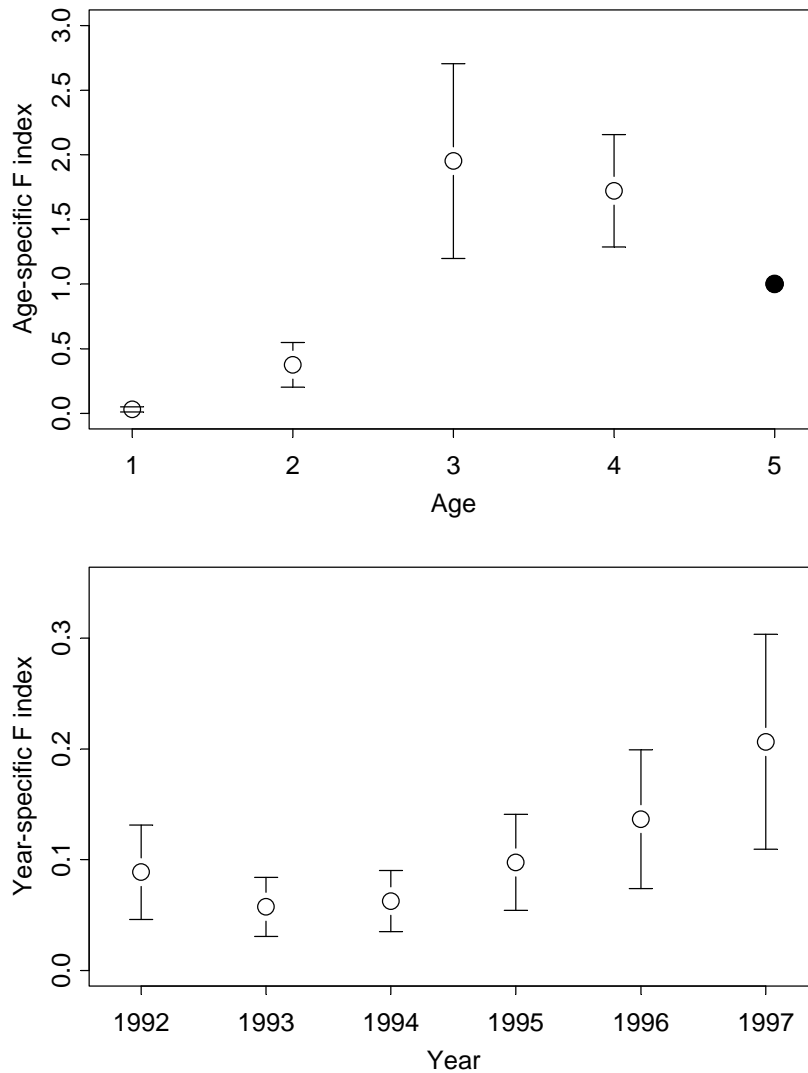
Figure 5. Boxplots of standardized recapture residuals broken down by various factors for the model with unconstrained fishing mortality rates (model 1).
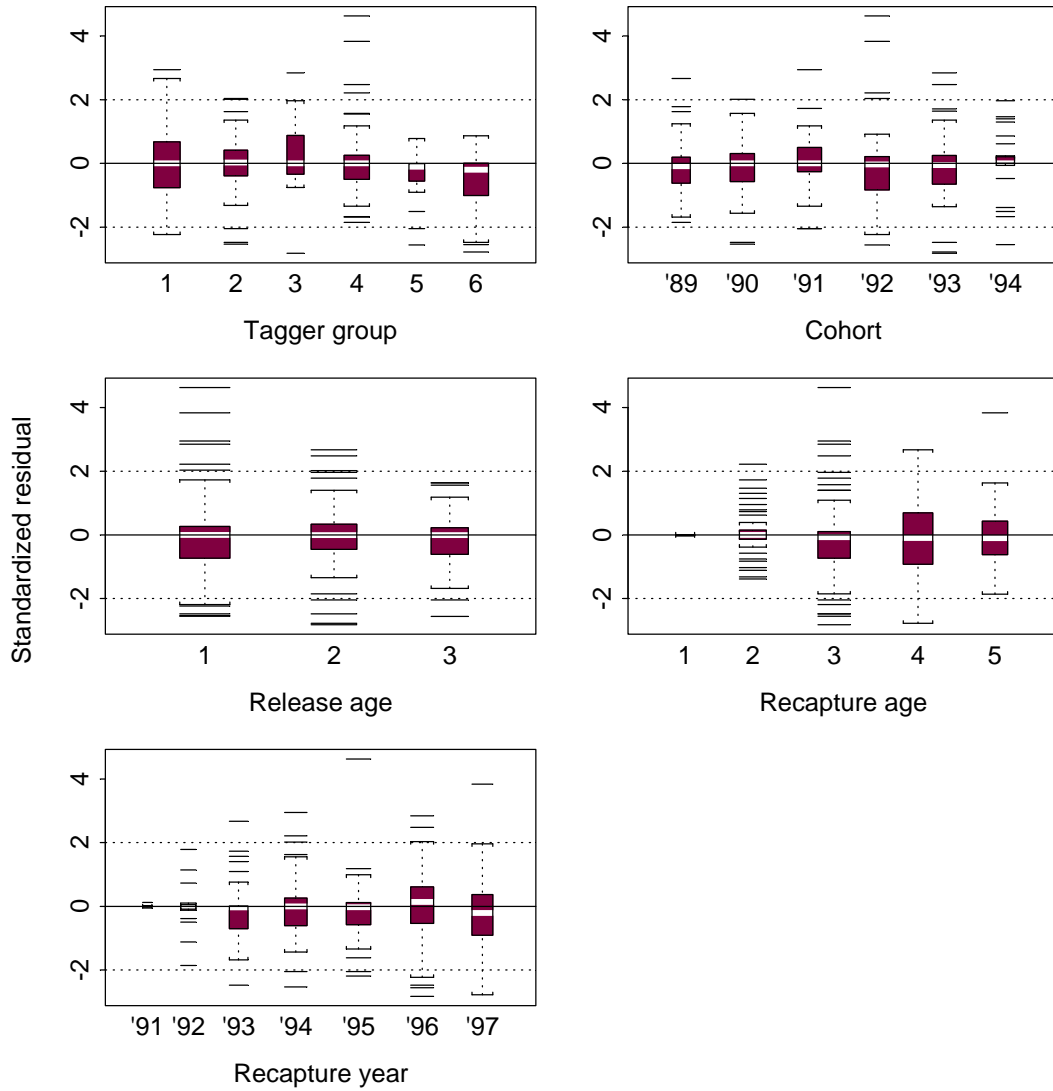
Figure 6. Effect of varying the coefficient of variation of the catch data ("catch CV") on the fishing mortality rate estimates for the 1990 cohort (top) and the age 1 population size estimates (bottom).
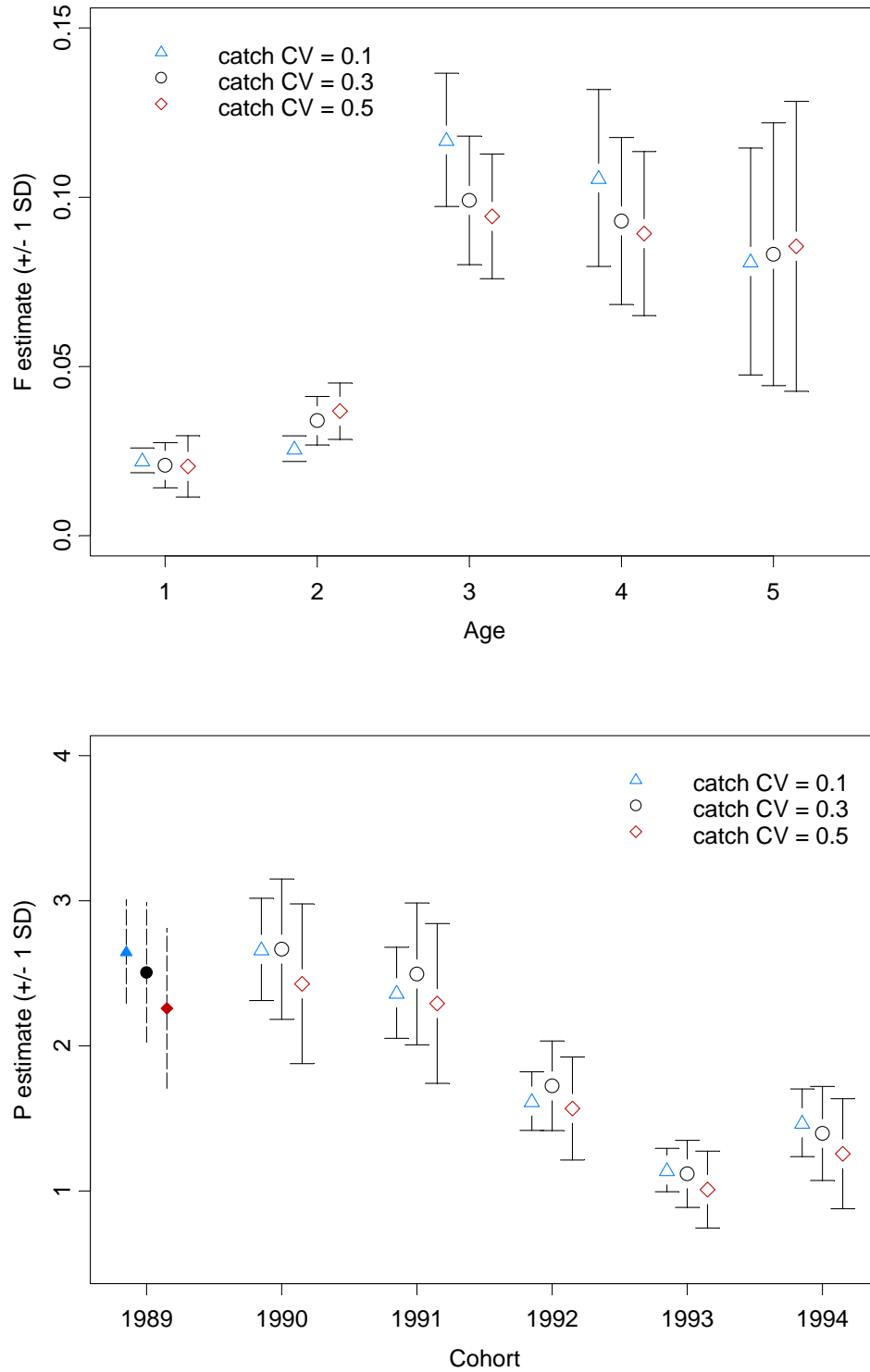
Figure 7. Effect of varying standard error of reporting rate estimates ("RR SE") on the fishing mortality rate estimates for the 1990 cohort (top) and the age 1 population size estimates (bottom).
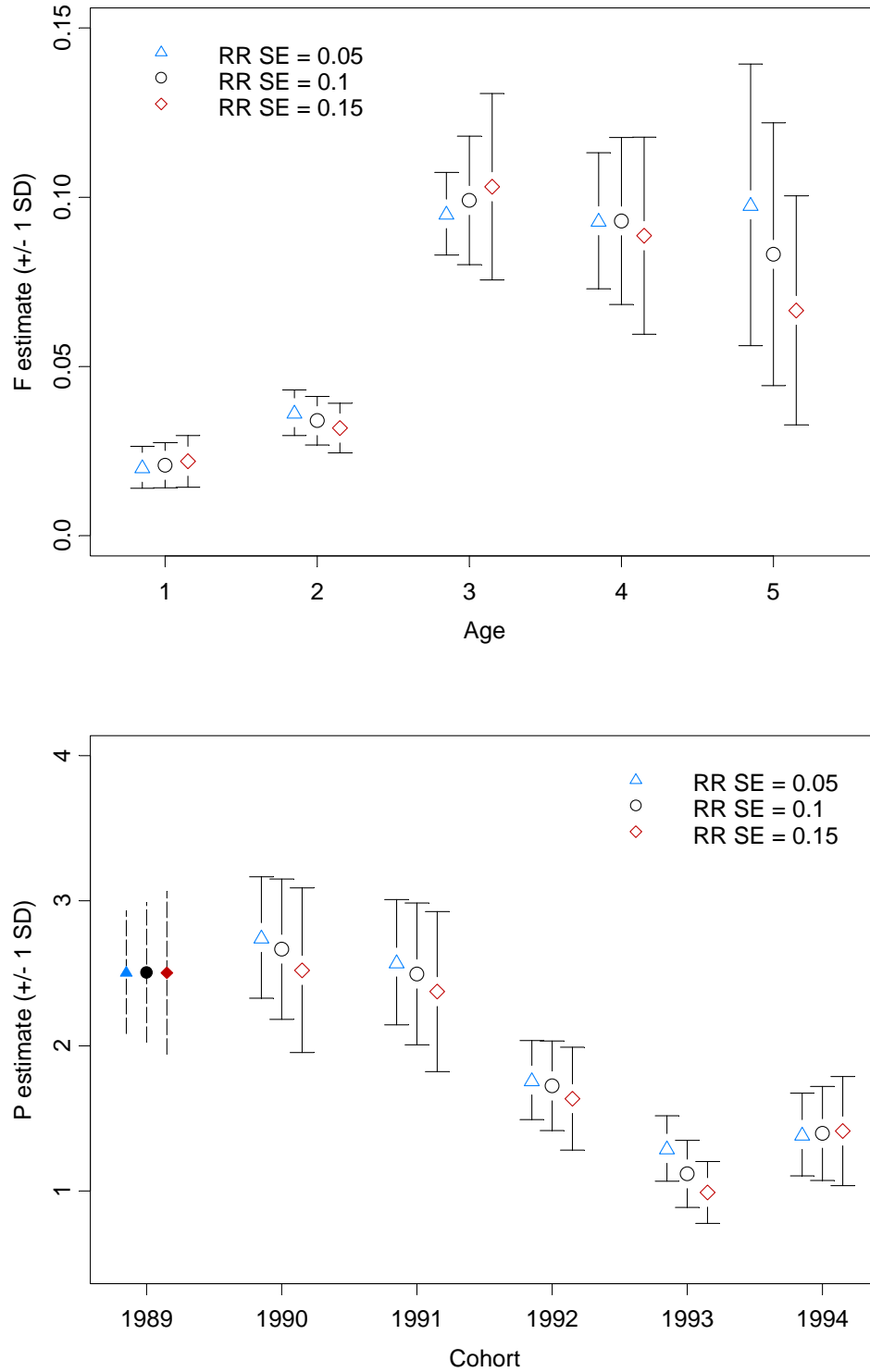
Figure 8. Effect of using reporting rate option 1 (high reporting rates) versus reporting rate option 8 (low reporting rates) on the fishing mortality rate estimates for the 1990 cohort (top) and the age 1 population size estimates (bottom).
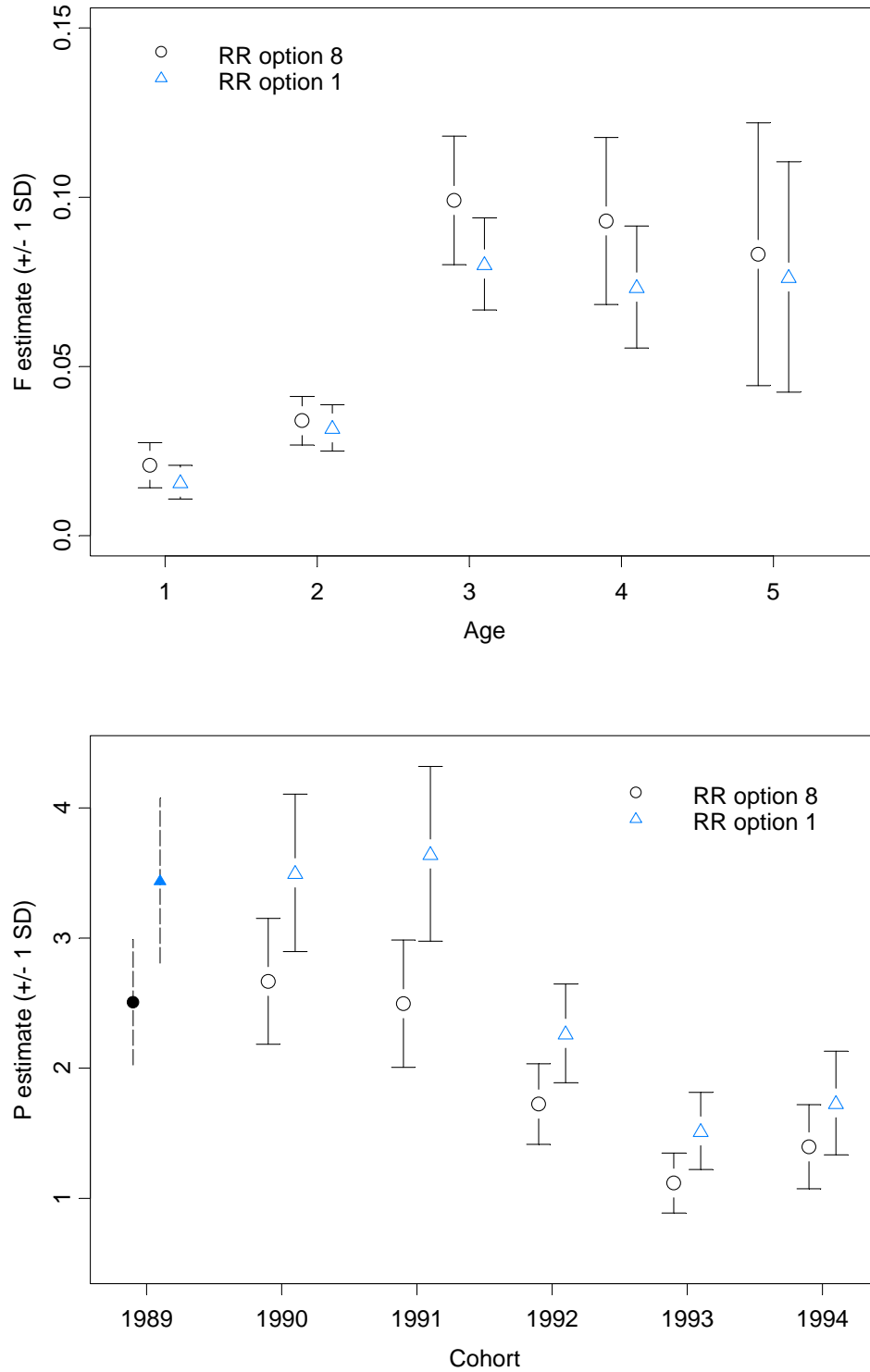
Annex A

Table A1. Parameter estimates obtained from model with unconstrained fishing mortality rates (model 1). Standard error estimates are given in parentheses below the point estimates. Note that values for the population size at the initial age of tagging, P_{A_k}, are in millions, and that $A_k = 1$ for all cohorts except 1989, for which $A_k = 2$. The F^* estimates of fishing mortality for newly tagged fish are not of primary interest and are not shown.

M_1	M_5
0.424	0.181
(0.031)	(0.195)

Cohort	F_1	F_2	F_3	F_4	F_5	P_{A_k}
1989	–	0.054	0.162	0.102	0.076	1.62
		(0.017)	(0.031)	(0.030)	(0.034)	(0.31)
1990	0.021	0.034	0.099	0.093	0.083	2.67
	(0.007)	(0.007)	(0.019)	(0.025)	(0.039)	(0.48)
1991	0.003	0.025	0.101	0.144	0.085	2.50
	(0.001)	(0.005)	(0.021)	(0.041)	(0.041)	(0.49)
1992	0.000	0.024	0.177	0.183	0.136	1.72
	(0.000)	(0.006)	(0.029)	(0.054)	(0.061)	(0.31)
1993	0.000	0.042	0.489	0.475	–	1.12
	(0.000)	(0.007)	(0.158)	(0.156)		(0.23)
1994	0.000	0.040	0.372	–	–	1.40
	(0.000)	(0.008)	(0.075)			(0.32)

Cohort	λ_1	λ_2	λ_3	λ_4	λ_5
1989	–	0.679	0.558	0.455	0.476
		(0.093)	(0.088)	(0.087)	(0.088)
1990	0.930	0.632	0.570	0.528	0.306
	(0.105)	(0.088)	(0.094)	(0.090)	(0.078)
1991	0.887	0.741	0.531	0.464	0.276
	(0.100)	(0.098)	(0.094)	(0.091)	(0.073)
1992	0.926	0.571	0.620	0.413	0.570
	(0.100)	(0.091)	(0.086)	(0.082)	(0.091)
1993	0.520	0.563	0.315	0.662	–
	(0.100)	(0.092)	(0.085)	(0.090)	
1994	0.725	0.480	0.695	–	–
	(0.100)	(0.084)	(0.104)		

Tagger Group	ξ	Ω
1	0.974	0.039
	(0.005)	(0.003)
2	0.961	0.049
	(0.008)	(0.004)
3	1.000	0.067
	(–)	(0.003)
4	1.000	0.092
	(–)	(0.004)
5	0.921	0.096
	(0.029)	(0.016)
6	0.951	0.179
	(0.016)	(0.011)

Table A2. Parameter estimates obtained from model with fishing mortality rates constrained to have separable age and year effects (model 2). Standard error estimates are given in parentheses below the point estimates. Note that values for the population size at the initial age of tagging, P_{A_k}, are in millions, and that $A_k = 1$ for all cohorts except 1989, for which $A_k = 2$. The F^* estimates of fishing mortality for newly tagged fish are not of primary interest and are not shown.

M_1	M_5
0.424	0.236
(0.031)	(0.206)

	Age				
	1	2	3	4	5
F_A	0.032	0.377	1.952	1.722	1.0
	(0.020)	(0.172)	(0.754)	(0.435)	--

	Year						
	1991	1992	1993	1994	1995	1996	1997
F_Y	0.541	0.089	0.058	0.063	0.098	0.137	0.206
	(0.378)	(0.043)	(0.027)	(0.028)	(0.043)	(0.063)	(0.097)

	Cohort					
	1989	1990	1991	1992	1993	1994
P_{A_k}	2.01	2.67	2.54	1.45	1.45	1.2
	(0.42)	(0.45)	(0.50)	(0.21)	(0.24)	(0.25)

Cohort	λ_1	λ_2	λ_3	λ_4	λ_5
1989	–	0.682	0.542	0.49	0.565
		0.092	0.076	0.070	0.075
1990	0.930	0.637	0.524	0.492	0.298
	0.105	0.079	0.071	0.064	0.053
1991	0.887	0.770	0.461	0.431	0.208
	0.100	0.086	0.063	0.056	0.040
1992	0.926	0.582	0.595	0.352	0.466
	0.100	0.076	0.063	0.048	0.065
1993	0.520	0.627	0.529	0.718	–
	0.100	0.072	0.061	0.068	
1994	0.725	0.397	0.677	–	–
	0.100	0.060	0.075		

Tagger Group	ξ	Ω
1	0.974	0.039
	(0.005)	(0.003)
2	0.961	0.049
	(0.008)	(0.004)
3	1.000	0.067
	(–)	(0.003)
4	1.000	0.092
	(–)	(0.004)
5	0.921	0.097
	(0.029)	(0.016)
6	0.951	0.179
	(0.016)	(0.011)

Appendix 16:

Application of a spatial integrated tag-recapture and catch model to 1990s southern bluefin tuna data

Geoff M. Laslett, J. Paige Eveson and Tom Polacheck

FRDC Project 2002/015

1   Introduction

In Appendix 11 we developed a spatial multi-year tagging and catch-at-age model that was designed to have spatial and temporal dynamics resembling those of the current southern bluefin tuna (SBT) fishery and juvenile stock. The fisheries were divided into two seasonal components - a summer surface fishery and a winter longline fishery. The summer surface fishery occurs only in the Great Australian Bight (GAB) off South Australia, whereas the winter longline fishery occurs in three regions distinct from where the surface fishery occurs (waters off of South Africa, the South-East Indian Ocean and the Tasman Sea). All fish migrate out of the GAB at the end of the summer season into one of the longline regions, but only a fraction of the fish migrate from the longline regions to the GAB at the end of the winter season (for simplicity, we assume that all migrations between regions occur at the end of a season). Movement rates are allowed to vary between regions and with ages (e.g., a differing percent of fish migrate from the Tasman Sea region to the GAB with age). To simplify the migration dynamics, we do not allow for direct migration of fish between the longline regions. Thus, over the course of a single year, a fish either remains within a single longline region or migrates between one of the longline regions and the GAB. Two variations on the basic movement dynamics were considered: one in which movement is represented as a Markov process, meaning that a fish has no memory with respect to its previous migrations; and one in which there is site-fidelity in the movement dynamics, so that a fish that migrates from a particular longline region to the GAB will always return to that same longline region at the end of the summer season.

In the current appendix, we apply the spatial multi-year tagging and catch-at-age model with Markovian movement dynamics to SBT data collected from 1991 to 1997. The model was presented in Appendix 11 in terms of a single cohort of fish. Because we have SBT data from multiple cohorts, we first discuss the model for multiple cohorts. In one sense this is trivial, because the multicohort model is just a juxtaposition of several single cohort models. There are, however, some issues that

become more apparent with multiple cohorts, such as: data from partial cohorts (i.e., some cohorts being tagged at only a subset of ages); model simplification through assuming that certain parameters are common to all cohorts or display simple trends with cohort.

In addition to the adjusting the model for multiple cohorts, some further modifications need to be made before applying the model to the SBT data. The models in Appendix 11 were developed in terms of how best to design future SBT tagging programs, therefore they assume that observer data are available in each longline region in each winter season and that tag seeding data are available in the GAB in each summer season with which to estimate reporting rates. In fact, this is not the case for the 1990s SBT data; although some limited observer data and tag seeding data exist, they are insufficient for direct inclusion in the model. Instead, we use the available data and a number of assumptions to come up with reporting rate estimates by year, age and region for SBT from 1991 to 1997 (more details are given later). These estimates are then brought into the spatial tagging and catch model as if they are known without error.

Moreover, recent analyses of the double-tagging data for SBT (Appendix 14) suggest that tag shedding can be substantial for some taggers, so estimates of mortality rates and abundance that do not take tag shedding into account are likely to be biased. As such, we modify the model to incorporate instantaneous and continuous tag shedding parameters. Using the same approach to the reporting rates, we bring estimates of the tag shedding parameters into the model as being known without error.

Finally, we modify the model to allow for an initial period of non-mixing in the tag-return probabilities. This is to allow for the fact that newly tagged fish will not be fully mixed with the untagged population immediately after tagging, and for the fact that tagging of SBT generally occurred during or near the end of the fishing season so tagged fish are only vulnerable for part of the season. The general approach used is the same as that used in the non-spatial analysis of the 1990s SBT

tag-recapture and catch data in Appendix 15.

2   The multi-cohort model

2.1   Probability of survival or being caught

We suppose that there are two seasons each year and that there are R regions. Fishing takes place in region 1 in season 1, and in regions 2 to R in season 2. It is convenient to work in terms of time periods t, taking values $1, 2, 3, 4, \ldots, T$, where there are two periods per year. The season is given by $s = t - 2[(t-1)/2]$. The year is given by $y = 1 + [(t-1)/2]$. In these definitions, and in the definition of age a below, $[x]$ means the integer part of x.

The probability of survival of a fish during a time period, conditional on its being alive at the beginning of the period, depends on both natural mortality and fishing mortality. Before we define these explicitly, we specify their dependence on age, time period and region. We assume natural mortality depends only on age, denoted by M_a. We assume the two seasons are of equal length and that natural mortality occurs at a constant rate so that the natural mortality in each season is $M_a/2$.

In the model presented in Appendix 11 for one cohort, we let fishing mortality depend on age, season and region. With only one cohort, age and season uniquely define the time period so in fact this is equivalent to letting fishing mortality depend on time period and region. It is feasible that fishing mortality will vary not only between time periods (because fishing practices may change over time) but also between fish of different ages (because different gear types may target different age classes). As such, when more than one cohort is being modelled, we allow for the fishing mortality to depend on time t, age a and region r, denoted by F_{atr}.

For the SBT fishery with $R = 4$, the F_{atr} values will have the structure:

$$
F_{at} = \begin{pmatrix} F_{at1} \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ for odd } t \qquad\qquad F_{at} = \begin{pmatrix} 0 \\ F_{at2} \\ F_{at3} \\ F_{at4} \end{pmatrix} \text{ for even } t
$$

Here $F_{at1}, F_{at2}, F_{at3}, F_{at4}$ denote non-zero values. Note that fishing only takes place in region 1 in the summer season, and in regions 2, 3, 4 for the winter season, so the fishing mortality is 0 in the complementary regions.

We now state explicitly the meaning of M_a and F_{atr}. The probability that a fish of age a, alive in region r at the beginning of time period t, survives until the end of the time period, is

$$
S_{atr} = e^{-M_a/2 - F_{atr}} \ .
$$

The probability that it does not survive, and is caught, is $(1 - S_{atr})f_{atr}$, where

$$
f_{atr} = \frac{F_{atr}}{M_a/2 + F_{atr}}
$$

is the conditional probability that it is caught rather than dies naturally. Likewise the probability that it does not survive, and dies naturally, is $(1 - S_{atr})(1 - f_{atr})$. This is a type of competing risks model, and the mathematical rationale behind it is explained in Annex A.

Consider a fish alive in region r_C at the beginning of period t_C at age a_C. Let $s_{t_C} = t_C - 2[(t_C - 1)/2]$. Then this fish will have age (in years) of $a = a(a_C, t_C, t) = a_C + [(t - t_C - 1 + s_{t_C})/2]$ in period $t \geq t_C$. Let R_t be the region that the fish is in at the start and during time interval t, and $L_t = 1$ if the fish is alive at the start of time period t. Hence

$$
\Pr\{R_{t+1} = r, L_{t+1} = 1 | a_C, t_C, r_C\} =
$$
$$
\sum_{r'=1}^{R} \Pr\{R_t = r', L_t = 1 | a_C, t_C, r_C\} S_{atr'} \, \pi_{r'r|at} \tag{1}
$$

where $\pi_{r'r|at}$ represents the probability of a fish of age a moving from region r' to r at the end of time period t. Looking at the right-hand side of this equation, we

read that the fish was alive in region r' at time t with probability $\Pr\{R_t = r', L_t = 1 | a_C, t_C, r_C\}$, survived the time period with probability $S_{atr'}$, and then moved from region r' to region r with transition probability $\pi_{r'r|at}$. Note that the transition probabilities do not depend on any of the fish's history prior to time t.

In our presentation of the model we assume that the transition probabilities are a function of the fish age $a = a(t)$ and season $s = s(t)$, rather than t directly. That is, we assume that fish of the same age in the same season would have the same migration patterns (transition probabilities) across all years. Therefore, we let Π_{as} denote the matrix of transition probabilities for a fish of age a at the end of season s. We assume the following structure for these matrices ($R = 4$). For season 1:

$$
\Pi_{a1} \;=\;
\begin{pmatrix}
0 & \pi_{12|a1} & \pi_{13|a1} & 1 - \pi_{12|a1} - \pi_{13|a1} \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
$$

For season 2:

$$
\Pi_{a2} \;=\;
\begin{pmatrix}
1 & 0 & 0 & 0 \\
\pi_{21|a2} & 1 - \pi_{21|a2} & 0 & 0 \\
\pi_{31|a2} & 0 & 1 - \pi_{31|a2} & 0 \\
\pi_{41|a2} & 0 & 0 & 1 - \pi_{41|a2}
\end{pmatrix}
$$

These transition matrices essentially say that at the end of season 1, a fish in region 1 moves to either region 2, 3 or 4, whereas a fish in region 2, 3 or 4 stays in that region. At the end of season 2, a fish stays in its current region or moves to region 1. Note however that a fish could make the following transitions: regions 1, 2, 1, 3 in seasons 1, 2, 1, 2 in time periods 1, 2, 3, 4 respectively. Thus, it need not remain faithful to its winter (season 2) region.

A16 - 6

Now let us suppose S_{at} is the $R \times R$ diagonal matrix with rth diagonal element S_{atr}. Also, let $\Pi_{a(t)s(t)}$ denote the matrix of transition probabilities for a fish of age a during time interval t. Then, equation (1) can be expressed in matrix form as

$$p_{t+1|a_Ct_Cr_C} = \Pi'_{a(t)s(t)} \, S_{at} \, p_{t|a_Ct_Cr_C} \, ,$$

where $p_{t|a_Ct_Cr_C}$ is a vector of length R whose rth entry is the probability that a fish, alive at age a_C in region r_C at the beginning of time period t_C, is alive in region r at the beginning of time period t. This equation can be calculated iteratively using as a starting vector that fact that the rth entry of $p_{t_C|a_Ct_Cr_C}$ is 1 for $r = r_C$ and 0 for $r \neq r_C$.

The probability that a fish, alive at age a_C in region r_C at the beginning of time period t_C, is caught during time period t in region r is

$$q_{tr|a_Ct_Cr_C} = f_{atr}(1 - S_{atr})p_{tr|a_Ct_Cr_C} \tag{2}$$

where $p_{tr|a_Ct_Cr_C}$ (the rth entry of vector $p_{t|a_Ct_Cr_C}$ above) is the probability that the fish is alive in region r at the beginning of time interval t, $1 - S_{atr}$ is the probability that the fish does not survive the time period, and f_{atr} is the conditional probability that it is caught (rather than dies naturally).

2.2   Probability of tag return

For a tagged fish, we are interested not only in the probability of the fish being caught but also of its tag being returned. Even when a fish that has been tagged is caught, the tag may not be returned for several reasons:

1. The tag was lost almost immediately after the fish was tagged, because it was not securely fastened. This happens with probability $1 - \xi$, say.

2. The tag was not lost immediately, but through natural attrition tags are lost at a constant rate Ω per time period.

3. The tag is not lost, but when the fish is caught, the tag is not reported, either deliberately or through oversight. We assume that the probability of a tag being returned, conditional on the fish being caught and having a tag, is λ, and that it is the same whether the fish has one or two tags still attached.

In our model we suppose that all fish are double-tagged because this was the case for the 1990s SBT tagging experiments. We refer to the left-hand tag as tag A and the right-hand tag as tag B. In our development of the model, we allow for the shedding parameters ξ and Ω to differ for tag A and tag B, but we assume that they do not differ with region, age of fish, or time period.

Initially, consider just tag A. The probability that a tagged fish of age a, alive in region r at the beginning of time period t, survives until the end of the time period with tag A still intact, is

$$S^A_{atr} = e^{-M_a/2 - F_{atr} - \Omega_A} .$$

The probability that it is caught in the time period with its tag, is $(1 - S^A_{atr})f^A_{atr}$, where

$$f^A_{atr} = \frac{F_{atr}}{M_a/2 + F_{atr} + \Omega_A}$$

is the conditional probability that it is caught rather than dies naturally or sheds its tag. This is a competing risks model similar to in the previous section, except now tag shedding is an additional risk to dying naturally or being caught; the mathematical details are given in Annex A.

Analogous expressions, S^B_{atr} and f^B_{atr}, can be made regarding tag B.

Now consider both tags. Using the argument outlined in Annex A, the probability that a fish of age a, alive in region r at the beginning of time period t, survives until the end of the time period with both tag A and tag B, is

$$S^{AB}_{atr} = e^{-M_a/2 - F_{atr} - \Omega_A - \Omega_B} .$$

The probability that it is caught in the time period with both its tags, is $(1 - S^{AB}_{atr})f^{AB}_{atr}$, where

$$f^{AB}_{atr} = \frac{F_{atr}}{M_a/2 + F_{atr} + \Omega_A + \Omega_B}$$

is the conditional probability that it is caught rather than dies naturally or sheds both its tags.

Suppose a fish is double-tagged in region r_C at the beginning of period t_C at age a_C. Let $s_{t_C} = t_C - 2[(t_C - 1)/2]$. Then this fish will have age (in years) of $a = a(a_C, t_C, t) = a_C + [(t - t_C - 1 + s_{t_C})/2]$ in period $t \geq t_C$. We focus our argument on just tag A for the moment. Let R_t be the region that the fish is in at the start and during time interval t, $L_t = 1$ if the fish is alive at the start of time interval t and $A_t = 1$ if tag A is still on the fish. Hence

$$\Pr\{R_{t+1} = r, L_{t+1} = 1, A_{t+1} = 1 | a_C, t_C, r_C\} =$$
$$\sum_{r'=1}^{R} \Pr\{R_t = r', L_t = 1, A_t = 1 | a_C, t_C, r_C\} S_{atr'}^A \, \pi_{r'r|at} \tag{3}$$

where $\pi_{r'r|at}$ represents the probability of a fish of age a moving from region r' to r at the end of time period t. Looking at the right-hand side of this equation, we read that the fish was alive in region r' at time t and still in possession of tag A with probability $\Pr\{R_t = r', L_t = 1, A_t = 1 | a_C, t_C, r_C\}$, survived the time period without shedding its tag with probability $S_{atr'}^A$, and then moved from region r' to region r with transition probability $\pi_{r'r|at}$. Note that the transition probabilities do not depend on any of the fish's history prior to time t.

Now let us suppose S_{at}^A is the $R \times R$ diagonal matrix with rth diagonal element S_{atr}^A. Then, equation (3) can be expressed in matrix form as

$$p_{t+1|a_C t_C r_C}^A = \Pi'_{a(t)s(t)} \, S_{at}^A \, p_{t|a_C t_C r_C}^A \,,$$

where $p_{t|a_C t_C r_C}^A$ is a vector of length R whose rth entry is the probability that a fish, alive at age a_C in region r_C at the beginning of time period t_C with tag A attached, is alive in region r at the beginning of time period t with tag A still attached. This equation can be calculated iteratively, where we are now using as a starting vector that fact that the rth entry of $p_{t_C|a_C t_C r_C}^A$ is ξ_A for $r = r_C$ and 0 for $r \neq r_C$, where ξ_A is the probability that tag A is not shed immediately after tagging.

The probability that a fish, alive at age a_C in region r_C at the beginning of time period t_C with tag A attached, is caught with its tag during time period t in

region r is

$$q^A_{tr|a_Ct_Cr_C} = f^A_{atr}(1 - S^A_{atr})p^A_{tr|a_Ct_Cr_C} \tag{4}$$

where $p^A_{tr|a_Ct_Cr_C}$ is the rth entry of vector $p^A_{t|a_Ct_Cr_C}$. Note that tag B may or may not be on the fish.

Analogously we can compute the probability $q^B_{tr|a_Ct_Cr_C}$ of the fish being caught with tag B intact. Again, tag A may or may not be present.

Now, let $p^{AB}_{tr|a_Ct_Cr_C}$ be the probability that a fish, alive at age a_C in region r_C at the beginning of time period t_C with tags A and B attached, is alive in region r at the beginning of time period t with both tags still attached. In vector notation,

$$p^{AB}_{t+1|a_Ct_Cr_C} = \Pi'_{a(t)s(t)} S^{AB}_{at} p^{AB}_{t|a_Ct_Cr_C}.$$

To calculate this equation iteratively, we now use as a starting vector that fact that the rth entry of $p^A_{t_C|a_Ct_Cr_C}$ is $\xi_A\xi_B$ for $r = r_C$ and 0 for $r \neq r_C$, where $\xi_A\xi_B$ is the probability that neither tag A nor tag B is shed immediately. Then, the probability that the fish is caught in region r during time period t with both tags still intact is

$$q^{AB}_{tr|a_Ct_Cr_C} = f^{AB}_{atr}(1 - S^{AB}_{atr})p^{AB}_{tr|a_Ct_Cr_C}. \tag{5}$$

Then, the probability that the fish is caught with at least one of its tags is

$$q^A_{tr|a_Ct_Cr_C} + q^B_{tr|a_Ct_Cr_C} - q^{AB}_{tr|a_Ct_Cr_C},$$

and, taking into consideration non-reporting, the conditional probability that at least one of its tags is returned is

$$g_{tr|a_Ct_Cr_C} = \lambda_{atr}(q^A_{tr|a_Ct_Cr_C} + q^B_{tr|a_C,t_C,r_C} - q^{AB}_{tr|a_Ct_Cr_C}).$$

Note that we are allowing the reporting rate (λ_{atr}) to vary with age of the fish, time period and region.

The probability that neither tag is returned is

$$\bar{g}_{|a_Ct_Cr_C} = 1 - \sum_{r=1}^{R}\sum_{t=1}^{T} g_{tr|a_Ct_Cr_C}.$$

We therefore obtain a set of multinomial probabilities:

$$
\begin{aligned}
g \;=\; & (\bar{g}, g_{11}, g_{12}, \ldots, q_{1R}, \\
& g_{21}, g_{22}, \ldots, g_{2R}, \\
& \ldots, \\
& g_{T1}, g_{T2}, \ldots, g_{TR})
\end{aligned}
\tag{6}
$$

Many of these probabilities will be 0. For readability we have dropped the conditional expression $(|a_C t_C r_C)$ in the subscript of all terms.

2.3   Inference for the tagging data

We consider tag-recapture data for fish tagged in T consecutive time periods. Tagging occurs at the beginning of each odd time period (season 1) in region 1, and at the beginning of each even time period (season 2) in regions $2, \ldots, R$. In each time period and region of tagging, fish of ages, say, 1 to K (i.e., from K cohorts) are tagged. Tags are subsequently returned from the surface fishery operating in region 1 in season 1, and from the longline fisheries operating in regions $2, \ldots, R$ in season 2.

Inference for these data is fairly straightforward. We refer to the fish tagged at age a_C in region r_C at the beginning of period t_C as tagging group (a_C, t_C, r_C). If tagging occurs at K ages in R regions and T time periods, and T is even to correspond to $T/2$ years, then there will be $I = KRT/2$ tagging groups — $KT/2$ corresponding to fish tagged in region 1 in odd time periods and $KT/2$ corresponding to fish tagged in each of regions $2, \ldots, R$ in even time periods. In practice, tagging may not be possible at all ages in every region and time period so the number of tagging groups may be less.

Suppose we put the tag-recapture data into a matrix n of dimension $I \times J$, for which the data in row i are the numbers of returns by time period and region from tagging group i. The first entry of each row is the number of unrecaptured tags,

so that $J = 1 + RT'$, where T' is the number of time periods in which recaptures occur $(T' \geq T)$. The probabilities of return are placed into a corresponding matrix G. Row i of G is the vector g from (6) for the relevant tagging group. Each row of data is treated independently as multinomial data.

The log-likelihood is (apart from an additive constant)

$$l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \log G_{ij} \,.$$

If there are structural zeroes in the data, this method of computing the log-likelihood will fail, because $G_{ij} = 0$ for structural zeroes, and $\log G_{ij} = -\infty$ then. The likelihood for structural zeroes is $0^0 = 1$, and the log-likelihood is 0. We can achieve this result by the following simple device. Let

$$G_{ij}^* = \begin{cases} G_{ij} & \text{if } n_{ij} > 0; \\ 0.01 & \text{otherwise.} \end{cases}$$

Then

$$l_{tag} = \sum_{i=1}^{I} \sum_{j=1}^{J} n_{ij} \log G_{ij}^* \,.$$

2.4   Catch-at-age data

We now look at the model for catch-at-age data. First concentrate on a single cohort. A complication with catch-at-age data is that, since the fish do not have tags, we do not know their origins. However, we do know their ages, hence we can work out the age a_0 and time period t_0 at which the cohort first entered the study (i.e., was first tagged). For example, consider a fish caught in time period 5. If it was age 3, then it would be from the cohort tagged at age 1 at the beginning of time period 1 $(a_0 = 1, t_0 = 1)$; if it was age 2, then it would be from the cohort tagged at age 1 in time period 3 $(a_0 = 1, t_0 = 3)$; and if it was age 4, then it would be from the cohort tagged at age 2 in time period 1 $(a_0 = 2, t_0 = 1)$. Note that a_0 and t_0 together uniquely define the cohort. Suppose the population size of these fish at the beginning of t_0 in region r_0 is $P_{a_0 t_0 r_0}$. Then the probability of a fish originating in

region r_0 being recaptured in region r at time period t is $q_{tr|a_0 t_0 r_0}$ from equation (2). Thus, the probability of being recaptured in region r during time period t regardless of its source is

$$\frac{\sum_{r_0=1}^{R} q_{tr|a_0 t_0 r_0} P_{a_0 t_0 r_0}}{\sum_{r_0=1}^{R} P_{a_0 t_0 r_0}} .$$

The expected number of fish of this cohort caught in region r during time period t is thus

$$\mu_{tr|a_0 t_0} = \sum_{r_0=1}^{R} q_{tr|a_0 t_0 r_0} P_{a_0 t_0 r_0} .$$

For the catch-at-age data, we adopt a Gaussian model with known coefficient of variation c_r different for each region. The catches in each region and time are assumed to be statistically independent.

If the catch in region r and time period t of fish from cohort (a_0, t_0) is $C_{tr|a_0 t_0}$, then the log-likelihood for the catch-at-age data for all regions, time periods and cohorts involved in the study is (apart from an additive constant)

$$l_{catch} = \sum_{t=1}^{T'} \sum_{r=1}^{R} \sum_{a_0 t_0} \left[-\log \mu_{tr|a_0 t_0} - \frac{1}{2} \frac{(C_{tr|a_0 t_0} - \mu_{tr|a_0 t_0})^2}{c_r^2 \mu_{tr|a_0 t_0}^2} \right] .$$

The Gaussian model is intended to capture all source of variation in the catch-at-age data, including multinomial process error and sampling variability. The latter is assumed to be dominant, because for SBT the data have been derived from taking a small sample of the catch to be aged (where age is either estimated from length or from annuli in the otolith), then scaling up the sample age distribution to the total catch in numbers (which may itself have been estimated from total weight). A more complete discussion of the error distribution for the catch-at-age data can be found in Appendix 5.

2.4.1 An alternative

The philosophy behind the above argument that catches in each region and time period can be assumed to be statistically independent is as follows: within each cohort, the catches for each year can be modelled, unconditionally, as independent counts.

The cohorts are independent, so the catches within each year are independent. It is possible to pursue a different argument that may reflect actual practice.

We focus on a particular time period t and region r. We suppose that M fish are caught, with a total weight of W. We assume that W is measured, but M is not counted. If W_i is the weight of the ith fish, we have

$$W = \sum_{i=1}^{M} W_i \,.$$

If M is Poisson and the W_i are independent and identically distributed with mean μ_w and variance σ_w^2, we have

$$
\begin{aligned}
\mathrm{E}[W] &= \mu_w \, \mathrm{E}[M] \\
\mathrm{Var}[W] &= (\mu_w^2 + \sigma_w^2) \mathrm{E}[M] \,.
\end{aligned}
$$

For each cohort of fish, we know the probability of its being caught in time period t and region r. Hence

$$\mathrm{E}[M] = \sum_{a_0 t_0} \sum_{r_0=1}^{R} q_{tr|a_0 t_0 r_0} P_{a_0 t_0 r_0} = \sum_{a_0 t_0} \mu_{tr|a_0 t_0} \,.$$

A subsample of size $N = n$ is selected, and each fish in this sample is weighed and aged. Denote these by W_i and A_i. Without loss of generality, we label these as the first n of the M fish that are caught. We can equally regard the age group A_i as being equivalent to the cohort (a_0, t_0), so we instead interpret A_i as the cohort. We assume, for simplicity, that there is no measurement error. For the likelihood, we need

$$\Pr\{W_i = w, A_i = (a_0, t_0)\} = \Pr\{W_i = w | A_i = (a_0, t_0)\} \Pr\{A_i = (a_0, t_0)\}$$

For simplicity, we assume that $W_i | A_i$ is normal, so that

$$\Pr\{W_i = w | A_i = (a_0, t_0)\} = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_{a_0,t_0}} \exp\left(-\frac{(w - \mu_{a_0,t_0})^2}{2\sigma_{a_0,t_0}^2}\right) \,.$$

Of course, other assumptions, such as $W_i | A_i$ lognormal, could be more realistic. It is trivial to adjust the argument to this case. We also have

$$\Pr\{A_i = (a_0, t_0)\} = \frac{q_{tr|a_0 t_0}}{\sum_{a_0 t_0} q_{tr|a_0 t_0}} \,.$$

A16 - 14

It is then an easy matter to compute the log-likelihood of the sample.

The likelihood should also include that for W. Note that this includes the subsampled n, but we assume that M is so much larger than n that W can be assumed independent of the subsample. If this is not the case, the n can be omitted from W. We assume that the catch size is sufficiently large for W to have a Gaussian distribution. Hence

$$W \sim N(\mu_w \, \mathrm{E}[M], \mu_w^2 \mathrm{Var}[M] + \sigma_w^2 \mathrm{E}[M]) \,.$$

Thus we can include the Gaussian log-likelihood for W. In practice, we would probably assume that W had a known coefficient of variation.

The trouble with this style of argument is that it is highly idealised. In practice, the subsampling of fish for ageing and weighing is likely to be more complex than simple random sampling, and the subsample will, almost certainly, exhibit overdispersion. The likelihood of the catch-at-age data will therefore be quite complex. It is more pragmatic to adopt the approximate independence argument that we have used in Section 2.4.

2.5   Overall log-likelihood

The parameters are estimated by maximising the total log-likelihood

$$l = l_{tag} + l_{catch} \,.$$

We are implicitly assuming that both types of data are statistically independent, as seems reasonable.

In maximising the likelihood, we treat the reporting rate parameters (λ_{atr}) and tag shedding parameters (Ω_A and Ω_B) as known. The parameters to be estimated are the age-specific natural mortality parameters, the age-, time- and region-specific fishing mortality parameters, the region-specific population sizes for each cohort at the time of first tagging, the age- and season-specific transition probabilities between regions, and the age-, time- and region-specific reporting rates. Note,

however, that when a cohort is tagged at K consecutive ages, only $K-1$ natural mortality rate parameters can be estimated (this is shown in Appendix 13, and is a well-known feature in non-spatial Brownie models).

2.6   Allowing for non-mixing of recently tagged fish

For a species with a geographical range as widespread as SBT, we do not expect newly tagged fish to be fully mixed with the untagged population immediately after tagging. Moreover, tagging of SBT often occurs during the fishing season, rather than at the start, so tagged fish are only vulnerable to fishing pressure for part of the season. As such, we follow the approach used in Appendix 15 to modify the tag return probabilities to allow for fishing mortality to differ between tagged fish in the time period of tagging and untagged fish in that same time period. We assume that tagged and untagged fish are fully mixed by the time period following release. In particular, for fish in tagging group (a_C, t_C, r_C), we replace $F_{a_C t_C r_C}$ with the parameter $F^*_{a_C t_C r_C}$ in the likelihood for the tagging data.

Exploration of this model using simulations yielded an important finding. When the model was applied to exact data (i.e., simulated without any error), almost all of the parameters were estimated correctly, but a few were not. In particular, the F_{atr} values in time period 1 and the population sizes were incorrect. This led us to investigate the issue of parameter identifiability in this model.

2.7   Identifiability in the model allowing for a period of non-mixing

In time period 1 the tagging data involve F^*_{a11} rather than F_{a11}. The only data that bear on F_{a11} are the catch-at-age data. Furthermore, the population parameters are involved only in the catch-at-age data. Accordingly, we explain the identifiability problem using an artificial example in which only catch-at-age data are available. We simplify the problem to its bare essentials.

Consider a two-region spatial model in which the initial population in region 1 is P_1 and in region 2 is P_2. We shall suppose that fish are caught in region 1 in time periods 1, 3, 5, ..., and in region 2 in periods 2, 4, 6, ... The natural mortality per time period is M, and the fishing mortality in the relevant region in time period t is F_t.

All fish migrate from region 1 to 2 at the end of the odd time periods, and a proportion π_t migrate from region 2 to 1 at the end of the even time periods. The fish that started out in region 2 in time period 1 either die naturally or remain in region 2 for period 2. The proportion of fish $1 - \pi_t$ that remain in region 2 at the end of period t (where t is even) stay in region 2 for the next two time periods.

We assume that we have catch-at-age data only, and that π_t, M and F_2, F_3, F_4, ... are all known exactly. The question is: can we estimate P_1, P_2 and F_1 from the catch-at-age data?

In the following table, we set out the expected catches for the first four time periods. For simplicity of notation, we set $Q = P_1\, e^{-(M+F_1)} + P_2\, e^{-M}$.

Time period t Expected catch C_t

1	$P_1\left[1 - e^{-(M+F_1)}\right]\frac{F_1}{M+F_1}$
2	$Q\left[1 - e^{-(M+F_2)}\right]\frac{F_2}{M+F_2}$
3	$Q\,e^{-(M+F_2)}\pi_2\left[1 - e^{-(M+F_3)}\right]\frac{F_3}{M+F_3}$
4	$Q\,e^{-(M+F_2)}\left[\pi_2 e^{-(M+F_3)} + (1 - \pi_2)e^{-M}\right]\left[1 - e^{-(M+F_4)}\right]\frac{F_4}{M+F_4}$

It is clear that the catches C_t in time periods $t \geq 2$ consist of Q multiplied by known factors, and they all merely reinforce our knowledge of Q. We are therefore left with two equations for three unknowns:

$$
\begin{aligned}
C_1 &= P_1\left[1 - e^{-(M+F_1)}\right]\frac{F_1}{M + F_1} \\
Q &= P_1\, e^{-(M+F_1)} + P_2\, e^{-M}.
\end{aligned}
$$

Hence P_1, P_2 and F_1 cannot be identified from the catch-at-age data.

Now suppose we have three regions, but there is no fishing in regions 2 and 3 in the odd periods, or in region 1 in the even periods. Let F_{2r} be the fishing mortality

in regions 2 and 3 in period 2, which we assume to be known. The expected catch in period 1 is

$$C_1 = P_1[1 - e^{-(M+F_1)}] \frac{F_1}{M + F_1} .$$

Let $Q_2 = P_1 e^{-(M+F_1)}\pi_{12} + P_2 e^{-M}$, where π_{12}, the probability of a fish going from region 1 to 2 at the end of time period 1, is known. The expected catch in period 2 in region 2 is

$$Q_2 [1 - e^{-(M+F_{22})}] \frac{F_{22}}{M + F_{22}} .$$

Let $Q_3 = P_1 e^{-(M+F_1)}\pi_{13} + P_3 e^{-M}$, where $\pi_{13} = 1 - \pi_{12}$. The expected catch in period 2 in region 3 is

$$Q_3 [1 - e^{-(M+F_{23})}] \frac{F_{23}}{M + F_{23}} .$$

The last two equations are equivalent to knowing Q_2 and Q_3, since the right-hand factors are known. We clearly cannot identify P_1, P_2, P_3 and F_1 from C_1, Q_2 and Q_3.

There are several ways to overcome this problem. Here we shall assume

$$F_{atr}^* \quad = \quad \kappa F_{atr} \tag{7}$$

for all relevant times, regions and age groups, where κ is an unknown parameter to be estimated. For the 1990s SBT experiments, tagging often took place near the end of the fishing season so that the F^* values are likely to be quite small anyway, and model (7) is a parsimonious solution to the estimation problem.

3 Analysis of the 1990s southern bluefin tuna data

3.1 The data

The southern bluefin tuna data used as input to the model are: tag release and return data from tagging experiments conducted from 1991 to 1997; catch-at-age data from the commercial fisheries from 1991 to 1997; estimates of tag reporting rates for years 1991 to 1997; estimates of tag shedding rates from the 1990s tagging

experiments. Note that a year, 1991 say, is comprised of 2 seasons, where season 1 is defined as 1 November 1990 to 30 April 1991 and season 2 is defined as 1 May 1991 to 31 October 1991. The tagging, catch and reporting rate data are divided into 4 regions: region 1 is the Great Australian Bight (GAB), where fishing only occurs during season 1; regions 2, 3 and 4 represent South African waters (SAfr), the South-East Indian Ocean (SEIO) and the Tasman Sea (TAS), respectively, and fishing only occurs in these regions in season 2 (see Appendix 4 for specific definitions of the 4 regions).

Details about the data collection processes and data screening criteria for the tagging and catch data sets can be found in Appendix 4. This appendix also contains a section specific to the spatial analysis detailing how these data were compiled (e.g., how they were divided into seasons and regions). The tag-return and catch-at-age data from 1991 to 1997 are tabulated in Annex B, Tables 1 and 2. With regard to the tag-return data, release data are included for fish tagged at ages 1 to 3; a small number of fish were also tagged at ages 4 and 5 (< 300 in total) but they are not included in the table or in our analysis because the number for any cohort was too small to provide meaningful estimates. Recapture data are only included for years 1991 to 1997 because we do not have any data with which to estimate reporting rates beyond 1997. With regard to the catch-at-age data, Table 1 includes entries for years 1991 to 1997 and ages 1 to 8, but only those corresponding to a year and age of recapture are included in the model. For example, in 1991, recaptures only exist for fish at ages 1, 2 and 3 so only the catch numbers for these ages are used.

An analysis of tag shedding rates for SBT using data from the 1990s tagging experiments (namely the number of recaptures with one tag versus both tags still attached) is presented in Appendix 14. Allowing for the shedding parameters to differ between the two tags made little difference to the results, so it was assumed that $\xi_A = \xi_B = \xi$ and $\Omega_A = \Omega_B = \Omega$. The tag shedding estimates used as input to the spatial model are $\xi = 0.98$ and $\Omega = 0.065/2$. These are taken from the final row of Table 3 in Appendix 14, except noting that the Ω value is being divided by

2 because the analysis in Appendix 14 used time periods of years so that Ω was the shedding rate per year, whereas here we want the shedding rate per half-year. The estimates of ξ and Ω were obtained using the data from all taggers combined and, therefore, represent 'average' shedding parameters for all taggers. Recall that they are assumed to be known exactly.

The tag reporting rates used our analysis are presented in Table 3 of Annex B. These estimates are based on the non-spatial reporting rate estimates derived in Appendix 19. To obtain regional estimates, we take a weighted average of the fishery-specific reporting rate estimates from Table 4a of Appendix 19 (for fisheries relevant to a given region), using the catch-at-age data by fishery as weights. Details are given in Annex B. Recall that, like the shedding rate estimates, the reporting rate estimates are also treated as exact in the model.

3.2   Parameter identifiability

In our development of the model we assumed that fish are tagged in every region and time period in which fishing occurs (i.e., in region 1 in all odd time periods and in regions 2 to 4, for SBT, in all odd time periods). Looking at the tagging data for SBT, many of the relevant cells have zero tag releases. Therefore, before analysing the real data, it was necessary to check whether the parameters were identifiable. Accordingly we simulated some exact tagging and catch-at-age data from the model that mimicked the general nature of the SBT data. We adopted the parsimonious model (7) for the fishing mortalities in the time period after tagging, both in the simulation and in the estimation. The maximum likelihood estimates agreed exactly with the true values for all parameters, including κ. This suggests that, despite the patchy nature of the tagging, all parameters are identifiable.

3.3   Results

We applied the model to the SBT data corresponding to cohorts 1990 through 1994. These five cohorts were tagged at ages 1 to 3 and have sufficient release and recapture

data to warrant inclusion. We only included recaptures up to a maximum of age 5, because the numbers of recaptures beyond age 5 are relatively small (due to a combination of natural mortality, fishing mortality, tag shedding, and low reporting rates) and because the assumption that natural mortality is constant for ages 2 and above is less likely to hold true as more older age classes are included.

Analysis of the data initially led to highly erratic estimates that defied sensible interpretation. It was clear that although the parameters were identifiable, they were not very well constrained by the data. Accordingly, we placed some restrictions on the parameters in the model in an attempt to achieve more sensible estimates. Several sets of restrictions were tried, especially with regard to the transition probabilities and the distribution of age 1 fish by region. Of these, the following set seemed most satisfactory in terms of parsimony (as judged by Akaike's information criteria) and providing plausible parameter estimates.

For fishing mortalities, we assumed that these could be decomposed into factors:

$$F_{atr} = F_{ar} F_{tr} .$$

That is, we are assuming age-based selectivity is constant within a region but that fishing intensity varies with region. This seems a plausible model from a fishing viewpoint, and is consistent with the underlying assumptions in the current SBT stock assessments. Note that region comes into both factors. Clearly this factorisation is not unique, in that multiplying F_{ar} by a constant k and dividing F_{tr} by k leads to the same product. To overcome this, we fixed $F_{ar} = 1.0$ for $a = 3$.

With respect to the transition probabilities, $\pi_{rr'|as}$ is the probability of moving from region r to r' for a fish of age a at the end of season s. At the end of season 1 (i.e., odd time periods), recall that we are assuming all fish leave region 1, such that $\pi_{11|a1} = 0$. At the end of season 2 (i.e., even time periods), we are assuming that the fish either stay in the region they are in, or they move back to region 1. Moreover, we assume

$$\pi_{r1|a2} = \alpha_a \beta_r .$$

In words, from each age group, the same relative proportion return to region 1 from each region, but the absolute proportion can change with age. We anticipate that α_a would decline with age (so that a smaller proportion of fish return to the GAB as they get older). The same multiplicative non-uniqueness applies to this factorisation as for the fishing mortalities. However, it is not absolutely clear *a priori* how to make the factorisation unique, so we let the parameters be free but bounded, anticipating that at least one β_r will be forced to a boundary, and thus constrained. Of course, $\pi_{rr|a2} = 1 - \alpha_a \beta_r$ and $\pi_{rr'|a2} = 0$ for $r' \neq 1$ or r.

Finally, all cohorts being modelled were tagged initially at age 1 in season 1, so we let $P_{1t_0 r_0}$ be the number of age 1 fish in region r_0 and time period t_0 from the cohort that is age 1 at the beginning of time period t_0, where t_0 is odd. We model this as

$$P_{1t_0 r_0} = P_{1t_0} \, q_{r_0}$$

where $\sum_{r_0}^{R} q_{r_0} = 1$. In other words, P_{1t_0} is the total age 1 population size over all regions at the beginning of time period t_0, and q_{r_0} is the proportion of age 1 fish allocated to region r_0. We assume that P_{1t_0} varies between years (odd time periods), but that q_{r_0} remains constant.

The coefficients of variation of the catch-at-age data were set at 0.3 for all regions. The parameter estimates are as follows:

```
natural mortalities:
M(a)
age 1   age 2
0.415   0.398


fishing mortality components:
F(t,r)
region season    1991  1992  1993  1994  1995  1996  1997
  1       1      3.50  0.43  0.18  0.23  0.49  1.26  3.66
```

```
2         2        20.00  0.66  0.24  0.18  0.19  0.53  1.54
3         2         0.05  0.02  0.01  0.02  0.07  0.12  0.10
4         2         0.23  1.80  1.02  0.15  0.27  0.08  0.19
```

```
F(a,r)
```

region	age 1	age 2	age 3	age 4	age 5
1	0.03	0.52	1.00	0.30	0.04
2	0.00	0.07	1.00	4.88	2.59
3	0.01	0.31	1.00	0.68	0.25
4	0.00	0.14	1.00	1.55	5.93

```
F* factor:
kappa 0.08
```

```
transition probabilities, season 1:
pi(1,r|a,1)
```

region	age 1	age 2	age 3	age 4	age 5
2	0.084	0.116	0.051	0.004	0.108
3	0.819	0.734	0.899	0.576	0.883
4	0.097	0.141	0.050	0.420	0.009

```
transition probability components, season 2:
pi(r,1|a,2)=alpha(a)*beta(r)
alpha(a)
```

age 1	age 2	age 3	age 4
0.099	0.352	0.524	0.605

```
beta(r)
```

region 2	region 3	region 4

```
   0.010      0.757      0.990
```

```
age 1 population sizes by year:
P(1,t0)
    1991       1992       1993       1994       1995
3668222   3706423   2115802   1483992    970461
```

```
proportion of the age 1 population in each region:
q(r0)
region 1  region 2  region 3  region 4
   0.472     0.000     0.483     0.044
```

```
average yearly fishing mortalities by year and age:
 year  age 1  age 2  age 3  age 4  age 5
 1991  0.041   ---    ---    ---    ---
 1992  0.005  0.037   ---    ---    ---
 1993  0.002  0.018  0.100   ---    ---
 1994  0.003  0.017  0.096  0.100   ---
 1995  0.006  0.039  0.194  0.185  0.107
 1996   ---   0.071  0.344  0.270  0.129
 1997   ---    ---   0.454  0.445  0.195
```

The natural mortality estimates are quite high, but are within the range of values assumed for age 1 and 2 fish in stock assessments for SBT. However, recall that, in our model, natural mortality has been constrained to be the same at ages 2 and above, so that the age 2 parameter value actually applies to fish of ages 2 to 5. Thus, to check that this constraint is not biasing the results, we re-ran the

analysis with natural mortality modelled as a linear function of age. The estimate at age 5 still came out close to 0.3, and the likelihood value and the other parameter estimates were barely changed.

The age component estimates of fishing mortality suggest that, at least for the period of the study, selectivity in the GAB (region 1) was dome-shaped, peaking at age 3 and declining to close to zero at ages 1 and 5. Similarly, selectivity in the South-East Indian Ocean (region 2) was also dome-shaped with a peak at age 3, which may reflect the fact that substantial numbers of juveniles are caught as by-catch in the Taiwanese albacore fishery in this region. However, it was lower at age 2 and higher at ages 4 and 5 than in the GAB. In the South African and Tasman Sea regions (regions 2 and 4), fishing mortality is greatest at ages 4 and 5.

The year component estimates suggest that fishing mortality in the GAB and South African fisheries was U-shaped over the years 1991 to 1997, being lowest in 1993 in the GAB and in 1994-1995 for South Africa. There is suggestion of this pattern for the SEIO as well, however the absolute values are much lower. For the Tasman Sea, the fishing mortality estimates were variable over the years of the study, but were highest in 1992 and 1993. Note that the estimates for 1991 in the GAB and, more so, South Africa are very high. This is the result of age 1 catches being higher in 1991 in these regions than in later years (refer to Table 1 in Annex B). In 1991, only data for age 1 fish are included in the model so the age 1 selectivity is estimated to be low based on later years and the 1991 year effect can be estimated as high as necessary to compensate. This would not have been possible if data from more ages in 1991 were being included. In order to include earlier data (and achieve a reasonable fit), the model would need to allow for selectivity to change over time. This was an issue in the non-spatial analysis as well, and we refer the reader to Appendix 15 for a more complete discussion.

The transition probability estimates suggest that the majority (75-90%) of fish move from the GAB to the SEIO at the end of season 1 at all ages, with the exception of age 4 where a large fraction (about 40%) also move to the Tasman

region. It is difficult to know if the latter finding is real or simply an artefact of limited data. The transition probability estimates back to the GAB at the end of season 2 are broken down into a region and age component. The region component suggests that almost no fish leave South Africa, whereas large fractions return from the SEIO and Tasman regions. The age component suggests that a greater proportion of fish returns to the GAB as age increases. This is the opposite trend to what we believe to be true for SBT, and is probably due in part to lack of recapture data in the longline regions and at older ages. It may also reflect problems with the age distribution of the catch data, because inadequate and biased catch sampling as well as discarding of small fish have been identified as probable issues in the SBT fishery during the time of the study.

Population abundance appears to have decreased from over 3.5 million age 1 fish in 1991 and 1992 to just under 1 million age 1 fish in 1995. The distribution of the age 1 population by region, as given by q_{r_0}, suggests that at the start of season 1, almost all age 1 fish are found in the GAB and the SEIO. However, it should be recognized that there is almost no information for the model to determine this initial distribution. It is of more interest to look at the distribution at the start of season 1 at older ages, which we can calculate iteratively using all the parameter estimates. We find the same outcome – that almost all fish are found in the GAB and SEIO – but that the split between the two regions varies with age. At age 1 fish are fairly evenly split between the two regions, but at age 2 the SEIO dominates, and the pattern reverts with age until fish are fairly even split again by age 5. The fact that the percent of the population in the GAB increases from age 2 to 5 is not surprising given that the transition probabilities back to the GAB at the end of season 2 were estimated to increase with age. However, we reiterate that this result contradicts our prior beliefs and is likely due to inadequate and/or inaccurate data. Concerns exist about potential overestimation of the proportion of small fish in the length/age estimates of the surface fishery catches. If this were the case, it may be the source of the increasing estimated rates of transition back to the GAB at older

ages and the relatively low proportion of age 1 fish in the GAB originally.

It is also of interest to compare the parameter estimates from the spatial model with those obtained from a non-spatial analysis of the same data. Appendix 15 presents results from a non-spatial analysis; however, the data and model used are not directly comparable (for example, data from an additional cohort were included; errors in the reporting rate and shedding rate estimates were incorporated; natural mortality was assumed to be linear with age; differences in shedding rates between taggers were accounted for). Thus, we modified the model in Appendix 15 to be more closely comparable to the spatial model and applied it the SBT data from cohorts 1990 to 1994. The results are as follows:

```
natural mortalities:
age 1   age 2
0.456   0.367
```

```
fishing mortalities by year and age:
 year  age 1  age 2  age 3  age 4  age 5
 1991  0.019   ---    ---    ---    ---
 1992  0.001  0.049   ---    ---    ---
 1993  0.001  0.019  0.095   ---    ---
 1994  0.001  0.020  0.098  0.099   ---
 1995  0.001  0.040  0.197  0.200  0.103
 1996   ---   0.054  0.270  0.275  0.142
 1997   ---    ---   0.399  0.405  0.209
```

```
age 1 population sizes by year:
    1991      1992      1993      1994      1995
2974970   3932870   1545480   1408060   1205560
```

The natural mortality rate estimates are quite similar from the two analyses, especially if we consider the large uncertainty generally associated with natural mortality estimates.

The age 1 population size estimates are also similar. Both the spatial and non-spatial analyses suggest a decrease of about 70% in the number of age 1 fish from 1992 to 1995. The 1991 estimates are somewhat inconsistent in that the estimate from the spatial analysis is almost the same as the 1992 estimate, whereas the estimate from non-spatial analysis is quite a bit lower; however, the difference is reasonably small in practical terms.

To compare the fishing mortality estimates, we calculated average yearly fishing mortalities for the spatial model as outlined in Appendix 11, section 3.5.1. These are presented above as the last item of the non-spatial results. The estimates are remarkably similar in the two analyses.

4   Concluding remarks

Overall, the results are promising and suggest that the spatial tagging and catch model has potential to provide useful estimates of mortality rates and abundance for SBT. The population-wide estimates are generally consistent with estimates from the 'equivalent' non-spatial model and from prior stock assessments. It is difficult to evaluate the regional estimates – some seem plausible given our limited understanding of the distribution and movement of juvenile SBT (such as the regional fishing mortality estimates), whereas others contradict our expectations (such as the increasing number of fish returning to the GAB with age, and the general absence of young fish in the South African region).

Of course, in interpreting the regional results it is important to bear in mind the limitations of the data. The tagging and recapture data for the longline regions, South Africa and the SEIO in particular, are very limited. Furthermore, the division of the catch data between South Africa and the SEIO is uncertain due to the lack of detailed catch information for these regions (especially for the Taiwanese fishery).

We found that the region-specific parameter estimates for the longline regions were quite sensitive to the model constraints chosen. There is potential for the results to be greatly improved if, in future, adequate tagging were to occur in the longline regions, and if the catch-at-age data were improved.

Both the spatial and non-spatial model results presented correspond to a particular set of reporting rate estimates. The difficulty in estimating reporting rates for SBT is well-known, and is a large source of uncertainty in any analyses that include SBT tagging data. In recent stock assessments, a range of reporting rate options have been considered (see Appendix 19). For the analyses in this appendix, we selected the option considered most plausible; however further analyses should be done using alternate reporting rate options to see how the parameter estimates are affected.

If the spatial model is to be pursued as a viable model for analysing SBT tagging and catch data, a number of modifications would need to be made. For example, significant differences in tag shedding rates were found to exist among taggers (see Appendix 14), and the model should be modified to allow for such differences. Moreover, we took reporting rate estimates and tag shedding estimates from previous analyses and used them in the model as if they were known without error. The uncertainty in these estimates needs to be accounted for. This is especially true for the reporting rates, for which we have already stressed that the uncertainty is high. It would be straightforward to take the same approach used in Appendix 15 and include a separate likelihood term for the reporting rate estimates and for the shedding estimates (keeping in mind that this approach accounts for error in the mean estimates; it does not account for the fact that the mean estimates may be biased). Due to the fact that the spatial model was already fairly complex and we were uncertain of how it would perform when applied to the SBT data, we chose to keep it as simple as possible to begin with and make additions/improvements only after we had investigated its performance and felt they were warranted.

Finally, we did not present standard error estimates for any of the param-

eters. These can be obtained using the standard method of inverting the Hessian matrix, and they were provided by the software package we used to fit the model. However, we chose not to include them because we know that they are too small. Before considering the precision of the parameter estimates, we need to make the model more realistically reflect the amount of uncertainty in the data. At minimum this means including error in the reporting rates and shedding rates, but preferably the model should also allow for overdispersion in the multinomial tag-return data, and the error disribution for the catch-at-age data should be more thoroughly investigated.

Given the limitations in the existing data, further analyses with the spatial model did not seem warranted. However, the results are promising with regard to the potential for applying the spatial modelling framework to data from future SBT tagging experiments provided these experiments are appropriately designed and implemented (e.g., with adequate spatial distribution of tag releases and adequate collection of data for the estimating reporting rates).

Annex A: Competing risks theory

We examine the basics of competing risks theory for the probability of a fish surviving or being caught. Let X be the time to the fish being caught, which, for simplicity, we assume to be exponentially distributed with parameter λ. Let Y be the time to the fish dying naturally, which we assume to be exponentially distributed with parameter μ. We assume that X and Y are statistically independent. We are interested in the probability that a fish is caught in the time interval (x_1, x_2). This is

$$
\begin{aligned}
\Pr\{x_1 < X < x_2, Y > X\} &= \int_{x_1}^{x_2} \lambda e^{-\lambda x} \Pr\{Y > x\} \, dx \\
&= \int_{x_1}^{x_2} \lambda e^{-\lambda x} e^{-\mu x} \, dx \\
&= \frac{\lambda}{\lambda + \mu} (e^{-(\lambda+\mu)x_1} - e^{-(\lambda+\mu)x_2}) \\
&= \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)x_1} [1 - e^{-(\lambda+\mu)(x_2-x_1)}] .
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\Pr\{x_1 < X < x_2, Y > X | X > x_1, Y > x_1\} \\
&= \frac{\lambda}{\lambda + \mu} [1 - e^{-(\lambda+\mu)(x_2-x_1)}] \\
&= \frac{\lambda(x_2 - x_1)}{\lambda(x_2 - x_1) + \mu(x_2 - x_1)} [1 - e^{-(\lambda+\mu)(x_2-x_1)}] \\
&= \frac{F}{F + M} [1 - e^{-F-M}] ,
\end{aligned}
$$

where $F = \lambda(x_2 - x_1)$ and $M = \mu(x_2 - x_1)$.

In fact, we use a slightly more general version of this argument, in which X and Y have piecewise exponential distributions. Suppose we partition time (or age) into J intervals with fixed cutpoints

$$
0 = x_0 < x_1 < x_2 < \ldots < x_J = \infty .
$$

Set $J_j = [x_{j-1}, x_j)$. The hazard function for X, defined quite generally as the ratio $\lambda(x) = f(x)/S(x)$ where $f(x)$ is the density function and $S(x)$ is the survival

function, is piecewise constant:

$$\lambda(x) = \lambda_j \text{ if } x \in J_j \,.$$

Under this model X is said to have a piecewise exponential distribution. The cumulative hazard is

$$\Lambda(x) = \int_0^x \lambda(u) \, du \,.$$

Hence the survival function is

$$S(x) = \exp(-\Lambda(x)) \,.$$

We assume that Y has a piecewise constant hazard function $\mu(x)$ on the same intervals J_j. It is straightforward to check that for piecewise exponentials

$$
\begin{aligned}
& \Pr\{x_{j-1} < X < x_j, Y > X | X > x_{j-1}, Y > x_{j-1}\} \\
= \; & \frac{\lambda_j}{\lambda_j + \mu_j}[1 - e^{-(\lambda_j + \mu_j)(x_j - x_{j-1})}] \\
= \; & \frac{\lambda_j(x_j - x_{j-1})}{\lambda_j(x_j - x_{j-1}) + \mu_j(x_j - x_{j-1})}[1 - e^{-(\lambda_j + \mu_j)(x_j - x_{j-1})}] \\
= \; & \frac{F_j}{F_j + M_j}[1 - e^{-F_j - M_j}] \,.
\end{aligned}
$$

This is the version we need. Note that F_j is the incremental cumulative hazard of being caught over the interval J_j. Similarly M_j is the incremental cumulative hazard of dying naturally over J_j.

For tagged fish, we are interested in the probability of the fish being caught with a tag still attached. This involves not only probabilities of survival and being caught, but also the probability of its tag(s) being shed prior to it being caught. We generalise the above competing risks argument to include tag shedding as one of the competing risks.

Let Z be the time to tag shedding. We assume that Z has the distribution

$$\Pr\{Z \le x\} = (1 - \xi) + \xi \, [1 - \exp(-\omega x)]$$

where $(1 - \xi)$ is the proportion of tags shed immediately after tagging and ω is the rate at which tags are shed over time. Then

$$
\begin{aligned}
\Pr\{x_1 < X < x_2, Y > X, Z > X\} &= \int_{x_1}^{x_2} \lambda e^{-\lambda x} \Pr\{Y > x, Z > x\}\, dx \\
&= \xi \frac{\lambda}{\lambda + \mu + \omega} e^{-(\lambda+\mu+\omega)x_1} [1 - e^{-(\lambda+\mu+\omega)(x_2-x_1)}].
\end{aligned}
$$

Now $\Pr\{X > x_1, Y > x_1, Z > x_1\} = \xi e^{-(\lambda+\mu+\omega)x_1}$. Hence

$$
\begin{aligned}
\Pr\{x_1 &< X < x_2, Y > X, Z > X | X > x_1, Y > x_1, Z > x_1\} \\
&= \frac{\lambda}{\lambda + \mu + \omega}[1 - e^{-(\lambda+\mu+\omega)(x_2-x_1)}] \\
&= \frac{F}{F + M + \Omega}[1 - e^{-(F+M+\Omega)}],
\end{aligned}
$$

where $F = \lambda(x_2 - x_1)$, $M = \mu(x_2 - x_1)$ and $\Omega = \omega(x_2 - x_1)$.

In words, the probability that the fish is caught *with a tag* during time period (x_1, x_2), given that it was alive at x_1 with a tag, is given by $\frac{F}{F+M+\Omega}[1 - e^{-(F+M+\Omega)}]$. To complete the picture, $\frac{M}{F+M+\Omega}[1 - e^{-(F+M+\Omega)}]$ is the conditional probability that the fish dies naturally *with a tag* during (x_1, x_2), and $\frac{\Omega}{F+M+\Omega}[1 - e^{-(F+M+\Omega)}]$ is the conditional probability that the fish sheds a tag during (x_1, x_2). In this last case, the fish might survive the time period, or it might die or be caught, but without the tag. The competing risks argument says simply that $1 - e^{-(F+M+\Omega)}$ is the conditional probability that (at least) one of the events will occur during the time period, and the ratios $\frac{F}{F+M+\Omega}$, $\frac{M}{F+M+\Omega}$ and $\frac{\Omega}{F+M+\Omega}$ give the further conditional probabilities governing which one occurs *first*.

Note that ξ disappears. It disappears from all the conditional arguments. Instead of starting with a probability of 1 that the fish is alive and has a tag at the beginning of the first time period, it has a probability of ξ.

If the fish is double tagged, we introduce two independent random variables Z_1 and Z_2 with parameters ω_1 and ω_2 respectively. Then

$$
\begin{aligned}
\Pr\{x_1 &< X < x_2, Y > X, Z_1 > X, Z_2 > X | X > x_1, Y > x_1, Z > x_1, Z_2 > x_1\} \\
&= \frac{F}{F + M + \Omega_1 + \Omega_2}[1 - e^{-(F+M+\Omega_1+\Omega_2)}],
\end{aligned}
$$

where $\Omega_i = \omega_i(x_2 - x_1)$. This represents the probability that the fish is caught during period (x_1, x_2) with two tags, given that it was alive at time x_1 with two tags.

These arguments can all be generalised to piecewise exponential distributions in which the intervals over which the hazard is constant are the same for each distribution. In general, though, we shall assume that F depends on age, time period and region, that M depends on age, but that ω_1 and ω_2 are constant and equal. Since the time periods are of equal lengths, this means that $\Omega_1 = \Omega_2 = \Omega$.

5 Annex B: Southern bluefin tuna data

Table 1 shows the catch-at-age data for the years 1991 to 1997 inclusive. The regions are the Great Australian Bight (GAB), South Africa (SAfr), the South-East Indian Ocean (SEIO) and the Tasman Sea (TS). The summer season is denoted S1, and the winter season S2. Each row gives the catch by age group for a region, year and season.

Table 1: Catch-at-age data for the SBT fishery

			Age							
Region	Year	Season	1	2	3	4	5	6	7	8
GAB	1991	S1	42235	48037	110688	2053	11	0	0	0
GAB	1992	S1	28744	10684	80658	11646	287	8	0	1
GAB	1993	S1	1705	19548	51543	8679	693	23	0	2
GAB	1994	S1	79	865	66111	22551	1358	131	18	0
GAB	1995	S1	342	18712	94839	18546	1903	145	1	0
GAB	1996	S1	0	19702	145388	43393	1447	29	0	0
GAB	1997	S1	1889	28797	174461	43553	9259	496	0	0
SAfr	1991	S2	273	3135	18118	22353	9440	3803	3715	3713
SAfr	1992	S2	12	681	15960	28919	14905	4361	2062	1898
SAfr	1993	S2	1	1833	14821	20607	23410	15304	7171	4369
SAfr	1994	S2	0	907	16588	12350	7174	7093	4638	2665
SAfr	1995	S2	0	611	6331	12347	10981	7647	7662	5166
SAfr	1996	S2	12	614	6097	11819	20484	9878	7125	6289
SAfr	1997	S2	10	599	11559	16317	17976	13854	9682	6699
SEIO	1991	S2	554	10458	21566	2479	1523	1352	1245	1066
SEIO	1992	S2	249	3718	20298	3195	1710	1005	974	1098
SEIO	1993	S2	65	2517	13869	9191	4606	3732	2185	1274
SEIO	1994	S2	151	7379	23577	13853	6059	5361	3729	2540
SEIO	1995	S2	63	10911	25055	17355	9116	7466	5684	4828
SEIO	1996	S2	404	5770	18572	14985	9682	8842	7342	6757
SEIO	1997	S2	64	2400	14067	7158	4602	4368	4649	4976
TS	1991	S2	66	14945	25872	11949	6965	3318	2424	1729
TS	1992	S2	203	18785	34616	23709	13346	8357	4217	2465
TS	1993	S2	257	14452	39840	27629	19805	12301	7802	5069
TS	1994	S2	6	1211	12146	23023	17008	11518	7629	4618
TS	1995	S2	1	511	6850	10906	14776	10656	6618	4341
TS	1996	S2	5	194	1654	5554	7329	6657	5550	3439
TS	1997	S2	1	210	3696	4993	10306	7612	6202	4722

We now tabulate the tag-recapture data. The majority of tag releases occurred in the GAB, but moderate releases also occurred in the South-East Indian Ocean (SEIO) and the Tasman Sea (TS). Table 2 gives the recapture numbers by age and region for each tagged group. 'NA' indicates that the entry is a structural 0 (i.e., when the recapture age is younger than the release age) or that the entry corresponds to a year beyond 1997 (the last year being included in the analysis).

Table 2: Tag-recapture data for the SBT fishery

Region	Release Year	Season	Age	Number Releases	Recaps	Region	1	2	3	4	5	6	7	8
GAB	1991	S1	1	3301	151	GAB	19	25	23	10	3	1	1	NA
						SAfr	0	0	4	4	1	0	1	NA
						SEIO	1	2	0	0	0	0	0	NA
						TS	0	14	19	9	9	3	2	NA
GAB	1991	S1	2	3127	372	GAB	NA	79	91	13	10	0	0	0
						SAfr	NA	1	2	7	3	2	1	0
						SEIO	NA	8	10	5	2	1	1	0
						TS	NA	15	44	32	19	16	5	5
GAB	1991	S1	3	810	99	GAB	NA	NA	50	1	0	1	0	0
						SAfr	NA	NA	2	1	3	0	1	0
						SEIO	NA	NA	0	1	0	0	0	1
						TS	NA	NA	11	5	13	6	0	3
GAB	1992	S1	1	2147	133	GAB	1	10	38	27	1	3	NA	NA
						SAfr	0	0	1	0	3	0	NA	NA
						SEIO	0	1	6	1	0	0	NA	NA
						TS	0	10	11	9	7	4	NA	NA
GAB	1992	S1	2	4646	397	GAB	NA	48	79	48	4	1	0	NA
						SAfr	NA	0	9	15	6	3	0	NA
						SEIO	NA	5	4	9	3	1	1	NA
						TS	NA	35	65	28	19	7	7	NA
GAB	1992	S1	3	1097	100	GAB	NA	NA	33	1	3	2	0	0
						SAfr	NA	NA	0	1	1	1	1	0
						SEIO	NA	NA	5	1	1	1	0	0
						TS	NA	NA	19	15	6	5	2	2
GAB	1993	S1	1	4898	390	GAB	2	29	178	74	36	NA	NA	NA
						SAfr	0	0	0	5	1	NA	NA	NA
						SEIO	0	9	6	5	5	NA	NA	NA
						TS	0	2	16	7	15	NA	NA	NA
GAB	1993	S1	2	2937	225	GAB	NA	31	40	40	8	1	NA	NA

						SAfr	NA	1	4	6	5	1	NA	NA
						SEIO	NA	0	6	3	2	2	NA	NA
						TS	NA	28	17	18	5	7	NA	NA
GAB	1993	S1	3	2777	202	GAB	NA	NA	37	44	17	2	0	NA
						SAfr	NA	NA	1	9	2	1	0	NA
						SEIO	NA	NA	1	5	2	2	3	NA
						TS	NA	NA	26	20	8	10	12	NA
GAB	1994	S1	1	9003	869	GAB	3	79	375	329	NA	NA	NA	NA
						SAfr	0	1	7	6	NA	NA	NA	NA
						SEIO	0	24	8	5	NA	NA	NA	NA
						TS	1	3	5	23	NA	NA	NA	NA
GAB	1994	S1	2	3158	323	GAB	NA	22	152	51	28	NA	NA	NA
						SAfr	NA	0	0	5	0	NA	NA	NA
						SEIO	NA	1	8	5	1	NA	NA	NA
						TS	NA	6	7	14	23	NA	NA	NA
GAB	1994	S1	3	3640	289	GAB	NA	NA	41	99	9	15	NA	NA
						SAfr	NA	NA	2	5	5	2	NA	NA
						SEIO	NA	NA	17	6	5	5	NA	NA
						TS	NA	NA	17	34	11	16	NA	NA
GAB	1995	S1	1	8585	706	GAB	0	48	581	NA	NA	NA	NA	NA
						SAfr	0	2	10	NA	NA	NA	NA	NA
						SEIO	0	35	15	NA	NA	NA	NA	NA
						TS	0	2	13	NA	NA	NA	NA	NA
GAB	1995	S1	2	5899	832	GAB	NA	65	356	309	NA	NA	NA	NA
						SAfr	NA	1	7	6	NA	NA	NA	NA
						SEIO	NA	13	17	14	NA	NA	NA	NA
						TS	NA	1	13	30	NA	NA	NA	NA
GAB	1995	S1	3	2629	229	GAB	NA	NA	35	79	53	NA	NA	NA
						SAfr	NA	NA	0	3	1	NA	NA	NA
						SEIO	NA	NA	4	11	5	NA	NA	NA
						TS	NA	NA	14	9	15	NA	NA	NA
GAB	1996	S1	1	82	3	GAB	0	3	NA	NA	NA	NA	NA	NA
						SAfr	0	0	NA	NA	NA	NA	NA	NA
						SEIO	0	0	NA	NA	NA	NA	NA	NA
						TS	0	0	NA	NA	NA	NA	NA	NA
GAB	1996	S1	2	2518	414	GAB	NA	56	310	NA	NA	NA	NA	NA
						SAfr	NA	0	4	NA	NA	NA	NA	NA
						SEIO	NA	17	11	NA	NA	NA	NA	NA
						TS	NA	2	14	NA	NA	NA	NA	NA
GAB	1996	S1	3	1511	315	GAB	NA	NA	103	178	NA	NA	NA	NA
						SAfr	NA	NA	2	1	NA	NA	NA	NA
						SEIO	NA	NA	7	8	NA	NA	NA	NA
						TS	NA	NA	2	14	NA	NA	NA	NA
GAB	1997	S1	1	884	1	GAB	0	NA	NA	NA	NA	NA	NA	NA

						SAfr	0	NA	NA	NA	NA	NA	NA	NA
						SEIO	1	NA	NA	NA	NA	NA	NA	NA
						TS	0	NA	NA	NA	NA	NA	NA	NA
GAB	1997	S1	2	592	15	GAB	NA	12	NA	NA	NA	NA	NA	NA
						SAfr	NA	0	NA	NA	NA	NA	NA	NA
						SEIO	NA	3	NA	NA	NA	NA	NA	NA
						TS	NA	0	NA	NA	NA	NA	NA	NA
GAB	1997	S1	3	526	91	GAB	NA	NA	90	NA	NA	NA	NA	NA
						SAfr	NA	NA	0	NA	NA	NA	NA	NA
						SEIO	NA	NA	1	NA	NA	NA	NA	NA
						TS	NA	NA	0	NA	NA	NA	NA	NA
SAfr	1995	S2	3	3	0	GAB	NA	NA	0	0	0	NA	NA	NA
						SA	NA	NA	0	0	0	NA	NA	NA
						SEIO	NA	NA	0	0	0	NA	NA	NA
						TS	NA	NA	0	0	0	NA	NA	NA
SEIO	1995	S2	1	9	0	GAB	0	0	0	NA	NA	NA	NA	NA
						SAfr	0	0	0	NA	NA	NA	NA	NA
						SEIO	0	0	0	NA	NA	NA	NA	NA
						TS	0	0	0	NA	NA	NA	NA	NA
SEIO	1995	S2	2	61	9	GAB	NA	0	6	3	NA	NA	NA	NA
						SAfr	NA	0	0	0	NA	NA	NA	NA
						SEIO	NA	0	0	0	NA	NA	NA	NA
						TS	NA	0	0	0	NA	NA	NA	NA
SEIO	1995	S2	3	32	1	GAB	NA	NA	0	1	0	NA	NA	NA
						SAfr	NA	NA	0	0	0	NA	NA	NA
						SEIO	NA	NA	0	0	0	NA	NA	NA
						TS	NA	NA	0	0	0	NA	NA	NA
TS	1991	S2	2	82	16	GAB	NA	0	3	0	0	0	0	0
						SAfr	NA	0	0	0	0	0	0	0
						SEIO	NA	0	0	0	0	0	0	0
						TS	NA	11	1	0	0	0	0	1
TS	1991	S2	3	1	0	GAB	NA	NA	0	0	0	0	0	0
						SAfr	NA	NA	0	0	0	0	0	0
						SEIO	NA	NA	0	0	0	0	0	0
						TS	NA	NA	0	0	0	0	0	0
TS	1992	S2	2	69	9	GAB	NA	0	3	0	0	0	0	NA
						SAfr	NA	0	0	0	0	0	0	NA
						SEIO	NA	0	0	0	0	0	0	NA
						TS	NA	0	1	2	3	0	0	NA
TS	1992	S2	3	13	2	GAB	NA	NA	0	0	0	0	0	0
						SAfr	NA	NA	0	0	0	0	0	0
						SEIO	NA	NA	0	0	0	0	0	0
						TS	NA	NA	0	1	1	0	0	0
TS	1993	S2	2	224	41	GAB	NA	0	15	5	0	0	NA	NA

						SAfr	NA	0	0	0	0	0	NA	NA
						SEIO	NA	0	1	0	0	0	NA	NA
						TS	NA	7	5	2	4	2	NA	NA
TS	1993	S2	3	132	23	GAB	NA	NA	0	2	1	0	0	NA
						SAfr	NA	NA	0	0	0	0	0	NA
						SEIO	NA	NA	0	0	0	0	0	NA
						TS	NA	NA	8	6	2	2	2	NA
TS	1994	S2	2	19	3	GAB	NA	0	2	1	0	NA	NA	NA
						SAfr	NA	0	0	0	0	NA	NA	NA
						SEIO	NA	0	0	0	0	NA	NA	NA
						TS	NA	0	0	0	0	NA	NA	NA
TS	1994	S2	3	97	14	GAB	NA	NA	0	7	0	0	NA	NA
						SAfr	NA	NA	0	0	0	0	NA	NA
						SEIO	NA	NA	0	0	0	0	NA	NA
						TS	NA	NA	0	3	1	3	NA	NA
TS	1995	S2	2	8	1	GAB	NA	0	0	1	NA	NA	NA	NA
						SAfr	NA	0	0	0	NA	NA	NA	NA
						SEIO	NA	0	0	0	NA	NA	NA	NA
						TS	NA	0	0	0	NA	NA	NA	NA
TS	1995	S2	3	64	7	GAB	NA	NA	0	3	0	NA	NA	NA
						SAfr	NA	NA	0	0	0	NA	NA	NA
						SEIO	NA	NA	0	0	0	NA	NA	NA
						TS	NA	NA	3	0	1	NA	NA	NA
TS	1996	S2	2	6	0	GAB	NA	0	0	NA	NA	NA	NA	NA
						SAfr	NA	0	0	NA	NA	NA	NA	NA
						SEIO	NA	0	0	NA	NA	NA	NA	NA
						TS	NA	0	0	NA	NA	NA	NA	NA
TS	1996	S2	3	5	0	GAB	NA	NA	0	0	NA	NA	NA	NA
						SAfr	NA	NA	0	0	NA	NA	NA	NA
						SEIO	NA	NA	0	0	NA	NA	NA	NA
						TS	NA	NA	0	0	NA	NA	NA	NA
TS	1997	S2	2	1	0	GAB	NA	0	NA	NA	NA	NA	NA	NA
						SAfr	NA	0	NA	NA	NA	NA	NA	NA
						SEIO	NA	0	NA	NA	NA	NA	NA	NA
						TS	NA	0	NA	NA	NA	NA	NA	NA
TS	1997	S2	3	27	0	GAB	NA	NA	0	NA	NA	NA	NA	NA
						SAfr	NA	NA	0	NA	NA	NA	NA	NA
						SEIO	NA	NA	0	NA	NA	NA	NA	NA
						TS	NA	NA	0	NA	NA	NA	NA	NA

Table 3 gives the reporting rate estimates by region, year, season and age, for seasons in which fishing occurs. These estimates were obtained by taking a weighted

average of the fishery-specific reporting rate estimates from Table 4a of Appendix 19 (for fisheries relevant to a given region), using the catch-at-age data by fishery as weights. In particular:

For the GAB, we use the reporting rate estimates based on tag seeding data and assume they hold for all components of the GAB catches (surface, farm and very limited longline), except for farm catches in 1996 where we need to account for mass deaths due to disease. Specifically, to estimate the reporting rate for year y and age a, we took an average of the reporting rate estimates for Fishery 2 (surface fishery), option b, and Fishery 3 (farm fishery), option b, from Table 4a of Appendix 19 for year y, weighting them by the non-farm and farm catches of age a fish in year y, respectively.

For the Tasman region (TAS), catches are comprised of Japanese longline catches from vessels with observers (Fishery 4 in Appendix 19 since virtually all observer catches occurred within the Australian EEZ), Japanese longline catches from vessels without observers (Fisheries 6 and 7 in Appendix 19), and New Zealand domestic catches (assumed to have a zero reporting rate). To estimate the reporting rate for year y and age a, we took an average of the reporting rate estimates for Fishery 4, Fishery 6 and Fishery 7, option a, from Table 4a of Appendix 19 for year y as well as an estimate of 0 for the NZ fishery, weighting them by the catches of age a fish in year y that occurred in the Tasman region in each of the respective four fisheries.

For the South-East Indian Ocean region (SEIO), catches are comprised of Japanese longline catches from vessels without observers (Fisheries 6 and 7 in Appendix 19) and other miscellaneous catches (for which the reporting rate is assumed to be 0%). There were some very limited catches from vessels with observers but we were unable to separate these data out, so we assume they have the same reporting rate as the unobserved catches. To estimate the reporting rate for year y and age a, we took a weighted average of the reporting rate estimates for Fishery 6 and Fishery 7, option a, from Table 4a of Appendix 19 for year y as well as an estimate of 0 for

the miscellaneous catches, where we weighted them by the catches of age a fish in year y that occurred in the SEIO region in each of the respective three fisheries.

Finally, for the South African region (SAfr), reporting rates for a given year and age are assumed to be 40% of the corresponding reporting rate for the SEIO region. This is consistent with option b for Fishery 8 of Appendix 19, where Fishery 8 corresponds to catches in the South African region. An explanation for this assumption can be found in the Polacheck et al. (1996) reference within Appendix 19.

Table 3: Reporting rates for the SBT fishery

			Age							
Region	Year	Season	1	2	3	4	5	6	7	8
GAB	1991	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
GAB	1992	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
GAB	1993	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
GAB	1994	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
GAB	1995	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
GAB	1996	S1	0.81	0.437	0.545	0.578	0.664	0.81	0.81	0.81
GAB	1997	S1	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
SAfr	1991	S2	0.119	0.107	0.019	0.14	0.141	0.14	0.139	0.14
SAfr	1992	S2	0.14	0.088	0.025	0.14	0.141	0.14	0.14	0.139
SAfr	1993	S2	0.273	0.168	0.079	0.08	0.132	0.108	0.122	0.151
SAfr	1994	S2	0.323	0.27	0.162	0.147	0.188	0.172	0.2	0.237
SAfr	1995	S2	0.087	0.072	0.038	0.047	0.072	0.064	0.074	0.09
SAfr	1996	S2	0.215	0.106	0.017	0.06	0.131	0.127	0.152	0.181
SAfr	1997	S2	0.309	0.207	0.035	0.091	0.2	0.186	0.215	0.246
SEIO	1991	S2	0.297	0.268	0.048	0.351	0.351	0.349	0.347	0.349
SEIO	1992	S2	0.35	0.221	0.063	0.351	0.351	0.349	0.349	0.348
SEIO	1993	S2	0.682	0.419	0.198	0.201	0.329	0.271	0.305	0.378
SEIO	1994	S2	0.808	0.676	0.405	0.367	0.47	0.429	0.5	0.594
SEIO	1995	S2	0.218	0.179	0.095	0.118	0.179	0.159	0.185	0.224
SEIO	1996	S2	0.537	0.264	0.042	0.15	0.329	0.318	0.379	0.453
SEIO	1997	S2	0.774	0.516	0.087	0.228	0.499	0.465	0.538	0.615
TS	1991	S2	0.589	0.505	0.472	0.435	0.425	0.416	0.411	0.41
TS	1992	S2	0.388	0.43	0.408	0.401	0.388	0.383	0.382	0.385
TS	1993	S2	0.724	0.768	0.732	0.717	0.712	0.71	0.711	0.713
TS	1994	S2	0.157	0.876	0.852	0.842	0.83	0.826	0.827	0.829
TS	1995	S2	0.369	0.416	0.393	0.365	0.331	0.308	0.319	0.325
TS	1996	S2	0.817	0.715	0.747	0.71	0.707	0.701	0.694	0.672
TS	1997	S2	1	1	0.985	0.994	0.994	0.986	0.98	0.95

Appendix 17:

An evaluation of abundance estimates from tagging programs when tag returns are only available from one component of a multi-component fishery: an example based on the 1990s southern bluefin tuna tagging program

Tom Polacheck, J. Paige Eveson and Geoff M. Laslett

FRDC Project 2002/015

Table of Contents

Abstract

The question of what information can be obtained from tagging experiments when tag return data with reliable reporting rates and catch at age data are available from only one component of a multi-component fishery is examined in the context of SBT fisheries. A Petersen-type mark and recapture estimator of abundance is used as the basis of our examination. The estimator is applied to tag return data from tagging experiments conducted on SBT in the 1990s to examine the type of performance that might be expected from the current CCSBT tagging program, given that reliable reporting rates are unlikely to be available for the longline component of the fishery. The results suggest that using only the SBT tag returns from the surface component of the entire juvenile fishery may still allow for information on juvenile abundances and/or trends, but this requires relatively consistent mixing patterns of tagged fish with the complete population of juvenile fish. It also requires that reliable estimates of reporting rates and of the age distribution of the surface catches are available, which emphasizes the need for developing appropriate statistical estimators for these quantities. The results presented for the 1990s SBT tagging program are reasonably consistent with the assumption of consistent and high levels of mixing. They also indicate no increase (and possibly a decrease) in the strength of cohorts at age 1 during the first half of the 1990s, and suggest a declining trend in abundance by ages 2 and 3 for the surviving members of these cohorts.

Introduction

Tagging experiments provide a potentially informative approach for reducing uncertainty in stock assessments by providing direct estimates of fishing and natural mortality rates and/or abundance (e.g. Polacheck et al. 1998, 2003; Hoenig et al. 1998; Pollock et al. 2002). This is particularly true for fisheries in which fishery-independent abundance surveys are infeasible and therefore commercial catch rate (CPUE) data must be depended on as the only measure of relative abundance (e.g. in pelagic longline and purse seine fisheries). However, reliable estimation of fishing and natural mortality rates and/or abundance from tagging data requires that estimates of reporting rates be available[1]. Estimation of reporting rates can be problematical for some fishery components in a multi-component fishery. In such cases, can tagging experiments still produce estimates that are useful for stock assessments? The answer to this question is likely to depend upon the nature of the fishery and the other data available from the fishery[2].

---

[1] Note if one is willing to assume reporting rates are constant, it is theoretically possible to simultaneously estimate reporting rates as well as fishing and natural mortality rates in a multi-year tagging program of the same cohort. However, the precision of the estimates are generally poor.

[2] For example, if tags are well mixed and there is good information on the catch by age for all fishery components, reporting rates for a missing component can be estimated based on the return rate of tagged fish in that component compared with the return rate of tagged fish in components with reliable reporting rates (e.g. Hearn et al. 2003).

Tagging programs were conducted on juvenile southern bluefin tuna (SBT) in the 1990s in order to provide estimates of fishing mortality and recruitment (e.g. Polacheck et al. 1998). The CCSBT is currently conducting an extensive tagging program with a similar objective (Anon. 2001a), for which tagging commenced in 2001/2002. SBT are harvested by fishing fleets from a number of countries; in particular, juvenile SBT are harvested by Australian purse seiners within the Great Australian Bight (GAB) and by longline fleets from Japan, Korea and Taiwan on the high seas. Observers are currently the only practical and feasible way to obtain estimates of reporting rates from pelagic longline vessels (Polacheck et al. 2004). In the previous 1990s SBT tagging experiments, observers on Japanese longline vessels (principally within Australia's EEZ) were used as a basis for estimating longline reporting rates (Polacheck et al. 1996, 1998). Even so, the estimates of reporting rates had a high degree of uncertainty associated with them (Polacheck et al. 1998). However, the Japanese fleet no longer operates with the Australian EEZ, so this source of observer data no longer exists. In the current situation, obtaining even minimal levels of observer coverage in high-seas longline fisheries has proven to be extremely difficult. Thus, although the CCSBT set a target of 10% observer coverage for all of its major fisheries in 2001 (Anon. 2001b), on the high seas only Japan had placed observers on its vessels in 2002 and the coverage was ~3.5% . No substantive increases were expected in 2003 for any of the fleets (Anon. 2003). The 2003 CCSBT Scientific Committee concluded that the current levels of observer coverage in the Japanese, Korean and Taiwanese longline fleets are not high enough to provide useful estimates of reporting rates from these fleets (Anon. 2003). Thus, reliable estimates of the reporting rates for the main longline fisheries will not be obtainable for at least the first several, if not all, years during which significant tag returns would be expected from the current CCSBT releases.

The current paper examines the question of what information can be obtained from tagging experiments when tag return data with reliable reporting rates and catch at age data are available from only one component of a multi-component fishery. The question is addressed in the context of SBT fisheries, and the approach taken is to consider the tagging and catch at age data in the context of a Petersen-type mark and recapture estimator of abundance (Seber 1973). A Petersen estimator is based on the ratio of the observed number of tags returned within samples taken from the population given the known number of tags released into the population. In a fishery context, the catch at age data constitutes a sample from the population. However, unlike most situations in which a Petersen-type estimator is used, the

size of the sample examined for tags is estimated rather than being known exactly. The approach developed is applied to data from tagging experiments conducted on SBT in the 1990s to examine the performance that might be expected from the current CCSBT tagging program, given that the reliable reporting rates are unlikely to be available for the longline component of the fishery.

Data and Background

Data from the multi-year, multi-age tagging experiments on juvenile SBT conducted in the 1990s (see Polacheck et al. 1998 and references therein) are used here. These experiments tagged fish in Western Australia (WA) and the Great Australian Bight (GAB)[3]. Fish were tagged in WA from 1991 through 1995 and in the GAB from 1991 through 1997. Fish of ages 0 to 2 were tagged in WA, and fish of ages 1 to 5 were tagged in the GAB. Only releases between ages 1 and 2 for WA and ages 1 and 3 for the GAB are considered here since only a small number of releases were outside these age ranges. The age of a fish when tagged was estimated based on its length using cohort slicing and the SBT growth curve currently being used by the CCSBT for its stock assessments. All tagging was done between November and April, so the ages were adjusted in order that fish tagged in November or December from a given year-class/cohort were placed in the same age grouping as those tagged after December. This adjusted age is referred to as a fish's "cohort" age[4], and it is the age used throughout this paper. Table 1 provides the number of tags released by area, cohort and age. The age of the fish tagged in each area reflects the predominant age classes found during tagging in each area.

In this paper, tag returns only from the GAB are considered. Juvenile SBT (ages 1 to 4) tend to spend their summers in coastal waters of Australia, where they are harvested by Australian surface fisheries, and their winters in deeper oceanic waters, where they are harvested by various longline fisheries. Age 0 to 2 fish are found in WA, while age 1 to 4 fish are commonly found in the GAB. The proportion of the global SBT stock for each of these age classes found in WA and the GAB during the summer months is not known; however, it is thought to be a relatively high but diminishing with age. Over the period covered by the

---

[3] Relatively small numbers of fish were tagged in other areas, primarily in waters off eastern Tasmania and from longline vessels. These releases have been excluded from the analyses presented here.
[4] SBT spawn between September and April. For the purpose of aging, all fish are assumed to have a birth date of January 1. Cohort age is defined as its estimated age from cohort slicing (i.e. its calendar age) if a fish was tagged or caught prior to June and as one plus its estimated age from cohort slicing if it was tagged or caught after the end of June.

tagging experiments, the Australian surface fishery shifted from predominately a pole and line fishery targeting fresh SBT for the Japanese sashimi market to a purse seine fishery for tuna farming. Nevertheless, for the fishing seasons of tag recoveries considered here (1991 through 1998) the estimated age composition of the surface catch in terms of cohort ages was relatively consistent, with some shift away from smaller/younger fish (Figure 1).

Methods

Basic estimator

The basic estimator used in this paper is:

$$P^*_{A,a,c,r} = \frac{C_{c,a} N_{A,c,r}}{R_{A,a,c,r}} \qquad (1)$$

where:

> A = age of tagging;
>
> a = age or ages of recapture (i.e. can be a vector of more than one age);
>
> c = cohort;
>
> r = the region of tagging (WA, GAB, or WA and GAB combined);
>
> $C_{c,a}$ = the catches from cohort c in the GAB at age a;
>
> $R_{A,a,c,r}$ = the number of recaptures in the GAB from cohort c at
>
> > age a that were released in region r at age A;
>
> $N_{A,c,r}$ = the number of releases from cohort c at age A in region r;
>
> $P^*_{A,a,c,r}$ = a measure of a cohort c's "strength" based on recaptures at
>
> > age a from releases at age A in region r.

As developed below, alternative interpretations of $P^*_{A,a,c,r}$ are possible depending upon assumptions made about mixing and the proportion of the juvenile stock in the GAB. Note $P^*_{A,a,c,r}$ is only calculated using returns from years after the year of release (i.e. for $a > A$) to allow for heterogeneity in recaptures during the year of tagging (e.g. short-term incomplete mixing during a season within a region; variability in the time of releases relative to the fishery; some releases having occurred in areas near the commercial fishery). In the results presented here tag reporting rates are assumed to be 100%.

Interpretation of $P^*_{A,a,c,r}$ assuming complete mixing

If there is full mixing of the tagged fish with the untagged fish from a cohort prior to there being any (substantial) differential fishing mortality, then $P^*_{A,a,c,r}$ provides an estimate of the size of a cohort at the age of tagging. This can easily be seen by deriving expressions for the expected catches and number of tag recaptures. Assuming that the fish tagged in a region are a representative sample of the fish in that region and that tagging does not affect their subsequent behaviour or mortality, the expected number of recaptures in the GAB of age a fish from cohort c that were released at age A in region r is:

$$R_{A,a,c,r} = \frac{f_{a,c}}{f_{a,c} + m_a} \rho_{A,a,c,r} N_{A,a,c,r} \left(1 - e^{-(f_{a,c}+m_a)}\right) e^{-z^+_{A,a-1,c,r}} \tag{2}$$

where

$f_{a,c}$ = the fishing mortality rate in the GAB for age a fish from cohort c;

m_a = the natural mortality rate for age a fish (assumed for simplicity to be constant across cohorts);

$\rho_{A,a,c,r}$ = the fraction of fish from cohort c that were in region r at age A that are in the GAB at age a (i.e. year $c+a$) during the fishing season;

$z^+_{A,a-1,c,r}$ = the cumulative natural and fishing mortality rates between ages A and $a-1$ (inclusive) for fish from cohort c that were in region r at age A

Similarly, the number of fish caught in the GAB at age a from cohort c (i.e. in year $c+a$) that were in region r at age A (i.e. the number of fish caught for which the tagged fish constitute a representative sample) is:

$$C_{A,a,c,r} = \frac{f_{a,c}}{f_{a,c} + m_a} \rho_{A,a,c,r} \varphi_{A,c,r} P_{A,c} \left(1 - e^{-(f_{a,c}+m_a)}\right) e^{-z^+_{A,a-1,c,r}} \tag{3}$$

where:

$\varphi_{A,c,r}$ = the fraction of cohort c that was in region r at age A;

$P_{A,c}$ = the size of cohort c at age A;

$C_{A.a,c,r}$ = the catch of age a fish in the GAB in year $a+c$ (from cohort c) that were in

region r at age A.

In a parallel manner, the expected catch in the GAB of age a fish from cohort c that were not in region r (referred to as r^*) at age A is simply:

$$C_{A.a,c,r^*} = \frac{f_{a,c}}{f_{a,c} + m_a} \rho_{A,a,c,r^*} \left(1 - \varphi_{A,c,r}\right) P_{A,c} \left(1 - e^{-(f_{a,c}+m_a)}\right) e^{-z^+_{A,a-1,c,r^*}} \qquad (4)$$

Adding equations 3 and 4 provides an expression for the total catch in the GAB from a cohort at age a (i.e, $C_{c,a}$ from equation 1). Under full mixing and before any differential fishing mortality, the ρ and z^+ parameters are equal across regions (i.e. $\rho_{A,a,c,r} = \rho_{A,a,c,r^*}$ and $z^+_{A,a-1,c,r} = z^+_{A,a-1,c,r^*}$), and by dividing the total catch expression (i.e. equation 3 + 4) by equation 2, it is straightforward to show that:

$$\frac{C_{c,a}}{P_{A,c}} = \frac{R_{A,a,c,r}}{N_{A,c,r}} \qquad (5)$$

In other words, $P^*_{A,a,c,r}$ provides an estimate of a cohort's size at the time/age of tagging ($P_{A,c}$). Note that $P_{A,c}$ can be calculated using returns from any age (or set of ages) after the age of tagging, allowing for multiple estimates of a cohort's size at the time of tagging.

If there was a short period of non-mixing (such that natural mortality could be ignored) in which fishing mortality was primarily in the area of releases and the recaptures were known for that period, $P^*_{A,a,c,r}$ could then be used to provide an estimate of a cohort's size at the time after mixing by reducing the number of releases by the number of short term recaptures. For example, in the SBT case, a small number of the GAB releases occurred during the fishing season (rather than at the end of the season) and in areas near the commercial fishery; taking into account the first year's recaptures in the surface fishery could address this problem. Because the numbers of first year recaptures were small, recaptures that occurred during the season of release have not been excluded in the results presented below, but a comparison of

the estimates when these recaptures were excluded showed that it had only a minimal effect on the estimates of $P^*_{A,a,c,r}$.

Interpretation of $P^*_{A,a,c,r}$ assuming incomplete mixing

If there is substantial non-mixing, then the relationship between $P^*_{A,a,c,r}$ and the size of the cohort ($P_{A,c}$) will depend upon: the fraction of the cohort was in the region of tagging at the time of tagging ($\varphi_{A,c,r}$); the relative fraction of the cohort that go into the GAB at age a that, at the time of tagging, were in the region of tagging compared to those that were not in the region of tagging (i.e. $\rho_{A,a,c,r}$ compared to ρ_{A,a,c,r^*}); and the differential in the fishing mortality rates that the two different groups of fish experience (i.e. $z^+_{A,a-1,c,r}$ compared to $z^+_{A,a-1,c,r^*}$). If all of these are highly variable over time, then estimates of $P^*_{A,a,c,r}$ relative to $P_{A,c}$ will be highly variable and will not provide any useful information about either absolute abundance or trends in abundance. However, there may be some situations in which the ratio $P^*_{A,a,c,r}/P_{A,c}$ will be relatively constant (i.e. a constant relative bias) and thus a time series of $P^*_{A,a,c,r}$ will provide a relative index of $P_{A,c}$.

Appendix 1 provides tables of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for the parameters $\varphi_{A,c,r}$, $\rho_{A,a,c,r}$, ρ_{A,a,c,r^*}, $z^+_{A,a-1,c,r}$ and $z^+_{A,a-1,c,r^*}$. These tables provide an indication of the degree of bias in $P^*_{A,a,c,r}$ that can occur. The bias in $P^*_{A,a,c,r}$ can be either positive or negative and its potential range is quite large (i.e. $P^*_{A,a,c,r}/P_{A,c}$ ranges from 0.31 to 7.76 for the range of the parameter values examined in Appendix 1). As such, without any additional information on the values for the parameters $\varphi_{A,c,r}$, $\rho_{A,a,c,r}$, ρ_{A,a,c,r^*}, $z^+_{A,a-1,c,r}$ and $z^+_{A,a-1,c,r^*}$, the absolute values of the estimates for $P^*_{A,a,c,r}$ would be of little value (even as possible bounds for $P_{A,c}$). Examination of the tables in Appendix 1 do suggest that there are some circumstances under which a time series of $P^*_{A,a,c,r}$ could provide a useful relative index of $P_{A,c}$. This would clearly apply if $\varphi_{A,c,r}$, $\rho_{A,a,c,r}$, ρ_{A,a,c,r^*}, $z^+_{A,a-1,c,r}$ and $z^+_{A,a-1,c,r^*}$ were constant over time. However, there appear to be some other situations in which the variability in the relative bias would be expected to remain relatively small. For example, if $\varphi_{A,c,r}$ and $\rho_{A,a,c,r}$ are

reasonably high and consistent (i.e. a large fraction of a cohort at the age of tagging is available for tagging and the proportion of these fish that go into to the area of recapture is high) combined with low rates of exploitation prior to the recovery for the tagged component of the population (conditions which are thought to apply to 1-year-old SBT), then $P^*_{A,a,c,r} / P_{A,c}$ would appear relatively insensitive to reasonable variation in ρ_{A,a,c,r^*} and $z^+_{A,a-1,c,r^*}$.

It is also worth noting that if $\rho_{A,a,c,r}$ and $z^+_{A,a-1,c,r}$ are relatively constant, then $P^*_{A,a,c,r}$ would provide a relative abundance index of the number of age a fish from a cohort in the area of recapture. In the case of SBT, given that fishing mortality rates on age 1 are near zero, a time series of estimates of $P^*_{A,a,c,r}$ from returns at age 2 would provide a good measure of relative abundance of age 2 fish in the GAB as long as $\rho_{A,a,c,r}$ was relatively constant (i.e. as long as in each year the same proportion of age 1 fish that were in the region of tagging go into the GAB at age 2). In this situation, the data provide a straightforward Petersen estimate except that the number of marked fish available for recapture is unknown but has been proportionally reduced by the same fraction in each year.

Variance and confidence intervals for $P^*_{A,a,c,r}$

There are two principle sources of errors in the estimates of $P^*_{A,a,c,r}$ (assuming 100% reporting rates). The first is the sampling error associated with the number of tags recaptured and the second is estimation error associated with the number of fish caught during a year. There is also potential error in the actual number of tags released by age from a cohort due to aging errors in the cohort sliced age estimates. This latter source of error is not considered here.

Tag return data are commonly modelled as multinomial. However, there is a number of factors that lead to recaptures being over-dispersed relative to a multinomial (e.g. heterogeneity in recapture probabilities as a result of schooling behaviour, or variability in selectivity among vessels). Previous analyses of these tagging data indicated over-dispersion in the data exists and a bootstrap approach was suggested as a mechanism to estimate the variance associated with sampling error in the number of recovered tags (Polacheck et al. 1998). Thus, a bootstrap approach was used here. Bootstrap samples were constructed by randomly selecting, with replacement, days from which tag releases occurred during a release

season and area. Days were selected until the bootstrap sample comprised the same number of days for which tagging actually occurred. All tags released on the selected days and their associated recoveries were included in the bootstrap sample. The bootstrap samples were thus conditioned on the number of days for which tagging occurred; the actual number of tags included in different bootstrap samples varied.

The appropriate approach to estimate the error associated with the catch at age estimates is less clear. The estimates of the catch at age are derived from a complex system of sampling the catch for length, converting the sampled length measurements to age estimates using cohort slicing, then scaling up the estimated sample age distribution using estimates of the total catch. Total catch in numbers is estimated from the landed weight of the catch divided by the estimated mean weight in the case of pole and line caught fish, and from video counts of fish during transfers from towing to farm cages in the case of the purse seine caught fish. For the purpose of getting an indication of how errors in the estimated catch at age data contribute to the error in the estimates of $P^*_{A,a,c,r}$, a Monte-Carlo approach was taken. In this approach, it was assumed that the major source of error in the catch at age estimates comes from estimating the age distribution; the total catch within a fishing season was assumed to be known exactly. Monte-Carlo re-sampling of the catch at age estimates was then performed assuming multinomial sampling with a pre-specified effective sample size. A sample size of 50 was used in the results presented here.

A single bootstrap sample was combined with one realization from the Monte-Carlo re-sampling of the catch at age data to produce a single "bootstrap/Monte-Carlo" estimate of $P^*_{A,a,c,r}$. This process was repeated 1000 times to derive an estimate of the coefficient of variation (CV) and confidence interval for $P^*_{A,a,c,r}$. Note that in some cases a bootstrap sample could contain a set of releases for which there were zero recaptures for the age range of recaptures being considered. In such cases, it was not possible to estimate $P^*_{A,a,c,r}$, and such a bootstrap sample was excluded in the calculation of the bootstrap mean, variance and confidence interval for that age range. This will underestimate the overall uncertainty. The bootstrap/Monte-Carlo estimates of the CVs and confidence intervals presented here are clearly conditional on the assumed effective sample size for the catch at age data, but they

provide both an indication of what these values might be and the relative contribution of the error associated with the tagging and catch at age data in the overall estimates of $P^*_{A,a,c,r}$.

Results

Figure 2 compares estimates of $P^*_{A,a,c,r}$ based on returns at age 2, 3, 4 and 2-4 pooled for fish released at age 1 in WA. With the exception of the estimate based on returns of age 2 fish for the 1994 cohort, the estimates based on the returns for different ages exhibit a large degree of consistency within a cohort (e.g. the 90% confidence error bars for the age-specific return estimates overlap with the pooled age estimates). Such agreement is consistent with full and complete mixing of the tagged fish with all the fish from a cohort. In the case of the 1994 cohort, the large difference stems from very few tagged fish having been returned at age 2 relative to the estimated number of age 2 fish caught, although the very wide error bars for this estimate indicate that there is large uncertainty associated with it. Figure 3 presents similar estimates as those in Figure 2 except the estimates are based on age 2 releases in WA. Figures 4 to 6 present similar results for releases in the GAB at age 1, 2 and 3 respectively. In general, the estimates based on returns at different ages tended to be relatively consistent within a cohort, although fewer actual comparisons could be made in some of these figures (e.g. Figure 3 is sparse because relatively few age 2 fish are tagged in WA; in Figure 6, which shows results for age 3 releases in the GAB, estimates are only possible using age 4 returns). The largest discrepancy in the age-specific estimates in these later figures is between the 1991 cohort estimates using age 3 versus age 4 returns from releases at age 2 in the GAB (Figure 5). In this case, it is not possible to distinguish which is the more likely "outlier". In this context, the estimate for the 1991 cohort using age 3 releases (based on age 4 returns) in Figure 6 also appears to be low based on the time trend in the estimates. It is perhaps worth considering that the 1991 cohort estimates based on age 4 returns (from age 2 and 3 GAB releases) and the 1994 cohort estimates based on age 2 returns (from age 1 WA releases) are dependent upon the estimates of the catch at age for the 1994/1995 and 1995/1996 fishing seasons. Sampling protocols to deal with the increasing farm catches were developing during this period. As such, the age compositional data from this period may be less reliable and warrant further examination.

Figure 7 compares the estimates of $P^*_{A,a,c,r}$ using age 1 releases from WA, the GAB, and both areas combined. Figure 8 shows similar results but for age 2 releases. In both cases, only

the estimates based on the pooled tag return data up through age 4 are shown. For both age 1 and 2 releases, the estimates of $P^*_{A,a,c,r}$ using WA and GAB releases appear reasonably consistent. This could be considered further indication of a high degree of mixing of tagged fish with all the fish from a cohort. The estimates of $P^*_{A,a,c,r}$ for the combined WA and GAB release data tend to follow closely the estimates for either the WA estimates (in the case of age 1 releases) or the GAB estimates (in the case of age 2 releases). This reflects the fact that releases from WA tend to dominate the overall age 1 releases, while GAB releases tend to dominate the age 2 releases for each cohort (Table 1). This simply reflects the areas where these age-classes of fish are generally found.

A further check of the consistency of the $P^*_{A,a,c,r}$ estimates in terms of mixing and as a possible measure of a cohort's absolute abundance at the age of tagging is to compare the estimates for each cohort across release ages. If the estimates represent estimates of absolute abundance, then the estimates at each successive release age should decrease, reflecting the natural and fishing mortality that occurred on each cohort. Figure 9 compares the estimates of abundance for ages 1 and 2 based on the combined releases from WA and the GAB and returns pooled across all recapture ages through age 4. Figure 10 provides a similar comparison but for age 1, 2 and 3 estimates based on releases from the GAB only. In the case of the releases from the GAB (Figure 10), the estimates for a cohort always decrease with age and the magnitude would appear reasonable. However, this would best be evaluated within an overall stock assessment that considered all catch from a cohort. In the case of the estimates from the combined releases (Figure 9), the estimates of age 1 abundance are generally greater than the age 2 estimates. However, in three of the six comparisons[5], the differences appear to be relatively small if natural mortality rates are considered. For age 1 SBT, natural mortality rates have been estimated to be relatively high (~0.3-0.5) based on alternate analyses of these same tagging data (Polacheck et al. 1998). As noted above, the estimates for the combined age 1 releases are dominated by releases from WA (with the exception of the 1995 and 1996 cohorts), while the age 2 releases are dominated by releases from the GAB. This would suggest that, to the extent that these combined estimates for ages 1 and 2 are considered inconsistent with an assumption of complete mixing, a greater fraction of the stock of age 1 fish represented by the tagged fish in WA ends up in the GAB than the fraction of the stock

---

[5] Note that the 1995 age 1 estimate is based on only 82 releases from the GAB (as there were no WA releases in this year) and 15 returned tags (see Tables 1 and 2), and perhaps should be excluded in this comparison.

of age 2 fish represented by the tagged fish in the GAB (i.e. too many age 1 WA releases were recaptured relative to age 2 GAB releases).

Discussion

The estimates of $P^*_{A,a,c,r}$ presented here suggest that using only the SBT tag returns from the surface component of the entire juvenile fishery may still allow for information on juvenile abundances and/or trends. This requires relatively consistent mixing patterns of tag fished with the complete population of juvenile fish. From the results presented for the 1990s SBT tagging program, the estimates of $P^*_{A,a,c,r}$ for a cohort at a particular age and based on a particular area of release (e.g. WA or GAB) are relatively consistent using returns at different ages, suggesting relatively consistent, if not complete mixing. Comparisons of the estimates for different ages of release also suggest a reasonable degree of consistency in mixing for releases at age 1, 2 and 3 from the GAB, but some concerns about lack of complete mixing within the juvenile population when estimates based on combined releases from WA and GAB are considered. The results suggest that possibly too high of a proportion of age 1 releases (particularly from WA) are subsequently recaptured in the GAB relative to age 2 releases. To the extent that the estimates presented here from the 1990s tagging experiments are considered to provide information on juvenile abundances, the results provide no indication of an increase (and possibly a decrease) in the strength for cohorts at age 1 from the first half of the 1990s, and suggest a declining trend by age 2 and 3 in abundance for the surviving members from these cohorts.

It should be emphasised that the estimates using only tag return data from the Australian fishery as presented here require that reliable estimates of reporting rate and of the age distribution of the surface catches are available. As such, the level of precision and potential biases in the resulting estimates of $P^*_{A,a,c,r}$ will be critically dependent upon the level of sampling for lengths in the surface fishery and the accuracy with which these are converted to age frequency estimates for the entire catch. The estimates of the CVs presented here for $P^*_{A,a,c,r}$ are only indicative as they are based on a rather arbitrary assumption of multinomial sampling error for the catch at age data with an effective sample size of 50. In most cases, the error in the catch at age tends to be an equal or dominant contributor (compared to the error in the tag returns) to the estimates of the CVs (Table 2). Thus, substantial reduction in

the effective sample sizes in these calculations would, in most cases, result in substantial increases in the CVs of the estimates. If meaningful estimates of precision are to be derived from these tagging experiments, it is critical that appropriate statistical models be developed for the actual estimation procedure for the catch at age data. Similarly, the estimates of the tag reporting rates in the surface fishery are critical. The CV associated with these could be a substantial contributor to the overall precision of the estimates if not estimated with reasonable precision. Moreover, any bias in the reporting rate would act as a multiplicative bias on $P^*_{A,a,c,r}$.

It should be noted that with the assumption of complete mixing and the existence of reporting rates estimates for the surface fishery, the tag returns from the surface fishery could be used along with estimates of the catch at age in both the surface and longline fisheries to estimate the reporting rates in the longline fisheries (see Hearn et al. 2003). The tag return data could then be used in a Brownie model estimation framework as envisioned in the original design of the SRP tagging program. Incorporation of any reported tags from the longline fishery would potentially add some, but little, information to the overall estimates (e.g. in a Brownie estimation framework, estimates of total mortality do not require estimates of reporting rates). However, the precision and accuracy of the mortality estimates would be highly dependent upon the variance and potential biases in the estimated catch at age data from the longline fisheries, since the estimates of the reporting rates are directly dependent upon these data. In addition to this dependence on the catch at age estimates from the longline fishery, a major disadvantage of using such a Brownie approach as compared to the Petersen-type approach used here is that it would provide little scope for diagnostics and testing assumptions of the underlying model. Moreover, the biases in the mortality rate estimates from any incomplete mixing would tend to be amplified by estimating the longline reporting rates this way[6].

Fully appropriate direct incorporation of the estimates of $P^*_{A,a,c,r}$ into a statistical catch at age stock assessment model similar to those being used for SBT would be problematical because of the need to account for the double use of the catch at age data (i.e. once in the estimates of $P^*_{A,a,c,r}$ and once as a directly fitted component in the assessment). Alternatively, the tag

---

[6] For example, in a case with low mixing into the areas of the longline fishery, a low number of tag returns from the longline fishery relative to the number of fish caught would incorrectly be considered to represent low reporting rates. Thus, the estimated number of actual recaptures would be too high, which would further bias upward the estimates of fishing mortality rates.

return data could be incorporated by using the stock assessment model to predict the expected number of returns from the Australian surface fishery conditional on the rest of the model structure and parameters (this would be similar to the approach currently being used in the stock assessment models and, functionally, would be more similar to a Brownie model approach). However, developing an appropriate likelihood for the tag return component (and thus determining the appropriate weight to be given to the tag data in the overall estimation) would be challenging.

Finally, one advantage of using the approach developed here when there is little or no information on the reporting rates from the longline fisheries is that it can provide an indicator of trends in juvenile abundances over the period of the tagging experiments independent of any assumptions about tag returns and catches in the longline fishery. Such indicators can provide a useful independent check on overall complex stock assessment results.

Literature Cited

Anon. 2001a. Report of the fifth meeting of the Scientific Committee. CCSBT. 19-24 March. Tokyo, Japan.

Anon. 2001b. CCSBT Report of the eighth annual meeting. 15-19 October 2001. Miyako, Japan.

Anon. 2003. CCSBT Report of the eighth meeting of the Scientific Committee. 1-4 September 2003. Christchurch, New Zealand.

Hearn, W.S., J.M. Hoenig, K.H. Pollock, and D.A. Hepworth. 2003. Tag reporting rate estimation: 3.Use of planted tags in one component of a multiple-component fishery. N. Am. J. Fish. Manage. 23:66-77.

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

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1996. Estimation of natural and fishing mortality for juvenile southern bluefin tuna based on multi-year tagging of cohorts. CCSBT-SC/96/11.

Polacheck, T., W. Hearn, C. Millar and C. Stanley. 1998. Updated estimates of mortality rates for juvenile SBT from multi-year tagging cohorts. CCSBT-SC/9808/20.

Polacheck, T., J.P. Eveson and G.M. Laslett. 2003. Exploring the trade-off between tag releases and observer coverage in the estimation of mortality rates through an integrated Brownie and Petersen mark-recapture estimation approach. CCSBT-ESC/0309/22.

Polacheck, T., J.P. Eveson and G.M. Laslett. 2004. Review of methods for estimating tag reporting rates and their applicability to SBT longline fisheries. CCSBT-ESC/0409/14.

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. Environmental and Ecological Statistics 9:57-69.

Seber, G.A.F. 1973. The estimation of animal abundance and related parameters. Charles Griffin. London. 506pp.

Acknowledgements

Table 1: Number of tag releases by cohort, area and cohort age. Note a small number of tags estimated to be age 3 were released in WA (88), and a small number of tags estimated be age 0, 4 or 5 were released in the GAB (291).

	WA		GAB		
Cohort	1	2	1	2	3
1988	0	0	0	0	810
1989	0	354	0	2773	1096
1990	2645	891	654	3755	2692
1991	2111	289	33	2648	3640
1992	4522	49	376	3109	2627
1993	8442	1756	561	4143	1511
1994	8170	0	415	2518	526
1995	0	0	82	592	0
1996	0	0	884	0	0

Table 2: Comparison of the bootstrap/Monte-Carlo CV estimates for $P^*_{A,a,c,r}$ when the estimates are based on: the bootstrap component for the tagging data only; the Monte-Carlo component for the catch-at-age data only; and both. R is the total number of tags recovered from a cohort for ages 2 to 4 in the case of age 1 releases, and for ages 3 to 4 in the case of age 2 releases.

				CV		
Release age	Cohort	$P^*_{A,a,c,r}$	R	tag & catch	tag only	catch only
1	1990	5.18	57	0.27	0.24	0.11
	1991	2.89	75	0.14	0.10	0.10
	1992	2.24	281	0.17	0.12	0.12
	1993	2.28	782	0.12	0.06	0.11
	1994	2.19	820	0.12	0.04	0.11
	1995	1.18	15	0.37	0.36	0.11
	1996	3.70	80	0.29	0.28	0.08
2	1989	2.17	104	0.17	0.13	0.11
	1990	2.53	127	0.13	0.06	0.11
	1991	2.86	80	0.14	0.09	0.11
	1992	1.91	203	0.14	0.07	0.12
	1993	1.49	665	0.13	0.06	0.12
	1994	1.00	425	0.12	0.04	0.12
	1995	1.01	93	0.12	0.04	0.11
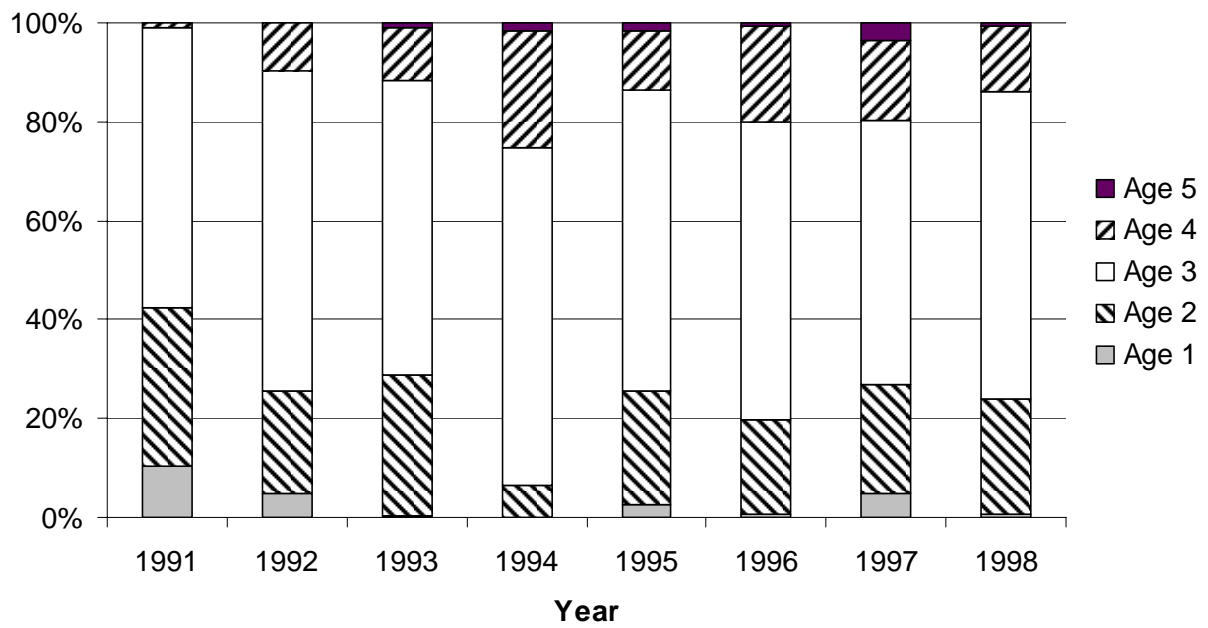
Figure 1: Estimated age composition of Australian surface fishery catches in the Great Australian Bight by fishing season (e.g. 1991 refers to the 1990/1991 fishing season).
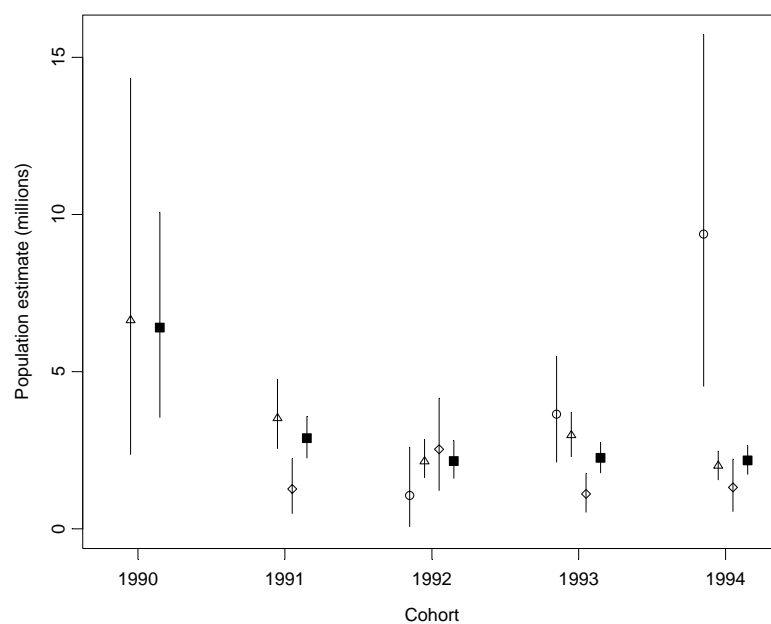
Figure 2: Petersen estimates for the number of 1-year-old SBT ($P^*_{A,a,c,r}$) based on age 1 releases from Western Australia. Circles represent estimates based on age 2 returns, triangles on age 3 returns, diamonds on age 4 returns, and solid squares on returns from ages 2-4 pooled. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown.
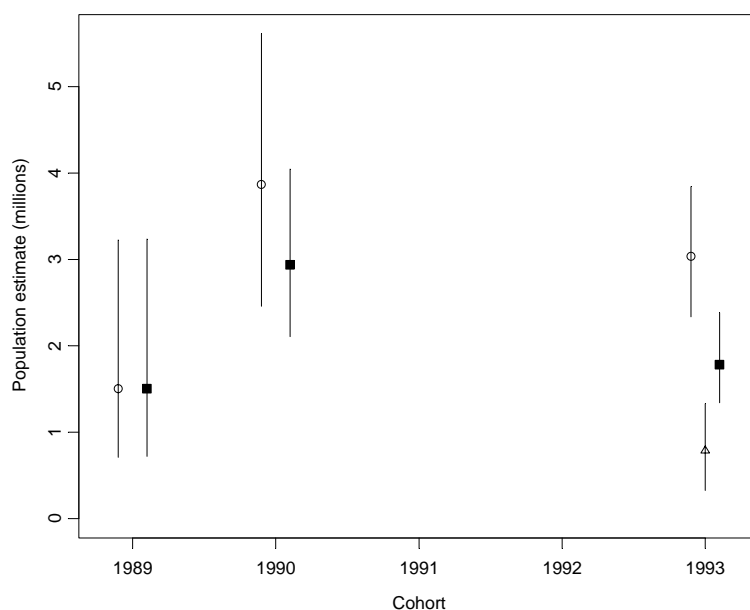
Figure 3: Petersen estimates for the number of 2-year-old SBT ($P^*_{A,a,c,r}$) based on age 2 releases from Western Australia. Circles represent estimates based on age 3 returns, triangles on age 4 returns, and solid squares on returns from ages 3 and 4 pooled. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown.
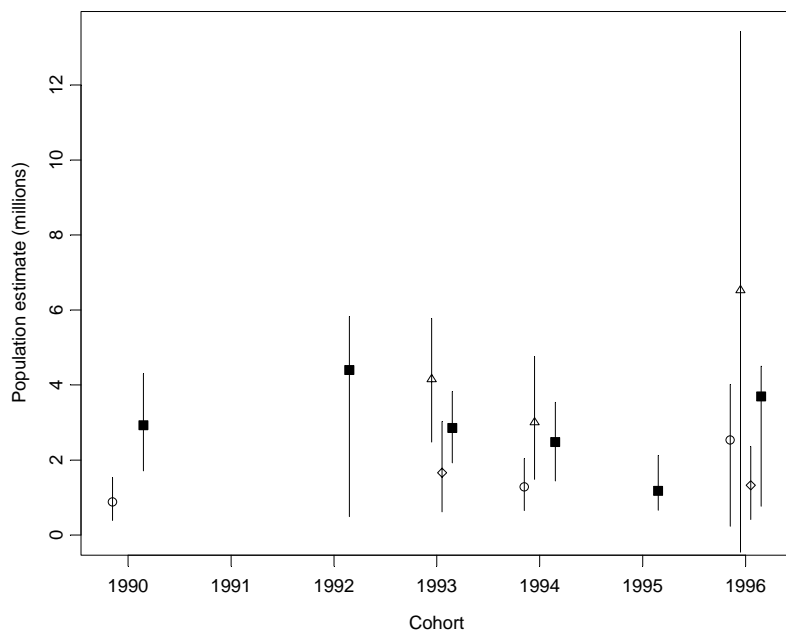
Figure 4: Petersen estimates for the number of 1-year-old SBT ($P^*_{A,a,c,r}$) based on age 1 releases from the Great Australian Bight. Circles represent estimates based on age 2 returns, triangles on age 3 returns, diamonds on age 4 returns, and solid squares on returns from ages 2-4 pooled. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown.
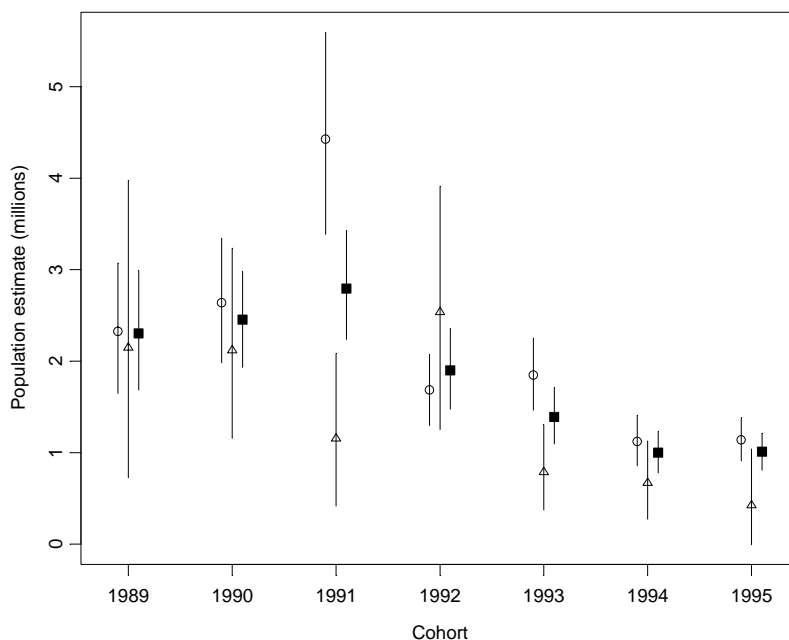
Figure 5: Petersen estimates for the number of 2-year-old SBT ($P^*_{A,a,c,r}$) based on age 2 releases from the Great Australian Bight. Circles represent estimates based on age 3 returns, triangles on age 4 returns, and solid squares on returns from ages 3 and 4 pooled. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown.
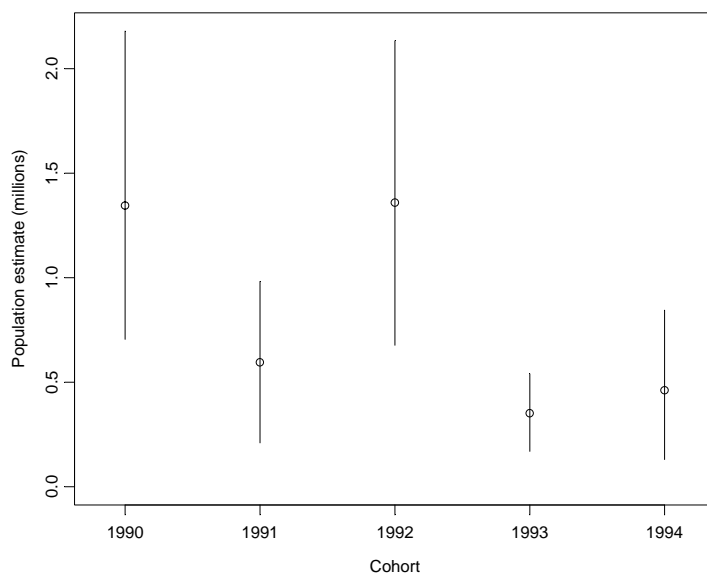
Figure 6: Petersen estimates for the number of 3-year-old SBT ($P^*_{A,a,c,r}$) based on age 3 releases from the Great Australian Bight. The estimates are based on age 4 returns. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown.
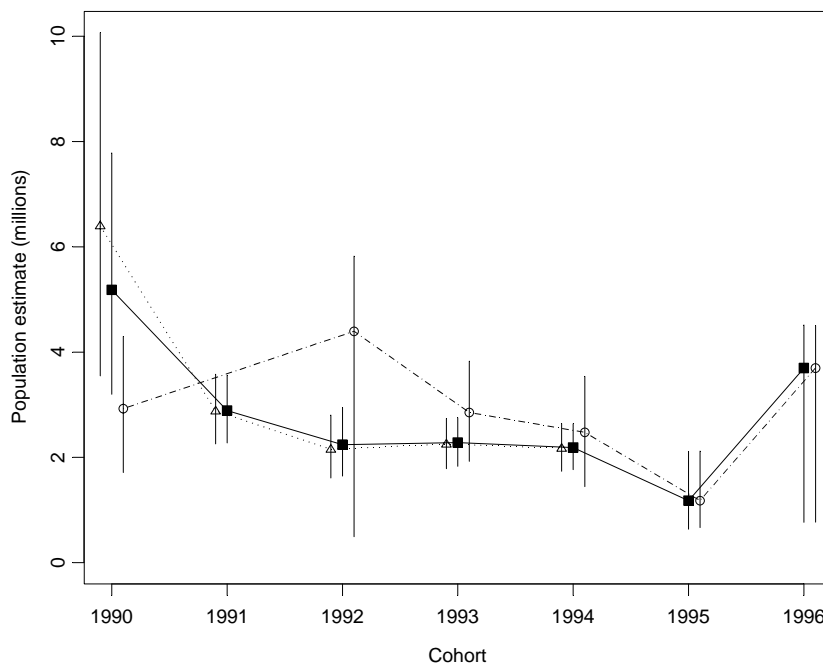
Figure 7: Comparison of Petersen estimates for the number of 1-year-old SBT ($P^*_{A,a,c,r}$) based on age 1 releases from Western Australia (triangles), the Great Australian Bight (circles), and both areas combined (solid squares). The estimates shown are based on the pooled returns and catches for ages 2 to 4. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown. Note that slight differences in the confidence intervals when there were only releases in one area (1995 and 1996) represent different realizations of 1000 bootstrap/Monte-Carlo estimates

.
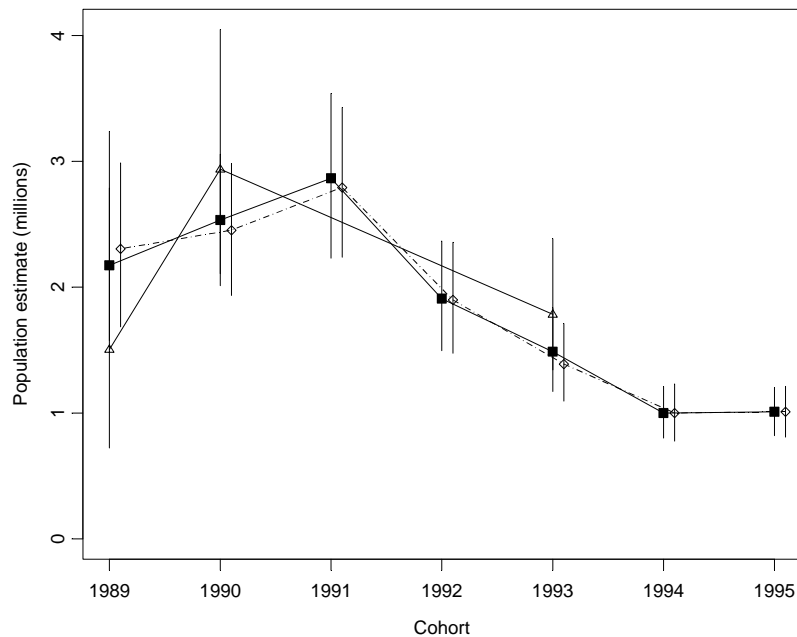
Figure 8: Comparison of Petersen estimates for the number of 2-year-old SBT ($P^*_{A,a,c,r}$) based on age 2 releases from Western Australia (triangles), the Great Australian Bight (circles), and both areas combined (solid squares). The estimates shown are based on the pooled returns and catches for ages 3 and 4. Error bars are estimated 90% confidence intervals (see text). Only estimates based on more than 10 returns are shown. Note that slight differences in the confidence intervals when there were only releases in one area (1994 and 1995) represent different realizations of 1000 bootstrap/Monte-Carlo estimates.
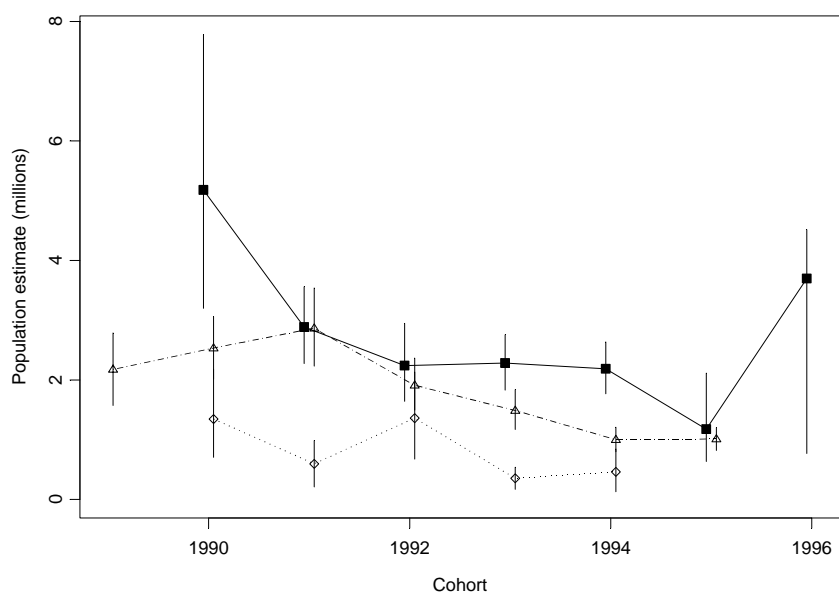
Figure 9: Comparison of Petersen estimates for the number of 1-, 2- and 3-year-old SBT ($P_{A,a,c,r}^{*}$) based on age 1 releases (solid squares), age 2 releases (triangles) and age 3 releases (diamonds), respectively, from Western Australia and the Great Australian Bight combined. Note there were no age 3 releases in Western Australia. The estimates shown are based on the pooled returns and catches for each age of release. Error bars are estimated 90% confidence intervals (see text).
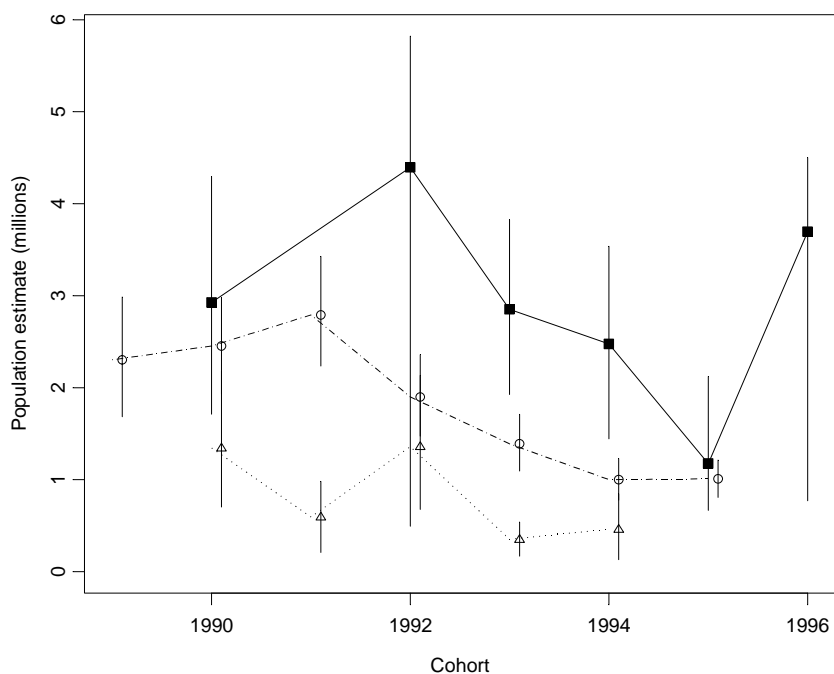
Figure 10: Comparison of Petersen estimates for the number of 1-, 2- and 3-year-old SBT ($P_{A,a,c,r}^{*}$) based on age 1 releases (solid squares), age 2 releases (circles), and age 3 releases (triangles), respectively, from the Great Australian Bight only. The estimates shown are based on the pooled returns and catches for each age of release. Error bars are estimated 90% confidence intervals (see text).

Appendix 1

Tables of values for $P^*_{A,a,c,r} / P_{A,c}$ for a range of values for the parameters $\varphi_{A,c,r}$, $\rho_{A,a,c,r}$,

ρ_{A,a,c,r^*}, $z^+_{A,a-1,c,r}$ and $z^+_{A,a-1,c,r^*}$. Note that in the results presented here $z^+_{A,a-1,c,r}$ and $z^+_{A,a-1,c,r^*}$

have been separated into a fishing and natural mortality component with natural morality

assumed to be independent of region r. Thus,

$$z^+_{A,a-1,c,r} = f^+_{A,a-1,c,r} + m^+_{A,a-1,c} \quad \text{and}$$

$$z^+_{A,a-1,c,r^*} = f^+_{A,a-1,c,r^*} + m^+_{A,a-1,c}$$

where f and m refer to fishing and natural mortality rates respectively. In the results presented

in this appendix $m^+_{A,a-1,c}$ has been fixed at 0.35.

Table 1: Comparison of the ratio of $P^{*}_{A,a,c,r} / P_{A,c}$ for a range of values for $\rho_{A,a,c,r^{*}}$ and $f^{+}_{A,a-1,c,r^{*}}$ when $\varphi_{A,c,r}$ is fixed at 0.8, $\rho_{A,a,c,r}$ is fixed at 0.8 and $f^{+}_{A,a-1,c,r}$ is fixed at 0.1. Note $m^{+}_{A,a-1,c} = 0.35$.

$\rho_{A,a,c,r^{*}}$	$f^{+}_{A,a-1,c,r^{*}}$															
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
0.10	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.82	0.82	0.82	0.82
0.20	0.86	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
0.30	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.87	0.87
0.40	0.91	0.91	0.91	0.91	0.91	0.91	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
0.50	0.94	0.94	0.94	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.92	0.92	0.92	0.92	0.92
0.60	0.97	0.96	0.96	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.94	0.94
0.70	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.97	0.97	0.97
0.80	1.02	1.02	1.02	1.01	1.01	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99
0.90	1.05	1.05	1.04	1.04	1.04	1.04	1.03	1.03	1.03	1.03	1.03	1.02	1.02	1.02	1.02	1.01
1.00	1.08	1.07	1.07	1.07	1.07	1.06	1.06	1.06	1.06	1.05	1.05	1.05	1.05	1.04	1.04	1.04

Table 2: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for ρ_{A,a,c,r^*} and f^+_{a-1,c,r^*} when $\varphi_{A,c,r}$ is fixed at 0.8, $\rho_{A,a,c,r}$ is fixed at 0.8 and $f^+_{A,a-1,c,r}$ is fixed at 0.5. Note $m^+_{A,a-1,c} = 0.35$.

ρ_{A,a,c,r^*}	$f^+_{A,a-1,c,r^*}$															
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
0.10	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84	0.84
0.20	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.87	0.87	0.87	0.87	0.87
0.30	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
0.40	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.94	0.94	0.94
0.50	1.01	1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98
0.60	1.05	1.04	1.04	1.04	1.04	1.04	1.03	1.03	1.03	1.03	1.02	1.02	1.02	1.02	1.01	1.01
0.70	1.09	1.09	1.08	1.08	1.08	1.07	1.07	1.07	1.07	1.06	1.06	1.06	1.06	1.05	1.05	1.05
0.80	1.13	1.13	1.12	1.12	1.12	1.11	1.11	1.11	1.10	1.10	1.10	1.10	1.09	1.09	1.09	1.08
0.90	1.17	1.17	1.16	1.16	1.16	1.15	1.15	1.15	1.14	1.14	1.14	1.13	1.13	1.13	1.12	1.12
1.00	1.21	1.21	1.20	1.20	1.20	1.19	1.19	1.18	1.18	1.18	1.17	1.17	1.17	1.16	1.16	1.15

Table 3: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for ρ_{A,a,c,r^*} and $f^+_{A,a-1,c,r^*}$ when $\varphi_{A,c,r}$ is fixed at 0.4, $\rho_{A,a,c,r}$ is fixed at 0.8 and $f^+_{A,a-1,c,r}$ is fixed at 0.1. Note $m^+_{A,a-1,c}=0.35$.

ρ_{A,a,c,r^*}	$f^+_{A,a-1,c,r^*}$															
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
0.10	0.48	0.48	0.47	0.46	0.46	0.45	0.45	0.44	0.44	0.43	0.43	0.43	0.42	0.42	0.42	0.42
0.20	0.57	0.55	0.54	0.52	0.51	0.50	0.49	0.48	0.47	0.47	0.46	0.46	0.45	0.45	0.44	0.44
0.30	0.65	0.63	0.60	0.58	0.57	0.55	0.54	0.52	0.51	0.50	0.49	0.48	0.47	0.47	0.46	0.46
0.40	0.73	0.70	0.67	0.65	0.62	0.60	0.58	0.56	0.55	0.53	0.52	0.51	0.50	0.49	0.48	0.47
0.50	0.81	0.78	0.74	0.71	0.68	0.65	0.63	0.61	0.59	0.57	0.55	0.54	0.52	0.51	0.50	0.49
0.60	0.90	0.85	0.81	0.77	0.73	0.70	0.67	0.65	0.62	0.60	0.58	0.57	0.55	0.54	0.52	0.51
0.70	0.98	0.93	0.88	0.83	0.79	0.75	0.72	0.69	0.66	0.64	0.61	0.59	0.57	0.56	0.54	0.53
0.80	1.06	1.00	0.94	0.89	0.84	0.80	0.76	0.73	0.70	0.67	0.64	0.62	0.60	0.58	0.56	0.55
0.90	1.15	1.08	1.01	0.95	0.90	0.85	0.81	0.77	0.74	0.70	0.67	0.65	0.62	0.60	0.58	0.57
1.00	1.23	1.15	1.08	1.01	0.96	0.90	0.85	0.81	0.77	0.74	0.70	0.68	0.65	0.63	0.60	0.58

Table 4: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for ρ_{A,a,c,r^*} and $f^+_{A,a-1,c,r^*}$ when $\varphi_{A,c,r}$ is fixed at 0.8, $\rho_{A,a,c,r}$ is fixed at 0.4 and $f^+_{A,a-1,c,r}$ is fixed at 0.1. Note $m^+_{A,a-1,c}=0.35$.

ρ_{A,a,c,r^*}	$f^+_{A,a-1,c,r^*}$															
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15
0.10	0.86	0.85	0.85	0.84	0.84	0.83	0.83	0.83	0.82	0.82	0.82	0.82	0.82	0.82	0.81	0.81
0.20	0.91	0.90	0.89	0.88	0.87	0.87	0.86	0.85	0.85	0.84	0.84	0.84	0.83	0.83	0.83	0.82
0.30	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88	0.87	0.87	0.86	0.86	0.85	0.85	0.84	0.84
0.40	1.02	1.00	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88	0.87	0.87	0.86	0.85	0.85
0.50	1.08	1.05	1.03	1.00	0.99	0.97	0.95	0.94	0.92	0.91	0.90	0.89	0.88	0.88	0.87	0.86
0.60	1.13	1.10	1.07	1.05	1.02	1.00	0.98	0.96	0.95	0.93	0.92	0.91	0.90	0.89	0.88	0.87
0.70	1.19	1.15	1.12	1.09	1.06	1.03	1.01	0.99	0.97	0.96	0.94	0.93	0.92	0.91	0.90	0.89
0.80	1.24	1.20	1.16	1.13	1.10	1.07	1.04	1.02	1.00	0.98	0.96	0.95	0.93	0.92	0.91	0.90
0.90	1.30	1.25	1.21	1.17	1.13	1.10	1.07	1.05	1.02	1.00	0.98	0.97	0.95	0.94	0.92	0.91
1.00	1.35	1.30	1.25	1.21	1.17	1.14	1.10	1.07	1.05	1.02	1.00	0.98	0.97	0.95	0.94	0.92

Appendix 17: An evaluation of abundance estimates using tags from only one component

Table 5: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for $\varphi_{A,c,r}$ and $\rho_{A,a,c,r}$ when ρ_{A,a,c,r^*} is fixed at 0.2, $f^+_{A,a-1,c,r}$ is fixed at 0.1, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.05. Note $m^+_{A,a-1,c} = 0.35$.

$\varphi_{A,c,r}$	$\rho_{A,a,c,r}$								
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.10	1.99	1.05	0.73	0.57	0.48	0.42	0.37	0.34	0.31
0.20	1.88	1.04	0.76	0.62	0.54	0.48	0.44	0.41	0.39
0.30	1.77	1.04	0.79	0.67	0.59	0.55	0.51	0.48	0.46
0.40	1.66	1.03	0.82	0.72	0.65	0.61	0.58	0.56	0.54
0.50	1.55	1.03	0.85	0.76	0.71	0.68	0.65	0.63	0.62
0.60	1.44	1.02	0.88	0.81	0.77	0.74	0.72	0.71	0.69
0.70	1.33	1.02	0.91	0.86	0.83	0.81	0.79	0.78	0.77
0.80	1.22	1.01	0.94	0.91	0.88	0.87	0.86	0.85	0.85
0.90	1.11	1.01	0.97	0.95	0.94	0.94	0.93	0.93	0.92

Table 6: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for $\varphi_{A,c,r}$ and $\rho_{A,a,c,r}$ when ρ_{A,a,c,r^*} is fixed at 0.6, $f^+_{A,a-1,c,r}$ is fixed at 0.1, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.05. Note $m^+_{A,a-1,c} = 0.35$.

$\varphi_{A,c,r}$	$\rho_{A,a,c,r}$								
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.10	5.78	2.94	1.99	1.52	1.24	1.05	0.91	0.81	0.73
0.20	5.25	2.72	1.88	1.46	1.21	1.04	0.92	0.83	0.76
0.30	4.72	2.51	1.77	1.40	1.18	1.04	0.93	0.85	0.79
0.40	4.18	2.29	1.66	1.35	1.16	1.03	0.94	0.87	0.82
0.50	3.65	2.08	1.55	1.29	1.13	1.03	0.95	0.89	0.85
0.60	3.12	1.86	1.44	1.23	1.10	1.02	0.96	0.92	0.88
0.70	2.59	1.65	1.33	1.17	1.08	1.02	0.97	0.94	0.91
0.80	2.06	1.43	1.22	1.12	1.05	1.01	0.98	0.96	0.94
0.90	1.53	1.22	1.11	1.06	1.03	1.01	0.99	0.98	0.97

Appendix 17: An evaluation of abundance estimates using tags from only one component

Table 7: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for $\varphi_{A,c,r}$ and $\rho_{A,a,c,r}$ when ρ_{A,a,c,r^*} is fixed at 0.6, $f^+_{A,a-1,c,r}$ is fixed at 0.4, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.05. Note $m^+_{A,a-1,c} = 0.35$.

$\varphi_{A,c,r}$	$\rho_{A,a,c,r}$								
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.10	7.76	3.93	2.65	2.02	1.63	1.38	1.19	1.06	0.95
0.20	7.01	3.61	2.47	1.90	1.56	1.34	1.17	1.05	0.96
0.30	6.26	3.28	2.29	1.79	1.49	1.29	1.15	1.05	0.96
0.40	5.51	2.95	2.10	1.68	1.42	1.25	1.13	1.04	0.97
0.50	4.76	2.63	1.92	1.56	1.35	1.21	1.11	1.03	0.97
0.60	4.01	2.30	1.74	1.45	1.28	1.17	1.09	1.03	0.98
0.70	3.25	1.98	1.55	1.34	1.21	1.13	1.06	1.02	0.98
0.80	2.50	1.65	1.37	1.23	1.14	1.08	1.04	1.01	0.99
0.90	1.75	1.33	1.18	1.11	1.07	1.04	1.02	1.01	0.99

Table 8: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for $\varphi_{A,c,r}$ and $\rho_{A,a,c,r}$ when ρ_{A,a,c,r^*} is fixed at 0.6, $f^+_{A,a-1,c,r}$ is fixed at 0.05, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.40. Note $m^+_{A,a-1,c} = 0.35$.

$\varphi_{A,c,r}$	$\rho_{A,a,c,r}$								
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.10	3.91	2.00	1.37	1.05	0.86	0.73	0.64	0.58	0.52
0.20	3.58	1.89	1.33	1.05	0.88	0.76	0.68	0.62	0.58
0.30	3.26	1.78	1.29	1.04	0.89	0.79	0.72	0.67	0.63
0.40	2.94	1.67	1.25	1.03	0.91	0.82	0.76	0.72	0.68
0.50	2.61	1.56	1.20	1.03	0.92	0.85	0.80	0.76	0.73
0.60	2.29	1.45	1.16	1.02	0.94	0.88	0.84	0.81	0.79
0.70	1.97	1.33	1.12	1.02	0.95	0.91	0.88	0.86	0.84
0.80	1.65	1.22	1.08	1.01	0.97	0.94	0.92	0.91	0.89
0.90	1.32	1.11	1.04	1.01	0.98	0.97	0.96	0.95	0.95

Appendix 17: An evaluation of abundance estimates using tags from only one component

Table 9: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for ρ_{A,a,c,r^*} and $\rho_{A,a,c,r}$ when $\varphi_{A,c,r}$ is fixed at 0.8, $f^+_{A,a-1,c,r}$ is fixed at 0.1, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.05. Note $m^+_{A,a-1,c} = 0.35$.

$\rho_{A,a,c,r}$	ρ_{A,a,c,r^*}									
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.10	1.01	1.22	1.43	1.64	1.85	2.06	2.27	2.48	2.69	2.90
0.20	0.91	1.01	1.12	1.22	1.33	1.43	1.54	1.64	1.75	1.85
0.30	0.87	0.94	1.01	1.08	1.15	1.22	1.29	1.36	1.43	1.50
0.40	0.85	0.91	0.96	1.01	1.06	1.12	1.17	1.22	1.27	1.33
0.50	0.84	0.88	0.93	0.97	1.01	1.05	1.09	1.14	1.18	1.22
0.60	0.84	0.87	0.91	0.94	0.98	1.01	1.05	1.08	1.12	1.15
0.70	0.83	0.86	0.89	0.92	0.95	0.98	1.01	1.04	1.07	1.10
0.80	0.83	0.85	0.88	0.91	0.93	0.96	0.98	1.01	1.04	1.06
0.90	0.82	0.85	0.87	0.89	0.92	0.94	0.96	0.99	1.01	1.03
1.00	0.82	0.84	0.86	0.88	0.91	0.93	0.95	0.97	0.99	1.01

Table 10: Comparison of the ratio of $P^*_{A,a,c,r}/P_{A,c}$ for a range of values for ρ_{A,a,c,r^*} and $\rho_{A,a,c,r}$ when $\varphi_{A,c,r}$ is fixed at 0.4, $f^+_{A,a-1,c,r}$ is fixed at 0.1, and $f^+_{A,a-1,c,r^*}$ is fixed at 0.05. Note $m^+_{A,a-1,c} = 0.35$.

$\rho_{A,a,c,r}$	ρ_{A,a,c,r^*}									
	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.10	1.03	1.66	2.29	2.92	3.55	4.18	4.82	5.45	6.08	6.71
0.20	0.72	1.03	1.35	1.66	1.98	2.29	2.61	2.92	3.24	3.55
0.30	0.61	0.82	1.03	1.24	1.45	1.66	1.87	2.08	2.29	2.50
0.40	0.56	0.72	0.87	1.03	1.19	1.35	1.50	1.66	1.82	1.98
0.50	0.53	0.65	0.78	0.90	1.03	1.16	1.28	1.41	1.54	1.66
0.60	0.51	0.61	0.72	0.82	0.93	1.03	1.14	1.24	1.35	1.45
0.70	0.49	0.58	0.67	0.76	0.85	0.94	1.03	1.12	1.21	1.30
0.80	0.48	0.56	0.64	0.72	0.79	0.87	0.95	1.03	1.11	1.19
0.90	0.47	0.54	0.61	0.68	0.75	0.82	0.89	0.96	1.03	1.10
1.00	0.46	0.53	0.59	0.65	0.72	0.78	0.84	0.90	0.97	1.03

Appendix 18:

Review of methods for estimating tag reporting rates and their applicability to southern bluefin tuna longline fisheries

Tom Polacheck, J. Paige Eveson and Geoff M. Laslett

FRDC Project 2002/015

Table of Contents

Abstract

This paper reviews alternative approaches to the use of observers for estimating tag reporting rates and discusses their applicability for obtaining estimates from the SBT longline fisheries. Five basic approaches were identified and reviewed. Of these, only the use of high reward tags appears to have any potential to provide robust and direct estimates. However, the logistic and reward issues associated with this approach would appear to make its realization problematical and also quite expensive. However, developing technologies may make approaches based on automatic detection systems more feasible in the future and, as such, might warrant further investigation if tagging is to be used as a long term monitoring strategy for SBT.

Introduction

As part of its Scientific Research Program, the CCSBT is undertaking a large-scale tagging program of juvenile SBT in order to obtain improved estimates of natural mortality and timely, reliable estimates of current juvenile fishing mortality rates (Anon. 2001). A critical component in the analyses of the data from such a tagging program is the ability to estimate reporting rates (i.e. the proportion of tagged fish, captured by fishermen, actually returned to the CCSBT). The 2003 CCSBT Scientific Committee concluded that the current levels of observer coverage in the Japanese, Korean and Taiwanese longline fleets are not high enough to provide useful estimates of reporting rates from these fleets and thus fishing mortality rates in these fleets (Anon. 2003). The overall implications of this conclusion for the ability of the SRP tagging program to meet its primary objectives were not certain and the Scientific Committee agreed to convene a Technical Group Meeting in conjunction with its next meeting to evaluate this question. One of the terms of reference for this working group was to "consider alternative methods, other than increasing observer coverage, for improving the estimates of reporting rates" (Anon. 2003). The current paper reviews approaches for estimating tag reporting rates and discusses their applicability for obtaining estimates from the SBT longline fisheries. Based on a review of the tagging literature, five basic approaches were identified and are reviewed here:

1. High reward tags
2. Tag seeding
3. Automatic tag detection systems
4. Extrapolation from estimates from the surface fishery
5. Model-based approaches

High reward tags

One approach to estimating reporting rates that has been used in tagging studies is to have a high reward associated with a relatively small fraction of the tags being released (Pollock et al. 2001). Those tags carrying a high reward must be clearly identifiable by a tag finder (e.g. different colours) and carry a sufficiently high reward that 100% of these tags will be returned by all tag finders. The ratio of the return rate (i.e. the number of tags returned divided by the number of releases) of "regular" tags compared to that of high reward tags provides a direct estimate of the reporting rate. Fishery-specific reporting rates can be estimated using this approach by considering within-fishery return rates.

For this approach to be viable, a sufficient number of high reward tags must be caught within each of the fishery components so that a viable reporting rate can be obtained, taking into

account overdispersion in the recapture probabilities and potential incomplete mixing. This might require that a relatively high number of high reward tags are released given the differential age-specific mortality rates between fishery components (e.g. a large number of high reward tags would be vulnerable to capture in the surface fishery at ages 2 and 3, prior to the ages of where fish are being captured in reasonable numbers by longline vessels). High reward tagging also requires that there is sufficient publicity that the fishermen on all vessels in each of the fishery components are aware of the existence of these high reward tags and the rewards they carry. There also must be well developed procedures in place to ensure that these high rewards actually are paid to those finding the tags. In addition, as already stated, the magnitude of the reward must be large enough that 100% of the high reward tags are returned. It is these latter two requirements that would appear to be the most difficult to meet within the SBT longline fisheries.

Past experience has shown that it can often be difficult to ensure that the tag finder is the person who actually receives the reward for a returned tag. This is due to the long time lags between tag finding and actual tag returns as a result of long fishing trips, logistic problems with contacting and getting returns back to individual crewmen, and problems with determining who to actually reward (e.g. the fishing master will often collect tags with the identity of the actual tag finder being lost, and thus removing the incentive to return tags from those actually responsible for finding them).

More problematical is determining a reward that would provide sufficient incentive to ensure a high degree of certainty that 100% of the recaptured high reward tags would be returned. Experience with the provision of rewards for archival tags indicates that a reward with a value on the order of $250A has not been sufficient to ensure near 100% reporting rates (i.e. return rates for archival tags with a $250 reward and conventional tags from longline vessels have been similar and low). The problem of suitable rewards is also compounded by the fact that in some fleets cash rewards are considered unacceptable and the incentive for any particular "in kind" reward is hard to assess.

In summary, high reward tagging does have the potential to provide estimates of the reporting rates from longline vessels. However, the logistic and reward issues would appear to make realization of this potential problematical and also potentially quite expensive.

Tag seeding

Tag seeding is a common approach that is used to estimate tag reporting rates in some commercial fisheries and is being used in the current SRP tagging program to provide estimates of reporting rates from the Australian surface fishery. The basis of tag seeding experiments is to plant tags into fish after they have been captured but before those handling and/or processing them have had a chance to actually recover the tags (e.g. prior to unloading of the catch for processing in a cannery). The proportion of planted tags recovered then provides a direct estimate of the reporting rate. It is critical that the tag seeding is done blind in the sense that the tag finders are unaware of the tag seeding and that tag seeded fish are indistinguishable from wild tagged fish. In the case of pelagic tuna longline fisheries, each fish is individually hauled on board and processed at the time of landing. As such, longlining provides no opportunity for tags to be seeded without the fishermen being aware of the tags being planted in the harvested fish. As such, tag seeding does not provide a viable alternative for estimating reporting rates for the SBT longline fisheries.

Automatic tag detection systems

Automatic tag detection systems (e.g. PIT tags) provide one approach for either eliminating the need to estimate reporting rates altogether or for estimating reporting rates. In the former case, it is essential that systems are in place that would ensure that all tuna landed pass through the detection system and the reliance on fishermen for tags is eliminated. Given the complex array of landings, transhipment and markets that are involved with SBT, a practical and acceptable system that was intended to detect all tags post-landing would be extremely difficult to implement, particularly within the timeframe of the current SRP, and would likely be expensive (both financially and in terms of labour) to operate and maintain. Alternatively, an automatic detection system could conceivably be developed for detecting tagged fish at the point of landing on a longline vessel. Such a system would need to be coupled with a system that ensured (or at least monitored) whether all fish were in fact passed by the automatic detection apparatus (e.g. video monitoring of the landing area?). Development and employment of such systems would require a strong level of commitment and appreciation for their need by industry. The initial cost of such a system would be expensive, but once deployed might be relatively cost effective compared to on-board observers, particularly if tagging was to be used as a long term monitoring approach.

Alternatively, automatic tag detection systems could be used in conjunction with conventional tagging to estimate reporting rates. In this case, a certain fraction of the fish tagged with a conventional tag would be also be tagged with an automatically detectable tag. Then, a proportion of the catch would be scanned for the automatically detectable tags at the time of landing. In this case, the automatically detectable tags would need to be cryptic with respect to the fishermen, which is the case for PIT and coded wire tags. The proportion of conventional tags that were retrieved from fish that were detected with an automatically detectable tag would then provide an estimate of the reporting rate.

This approach requires that both a representative and substantial fraction of the landed catch be scanned for automatically detectable tags. Current technology for automatically detectable tags requires that fish are individually scanned with a detection device or that fish come within a few meters or less of the detection device. This, combined with the large range and variety of landing and marketing options for SBT, would make obtaining representative, large scale samples post-harvesting extremely difficult. Thus, to obtain large scale representative sampling would require a large number of on-the-ground technicians to do the scanning with ability to obtain relevant information on the detected scanned fish (i.e. location and time of capture) and access to the different landing/marketing venues.

It should be noted that technologies are evolving. The potential exists for automatic tag detection systems with much greater detection ranges that require less human involvement and still have high probabilities of detection without direct human monitoring. In this case, the use of automatically detectable tags might be a feasible and cost effective alternative in the future.

Finally, industry has expressed concern about potential consumer issues with past proposals to use automatically detectable tags (e.g. PIT tags). This relates to the placement of the tag and the potential for the tag to go undetected prior to being sold and/or consumed. Placing the tags within the body cavity is not viable if detection is to take place after the fish have been landed since fish are gutted at this point. Previous proposals have suggested placing PIT tags in the head; however, the heads of SBT are sold and eaten. Industry objected to placing such tags within the head because of possible liability if a tag were eaten, plus the potential for

associated negative publicity. Developing acceptable tag placement arrangements which overcome such concerns would also be required prior to any large scale use of automatically detectable tags.

Extrapolation from estimates from the surface fishery

Tag reporting rates can be extrapolated from the surface fishery assuming complete mixing and that reliable estimates of reporting rates are available for the surface fishery (Hearn et al. 2003). This can be done by comparing, for each age class, the ratio of the number of tagged fish returned to the total number of fish caught in the longline fishery to the ratio of the number of tagged fish returned (after correcting for the estimated reporting rate) to the total number of fish caught in the surface fishery. The age-specific reporting rates can then be combined to provide an overall estimate of the reporting rate for a fishery within a year, assuming that reporting rates within a fishery are independent of age of the recaptured fish.

Extrapolations from the surface fishery require that reliable estimates of the catch at age data are available for both the surface fishery and the longline fishing components for which reporting rates are to be estimated. The precision of extrapolated reporting rate estimates for a longline component will be dependent upon the precision of the catch at age estimates for both the surface and longline components. A poorly estimated catch at age distribution in either component will translate into poorly estimated reporting rates.

If the assumption of complete mixing is violated and not accounted for, then biases will be introduced in tag-based estimates of mortality rates and population size, and these biases will be compounded by biases in reporting rate estimates when extrapolation methods are used to estimate them. This is because in the extrapolated reporting rate estimates, low return rates will be taken to signify low reporting rates. Thus, the estimated actual number of returns will be inflated. However, if the low return rate is due to incomplete mixing of tagged fish into areas of longline vulnerability, the actual number of returns should not be inflated (and this would be the case if an independent estimate of the reporting rate was available for the longline component). In addition to the compounding of biases, the reliance on extrapolated reporting rates can preclude the use of diagnostic tests for non-mixing (e.g. it is not possible to separate non-reporting in an area/time strata from lack of tagged fish being in that area/time strata). Moreover, the use of extrapolated reporting rates would prevent the application of more spatially explicit tag recovery models to account for heterogeneity in recapture probabilities as a result of non-mixing.

Model-based approaches

Within a Brownie estimation framework, it is theoretically possible to estimate tag reporting rates from the tagging data alone if reporting rates and natural mortality rates are constant over time[1]. However, except within specific special circumstances, the estimates are extremely imprecise (Hoenig et al. 1998; Pollock et al. 2001). Constancy in the reporting rate for the CCSBT SRP tagging program seems unlikely because of the evolving nature of the return promotional activities combined with the differential and changing age-specific catches within the various fishery components (since overall reporting rates for any age class in a given year are a weighted average of the reporting rates within each fishery component weighted by the catch in that component). Moreover, such model-based estimates of

---

[1] Some relaxation of these constancy assumptions may be possible if multiple cohorts as well as multiple years of tagging of the same cohorts are done.

reporting rates would preclude the use of diagnostic tests for non-mixing or the application of spatially explicit models to account for it.

Discussion

Substantial problems appear to exist with all five approaches identified and reviewed as alternatives to the use of observers for estimating tag reporting rates in the context of SBT longline fisheries. Two of the approaches, namely extrapolation and model-based approaches, are indirect (i.e. they use no direct observational data on the actual reporting rates). The model-based approach provides highly imprecise estimates in most cases and requires restrictive assumptions about constancy in reporting rates and natural mortality rates. Extrapolation from the surface fishery requires accurate and precise estimates of the reporting rates in the surface fishery and of the size/age composition of the catch from all fleets. More importantly, the estimated reporting rates for the longline fisheries are dependent upon the assumption of complete mixing. Violation of this assumption will introduce biases in extrapolated reporting rate estimates, which can compound biases that already exist in tagging estimates of mortality rates and population size when non-mixing is not accounted for. Moreover, extrapolation methods preclude diagnostic testing for non-mixing as well as the use of spatially explicit tag recovery models.

Two of the three approaches that provide direct estimates of reporting rates are logistically infeasible, at least at the current time (i.e. automatic detection systems and tag seeding). However, developing technologies may make automatic detection systems more feasible in the future and might warrant further investigation if tagging is to be used as a long term monitoring strategy. The third direct method, high reward tagging could potentially provide a way to estimate reporting rates. However, it is not clear what would serve as a sufficiently attractive, yet still affordable, award to ensure 100% reporting of such tags, and there are also substantial logistical problems with its implementation. It should be noted that it would be possible to combine the high reward tagging and the observer approaches for estimating reporting rates. The combined data would provide an improvement over data from a single approach (Pollock et al. 2002). However, the relative trade-off and improvements is not straightforward to determine and would depend upon a large number of relatively unknown factors. It is also possible to use high reward tags in just one component of the longline fishery to estimate reporting rates for that component, and then use these estimates along with the extrapolation method to estimate reporting rates in the other longline components; however, all of the assumptions and potential problems associated with the extrapolation method would still apply.

In summary, of the alternative approaches to the use of observers for providing estimates of tag reporting rates from longline vessels, only the use of high reward tags appears to have any potential to provide robust and direct estimates. However, logistic and reward issues make realization of this potential problematical and also potentially quite expensive.

Literature Cited

Anon. 2001. Report of the fifth meeting of the Scientific Committee. CCSBT. 19-24 March. Tokyo, Japan.

Anon. 2003. CCSBT Report of the eighth Meeting of the Scientific Committee. 1-4 September 2003. Christchurch, New Zealand.

Hearn, W.S., J.M., Hoenig, K.H. Pollock, and D. A. 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.

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

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. 2002. 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.

Acknowledgements

Appendix 19:

Updated estimates of tag reporting rates for the 1990s southern bluefin tuna tagging experiments

J. Paige Eveson and Tom Polacheck

FRDC Project 2002/015

Introduction

Tag release and return data from juvenile southern bluefin tuna (SBT) from tagging experiments conducted in the 1990s have been one of the primary data sources for recent SBT stock assessments and for conditioning the operating model being developed for evaluating management procedures for SBT. Estimates of reporting rates are integral for the use of the tagging data within these contexts. Estimates of tag reporting rates for the 1990s tagging experiments were first developed in 1996 (Polacheck et al. 1996) and were subsequently updated to incorporate new, updated and revised data (Polacheck et al. 1997, 1998; Preece et al. 2001). In particular, estimates of the catch by year, age and fishery are critical for the estimation of reporting rates for SBT, and there have been substantive changes to these data over the years reflecting improvements in the estimation of SBT growth and revisions to the estimation of total catch and/or its size distribution. The latest update to the reporting rate estimates was made in 2001 (Preece et al. 2001). Since then, there have been further revisions to the catch data, and, in addition, the operating model used for management procedures evaluation does not use catch data compiled by calendar year, which was the basis for the previous reporting rates, but by fishing seasons instead. At the 2004 CCSBT Scientific Committee meeting, it was agreed that the tag release and recapture data should be re-compiled to reflect the non-calendar year fishing seasons used in the operating model (Anon. 2004, Annex 5). Updated tagging data reflecting this change were provided by CSIRO to the CCSBT in September 2004. However, no provision was made for updating the reporting rate estimates to reflect the most recent catch estimates and the change in the definition of year.

This paper presents updated estimates of reporting rates that use the most recent catch estimates and are consistent with the definitions of fishing seasons used in the operating model. These updated reporting rates were calculated in a similar manner to the way they were calculated for the 2001 assessment, as documented in Appendix 2 of CCSBT-SC/0108/21 (Preece et al. 2001). A few changes were made to the way in which the data were compiled and to the reporting rate options; these changes are described in the present document.

Material and Methods

Definition of Fisheries

Reporting rates are first calculated for separate fishery components and then combined to provide an overall estimate for each fishing season. The fishery definitions used here remained the same as those used previously:

Fishery 1. Australian domestic longline and other miscellaneous catch outside South Australia (there are a few troll and purse seine operations outside SA)
Fishery 2. South Australia surface fishery (mainly pole and line and purse seine, but also a handful of trolling operations)
Fishery 3. Australian farm fishery
Fishery 4. Japanese longline catch *inside* the AFZ on vessels *with* observers

Fishery 5. Japanese longline catch *outside* the AFZ on vessels *with* observers in statistical areas 3-9

Fishery 6. Japanese longline catch *inside* the AFZ on vessels *without* observers

Fishery 7. Japanese longline catch *outside* the AFZ on vessels *without* observers in statistical areas 3-8

Fishery 8. Japanese longline catch *outside* the AFZ on vessels *without* observers in statistical area 9

Note that the Japanese longline fisheries include Australian and New Zealand joint venture operations.

Tag Return Data

The criteria used to filter the release and recapture records remained the same as described in CCSBT-SC/0108/21. The only difference to way in which the tag return data was compiled is that recapture year is defined as November 1 of the previous year to October 31 of the given year, rather than calendar year (which was used in the 2001 analysis).[1] The tag-return data provided by CSIRO to the CCSBT in September 2004 were also compiled using this adjusted year definition. This change was made so that the reporting rates would better correspond to fishing seasons. For example, using January 1 as the start of a year (i.e. calendar year) splits the surface fishing season into two years, whereas using July 1 as the start of a year splits the longline fishing season into two years; using November 1 is a reasonable compromise.

Table 1 gives an updated summary of the number of tags released by year and cohort, and the corresponding number of tags recaptured by year.

Catch Data

The catch data were generally compiled in the same way as described in CCSBT-SC/0108/21. The biggest difference is that catches were compiled by adjusted year (starting November 1), rather than calendar year, for reasons already discussed for the tag return data. A few more minor differences are as follows:

- Some of the historical catch data was updated for the 2004 CCSBT data exchange; in particular, Japan provided updated longline data for the early 1990s and New Zealand provided updated joint venture data.
- Two alternatives are presented for the observer catch data (Fisheries 4 and 6) because of uncertainty about the return of tags when observers were on board a vessel, but not actually observing the catch. This issue arises because a significant percentage of the catch was not actually observed while observers were aboard vessels (~30% for observer vessels within the AFZ and ~17% for RTMP observer vessels). In the first alternative, we assume that all tags are returned from all catches on observer vessels. In the second alternative, we assume that that the reporting rate is 100% only for the catches actually observed by the observer, and

---

[1] This adjusted year definition has always been used for release year and cohort (in both the present analysis and in past analyses).

when the observer is not observing the catch that the reporting rate is the same as for unobserved vessels. In reality, the reporting rates for the unobserved catch while an observer is on board is likely to be somewhere in between 100% and the rate for vessels without observers as the presence of observers is likely to promote tag returns.

Unfortunately the tag data base does not provide any insight into this issue because it only contains information on whether the tag came from a vessel with an observer, not on whether the fish from which the tag came was actually observed. Furthermore, the observer data base does not contain information on when and which tags were recovered while an observer was on board observing the catch. This means that no matter which assumption is used there may be some miss-assignment of tags to whether they came from the observed component. If we assume that 100% of tags are returned from all catches when an observer was on board a vessel, then the age distribution of the observer catches (determined from fish whose lengths were measured) needs to be scaled up to the total number of fish caught on observer vessels. If we assume that tags are only returned from catches actually observed by the observer, then essentially no scaling up is necessary because the number of fish measured for length is almost the same as the number of fish actually observed (a very small number of observed fish are not measured, so we scale up the age distribution to account for these fish)[2].

Reporting Rate Options

In calculating a final reporting rate for each year and age class, we first need to calculate the reporting rate for each fishery. The options considered for each fishery are given in Table 2. There are only minimal changes from 2001, as follows:

- Option b) for Fisheries 2 and 3 (Australian surface and farm fisheries, respectively) is now based on tag seeding data instead of relative returns rates compared to Fishery 1 (Australian longline catches). Using the relative return rate to Fishery 1 did not have a very scientific basis given the small sample size for this fishery and thus high variances. Furthermore, the reporting rates calculated this way tended to be equal or close to 100%. On the other hand, recent analyses of data from tag seeding experiments suggest that reporting rates may have been substantially less than 100% for the Australian farm fishery during the 1990s. Thus, tag seeding experiments conducted in the farms in 1996 and 1997 suggested significantly lower levels of reporting: 76% and 86% respectively (Polacheck 2004). We took the average (81%) and assumed that it is representative of the reporting rates in years 1991 to 1997 in both the farm fishery and the surface fishery.
- Option b) for Fishery 7 (out-of-zone Japanese longline fishery without observers) in 1997 now uses all ages, rather than just ages 5 and older, because the Japanese industry policy of non-retention of fish less than 25kg on unobserved vessels,

---

[2] Note that in previous estimates of reporting rates, option 2 (un-scaled data) was used because the problem of determining which catch data within the observer data base to use as the "observed" catch had not been fully recognized or addressed. nor the issue of which tags were "observer reported tags". Thus, the un-scaled catch (i.e. catches that were actually observed) had been used while the apparent intention was that the scaled catch should have been used.

which was in place in 1995 and 1996, was abandoned in 1997. This change should provide for a more precise estimate of the reporting rate in 1997 (i.e. as a result of the additional tag return data), although the previous exclusion should not have biased the estimate.

As outlined in CCSBT-SC/0108/21, option b) for Fishery 7 would ideally be calculated as the relative return rate compared to Fishery 5; however, lack of RTMP data makes this unfeasible (RTMP data were not made available for 1996 and 1997 and are fairly sparse in other years of the program, i.e. 1992 to 1995). This fact, along with the non-retention of small fish in Fishery 7 in 1995 and 1996, leads to the rather complicated option b) for Fishery 7.

Eight combinations of the reporting rate options for the eight fisheries were considered. Fisheries 1, 4, 5, and 6 have only one option. For the remaining fisheries, the combinations of options that we considered are summarized in Table 3. These are the same combinations presented in CCSBT-SC/0108/21. Now, however, we have the added complexity that we can use either the scaled or un-scaled data for the observer catches in calculating the reporting rates corresponding to these eight options. Until a decision has been reached on which of these two alternatives is believed to be more valid, the eight reporting rate options have been calculated using both the scaled and un-scaled data.

Combining Fishery Specific Reporting Rates

Using the reporting rates calculated for all of the fisheries, we then calculated age-specific reporting rates for each year as a weighted average of the reporting rates for all fisheries (Polacheck et al. 1997; Hearn et al. 1999). The reporting rate for each fishery is weighted by the proportion of catch at age in each fishery. Note that 'all fisheries' includes any catches not accounted for in Fisheries 1 to 8, for which the reporting rate is assumed to be zero.

Results

Table 4 presents updated year-specific reporting rate estimates for the fisheries and reporting rate options specified in Table 2. Results are presented using both the scaled and un-scaled observer catch data. Table 5 provides estimates of the year- and age-specific reporting rates (averaged over fisheries) for each of the eight reporting rate options given in Table 3. Again, results are presented using both the scaled and un-scaled observer catch data.

Discussion

The first eight reporting rate options presented here are analogous to those presented previously for the SBT tagging experiments conducted in the 1990s (Polacheck et al. 1996, 1998; Preece et al. 2001). They have been updated (1) to incorporate recent updates to the estimates of SBT catches and their size distributions; (2) to include corrected estimates of the SBT catches while observers were on board vessels; (3) to

include tag seeding estimates of reporting rates for the Australian surface fishery and (4) to estimate the reporting rates for a fishing year (in contrast to previous calendar year estimates). It should be noted that among the eight options only option 8 is actually information based for the major non-observed fisheries for which a non-zero reporting rate is estimated (i.e. the Australian surface fishery, the Japanese longline fishery in the AFZ, the Japanese longline fishery in Areas 3-8 outside the AFZ and the Japanese longline fishery in Area 9). For the other seven options, the reporting rate for at least one of these fisheries is based on what can be considered the most optimistic assumption for that fishery (e.g. 100% reporting rates for the Australian surface, out-of-zone reporting rates are the same as in-zone, etc.). In this sense, option 8 could be considered the most "realistic" or plausible.

It should be emphasized that the eight reporting options do not span the range of uncertainty in the actual reporting rates. The eight options originally provided a measure of the sensitivity/robustness of the resulting mortality rate estimates and fishing rate mortality rate trends over time to uncertainty in the reporting rates. In particular, they provided a measure of the sensitivity of having direct information for each of the major fishery components contributing to the overall reporting rate in contrast to what would be the most optimistic assumption in the absence of any direct information or data. As such, a comparison of the two sets of four options for each major fishery provides a measure of its contribution to the uncertainty (i.e. options 1-4 relative to options 4-8 for the Australian surface and farm fisheries; options 1,2,5,6 versus 3,4,7,8 for the Japanese longline fishery areas 3 to 8; odd versus even options for the Japanese South African longline fishery (area 9)). In this sense, option 1 could be considered a reasonable upper bound for the reporting rates. However, none of the options could be considered as a reasonable lower bound. Conditional on the available data and information, option 8 could be considered as a "best" estimate (although they are the lowest) as it uses the "best" direct information for each of the main fishery components. However, it should be stressed that there is large uncertainty about these reporting rate estimates as the information/data available for their estimation is quite limited. Even in those cases where direct data are used, rather restrictive assumptions are required (e.g. no temporal changes in the reporting rates in the surface fishery over time). Finally, it should also be noted that these estimates of reporting rates are dependent upon the estimates of the catch at age for all of the different SBT fisheries. Large uncertainty is associated with these in some cases and this uncertainty is not reflected within any of the options presented.

We have also included in this paper are eight additional reporting rate options which take into account the uncertainty about tag reporting when observers are on board a vessel but not actually in the act of observing the catch (i.e. the un-scaled estimates). It should be noted that if 100% of the tags were returned while an observer was on board, independent of whether he was actually observing, then the scaled and un-scaled estimates would be expected to be the same if returned tags could be correctly assigned to whether in fact they came from a fish when the observer was observing or not. In this case, the un-scaled estimates would be preferable as any non-reporting of tags when the observer was not observing would introduce a positive bias into the reporting rates (e.g. the assumption of 100% reporting rates for the observed fishery would be violated), while the un-scaled estimates would be unbiased. However, as noted above, given the way the data have been compiled, it is not possible to determine which tags returned from observer vessels came from fish that were

actually observed. If some tags came from the unobserved catches, this would introduce a negative bias into the un-scaled estimates (e.g. too many tags would be considered to have been returned from the observed portion of the catch and too few from the unobserved). It is undoubtedly the case that at least some tags from the unobserved portion of the catch were returned to the observers while they were on board. Thus, both the un-scaled and scaled estimates potentially contain negative or positive biases respectively. The difference between the two provides a measure of the extent of the possible bias but the data by themselves do not allow the extent of actual bias to be determined. The differences between the scaled and un-scaled estimates are substantive. These differences emphasize the importance of ensuring that detailed and accurate data from observers are recorded and made available if observer data are to be used for the estimation of reporting rates.

Literature Cited

Anon. 2004. CCSBT. Report of the Extended Scientific Committee for the Ninth Meeting of the Scientific Committee. 13-16 September 2004. Seogwipo City, Jeju, Republic of Korea.

Hearn, W., T. Polacheck, K. Pollock and W. Whitelaw. 1999. Estimation of tag reporting rates in age-structured multi-component fisheries where one component has observers. Can. J. Fish. Aquat. Sci. 56: 1255-1265.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw, and C. Stanley. 1996. Estimation of natural and fishing mortality for juvenile southern bluefin tuna based on multi-year tagging of cohorts. CCSBT/SC/96/11.

Polacheck, T., W. Hearn; C. Millar, W. Whitelaw and C. Stanley. 1997. Updated estimates of mortality rates for juvenile SBT from multi-year tagging of cohorts 1997. CCSBT/SC/9707/26.

Polacheck, T., W. Hearn, C. Millar, and C. Stanley. 1998. Updated estimates of mortality rates for juvenile SBT from multi-year tagging cohorts. 1998. CCSBT-SC/9808/20.

Polacheck, T. and C. Stanley. 2004. Update on tag seeding activities and preliminary estimates of reporting rates from the Australian surface fishery based on tag seeding experiments. CCSBT-ESC/0409/15.

Preece, A., T. Polacheck, D. Kolody, P. Eveson, D. Ricard, P. Jumppanen, J. Farley, and T. Davis. 2001. Summary of the primary data inputs to CSIRO's 2001 stock assessment models. CCSBT-SC/0108/21.

Table 1. Summary of release and recapture numbers used in the reporting rate analysis.

Cohort	Release Year	# of Releases	# of Recaptures by Year						
			1991	1992	1993	1994	1995	1996	1997
1988	1991	810	63	8	16	7	1	4	1
1988	1992	1	0	0	0	0	0	0	1
1989	1991	3127	102	147	58	34	21	7	5
1989	1992	1097	0	57	18	11	10	4	2
1989	1993	22	0	0	2	0	0	1	0
1989	1994	4	0	0	0	0	0	0	0
1990	1991	3299	20	40	46	23	15	4	4
1990	1992	4646	0	88	157	100	33	12	9
1990	1993	2777	0	0	65	78	31	15	15
1990	1994	111	0	0	0	4	2	0	0
1990	1995	3	0	0	0	0	0	0	0
1991	1992	2144	0	1	21	56	37	12	7
1991	1993	2937	0	0	60	68	69	21	11
1991	1994	3640	0	0	0	77	146	30	41
1991	1995	101	0	0	0	0	1	3	1
1991	1996	1	0	0	0	0	0	0	0
1992	1993	4898	0	0	2	40	202	93	63
1992	1994	3158	0	0	0	29	167	77	55
1992	1995	2629	0	0	0	0	54	102	75
1992	1996	24	0	0	0	0	0	1	1
1993	1994	9003	0	0	0	4	110	399	370
1993	1995	5899	0	0	0	0	83	396	367
1993	1996	1511	0	0	0	0	0	115	205
1993	1997	17	0	0	0	0	0	0	7
1994	1995	8585	0	0	0	0	0	87	637
1994	1996	2518	0	0	0	0	0	75	344
1994	1997	526	0	0	0	0	0	0	91
1995	1996	82	0	0	0	0	0	0	3
1995	1997	592	0	0	0	0	0	0	15
1996	1997	884	0	0	0	0	0	0	1

Table 2. Summary of the reporting rate options considered for each fishery.

Fishery	Description	Reporting Rate
1	AUS LL & misc.	100%
2	AUS surface	a) 100% b) 81% (based on tag seeding experiments[3])
3	AUS farm	a) 100% except in 1996 where it is estimated as 53% due to the mass farm deaths[4]. b) 81% (based on tag seeding experiments) except in 1996 where it is estimated as 53% of 81%.
4	JPN in-AFZ with observers	100%
5	JPN out-of-AFZ with observers	100%
6	JPN in-AFZ without observers	Calculated as the relative return rate compared to Fishery 4.
7	JPN out-of-AFZ without observers areas 3 to 8	a) Same as Fishery 6. b) 1991-1994: calculated as the relative return rate compared to Fisheries 4 and 5 combined; 1995: same as 1991-1994 except only using fish 5 years of age and older; 1996: calculated as the relative return rate for Fisheries 5 and 7 combined compared to Fishery 4, and only using fish 5 years of age and older; 1997: same as 1996 but using all ages.
8	JPN out-of-AFZ without observers area 9	a) Same as Fishery 7a) b) 40% of Fishery 7a) [5] c) Same as Fishery 7b) d) 40% of Fishery 7b)

Table 3. The 8 combinations of the reporting rate options presented in Table 2 that we considered.

Fishery	Combination							
	1	2	3	4	5	6	7	8
2	a	a	a	a	b	b	b	b
3	a	a	a	a	b	b	b	b
7	a	a	b	b	a	a	b	b
8	a	b	c	d	a	b	c	d

---

[3] See Polacheck et al. 2004 for further information.
[4] See Polacheck et al. 1998 for further information.
[5] See Polacheck et al. 1996 for reasoning behind 40%.

Table 4. Year-specific reporting rate estimates for the fisheries and options presented in Table 2.

a) Results using scaled up observer data (assumes 100% of tags are returned from observer vessels regardless of whether all catches are directly observed).

Fishery	Option	1991	1992	1993	1994	1995	1996	1997
1	-	1	1	1	1	1	1	1
2	a	1	1	1	1	1	1	1
2	b	0.81	0.81	0.81	0.81	0.81	0.81	0.81
3	a	1	1	1	1	1	0.529	1
3	b	0.81	0.81	0.81	0.81	0.81	0.429	0.81
4	-	1	1	1	1	1	1	1
5	-	1	1	1	1	1	1	1
6	-	0.371	0.363	0.687	0.83	0.313	0.707	1
7	a	0.371	0.363	0.687	0.83	0.313	0.707	1
7	b	0.331	0.415	0.492	0.327	0.177	0.413	1
8	a	0.371	0.363	0.687	0.83	0.313	0.707	1
8	b	0.331	0.415	0.492	0.327	0.177	0.413	1
8	c	0.148	0.145	0.275	0.332	0.125	0.283	0.4
8	d	0.132	0.166	0.197	0.131	0.071	0.165	0.4

b) Results using un-scaled observer data (assumes 100% of tags are returned from observer vessels when the observer is actually observing catches, and that the reporting rate is the same as for unobserved vessels when the observer is not observing).

Fishery	Option	1991	1992	1993	1994	1995	1996	1997
1	-	1	1	1	1	1	1	1
2	a	1	1	1	1	1	1	1
2	b	0.81	0.81	0.81	0.81	0.81	0.81	0.81
3	a	1	1	1	1	1	0.529	1
3	b	0.81	0.81	0.81	0.81	0.81	0.429	0.81
4	-	1	1	1	1	1	1	1
5	-	1	1	1	1	1	1	1
6	-	0.234	0.237	0.45	0.524	0.209	0.408	0.957
7	a	0.234	0.237	0.45	0.524	0.209	0.408	0.957
7	b	0.231	0.295	0.363	0.243	0.134	0.285	0.772
8	a	0.234	0.237	0.45	0.524	0.209	0.408	0.957
8	b	0.231	0.295	0.363	0.243	0.134	0.285	0.772
8	c	0.094	0.095	0.18	0.21	0.084	0.163	0.383
8	d	0.092	0.118	0.145	0.097	0.053	0.114	0.309

Table 5. Year- and age-specific reporting rate estimates (averaged over all fisheries) for the eight options presented in Table 3.

a) Results using scaled up observer data (assumes 100% of tags are returned from observer vessels regardless of whether all catches are directly observed).

Option 1:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.988	0.971	0.937	0.692	0.89	0.546	0.99
2	0.785	0.617	0.865	0.717	0.735	0.504	0.951
3	0.734	0.648	0.769	0.814	0.767	0.589	0.906
4	0.307	0.413	0.657	0.773	0.492	0.557	0.841
5	0.305	0.346	0.659	0.749	0.355	0.523	0.814
6	0.249	0.31	0.632	0.711	0.289	0.488	0.751
7	0.276	0.304	0.635	0.725	0.292	0.502	0.754
8	0.309	0.332	0.651	0.746	0.304	0.511	0.755

Option 2:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.987	0.971	0.937	0.692	0.916	0.538	0.988
2	0.779	0.616	0.85	0.695	0.776	0.494	0.946
3	0.72	0.634	0.731	0.765	0.775	0.587	0.89
4	0.216	0.337	0.547	0.71	0.477	0.522	0.753
5	0.213	0.255	0.477	0.658	0.315	0.36	0.638
6	0.185	0.257	0.451	0.597	0.251	0.375	0.54
7	0.195	0.255	0.481	0.611	0.238	0.395	0.568
8	0.202	0.268	0.5	0.635	0.25	0.386	0.596

Option 3:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.987	0.973	0.926	0.522	0.911	0.324	0.99
2	0.779	0.626	0.849	0.521	0.759	0.469	0.951
3	0.729	0.654	0.729	0.721	0.77	0.585	0.906
4	0.285	0.438	0.574	0.641	0.467	0.51	0.841
5	0.28	0.38	0.535	0.501	0.29	0.35	0.814
6	0.229	0.339	0.511	0.436	0.215	0.333	0.751
7	0.253	0.333	0.521	0.45	0.209	0.338	0.754
8	0.281	0.365	0.531	0.462	0.217	0.327	0.755

Option 4:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.986	0.973	0.926	0.522	0.911	0.32	0.988
2	0.773	0.624	0.838	0.512	0.758	0.469	0.946
3	0.717	0.638	0.703	0.702	0.769	0.584	0.89
4	0.204	0.352	0.495	0.616	0.456	0.491	0.753
5	0.198	0.276	0.405	0.465	0.267	0.255	0.638
6	0.171	0.278	0.381	0.391	0.193	0.267	0.54
7	0.18	0.277	0.411	0.405	0.179	0.275	0.568
8	0.186	0.292	0.423	0.418	0.187	0.254	0.596

Option 5:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.934	0.885	0.937	0.692	0.753	0.545	0.808
2	0.666	0.57	0.777	0.704	0.65	0.414	0.78
3	0.614	0.553	0.691	0.712	0.638	0.479	0.743
4	0.299	0.388	0.633	0.716	0.438	0.477	0.726
5	0.305	0.346	0.656	0.741	0.346	0.517	0.773
6	0.249	0.31	0.632	0.711	0.288	0.487	0.747
7	0.276	0.304	0.635	0.725	0.292	0.502	0.754
8	0.309	0.332	0.651	0.746	0.304	0.511	0.755

Option 6:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.933	0.885	0.937	0.692	0.753	0.538	0.805
2	0.66	0.568	0.762	0.681	0.649	0.413	0.775
3	0.6	0.539	0.653	0.663	0.637	0.478	0.727
4	0.208	0.313	0.523	0.653	0.418	0.444	0.639
5	0.213	0.255	0.474	0.65	0.305	0.355	0.597
6	0.185	0.257	0.451	0.596	0.25	0.375	0.537
7	0.195	0.255	0.481	0.611	0.238	0.395	0.568
8	0.202	0.268	0.5	0.635	0.25	0.386	0.596

Option 7:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.934	0.887	0.926	0.522	0.748	0.324	0.808
2	0.66	0.578	0.761	0.507	0.632	0.388	0.78
3	0.61	0.56	0.651	0.619	0.632	0.476	0.743
4	0.277	0.414	0.549	0.584	0.408	0.432	0.726
5	0.28	0.38	0.532	0.493	0.281	0.345	0.773
6	0.229	0.339	0.511	0.435	0.214	0.333	0.747
7	0.253	0.333	0.521	0.449	0.209	0.338	0.754
8	0.281	0.365	0.531	0.462	0.217	0.327	0.755

Option 8:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.933	0.887	0.926	0.522	0.748	0.32	0.805
2	0.654	0.577	0.75	0.498	0.632	0.387	0.775
3	0.597	0.543	0.625	0.6	0.632	0.475	0.727
4	0.196	0.327	0.471	0.559	0.397	0.413	0.639
5	0.198	0.276	0.402	0.457	0.258	0.25	0.597
6	0.171	0.278	0.381	0.39	0.192	0.267	0.537
7	0.18	0.277	0.411	0.405	0.179	0.275	0.568
8	0.186	0.292	0.423	0.418	0.187	0.254	0.596

b) Results using un-scaled observer data (assumes 100% of tags are returned from observer vessels only when the observer is actually observing catches, and that the reporting rate is the same as for unobserved vessels when the observer is not observing).

Option 1:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.985	0.965	0.88	0.456	0.902	0.317	0.989
2	0.74	0.536	0.773	0.517	0.75	0.467	0.949
3	0.702	0.603	0.671	0.732	0.763	0.583	0.904
4	0.213	0.321	0.497	0.633	0.452	0.504	0.831
5	0.197	0.235	0.456	0.524	0.277	0.338	0.791
6	0.161	0.213	0.433	0.48	0.212	0.315	0.724
7	0.177	0.208	0.434	0.484	0.209	0.32	0.726
8	0.198	0.226	0.446	0.497	0.216	0.312	0.726

Option 2:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.985	0.965	0.88	0.456	0.902	0.312	0.987
2	0.736	0.535	0.763	0.503	0.75	0.467	0.944
3	0.693	0.594	0.647	0.701	0.763	0.583	0.889
4	0.155	0.272	0.425	0.593	0.438	0.485	0.747
5	0.139	0.175	0.336	0.466	0.25	0.244	0.623
6	0.12	0.178	0.314	0.407	0.186	0.25	0.522
7	0.126	0.175	0.333	0.412	0.172	0.258	0.548
8	0.131	0.185	0.346	0.426	0.18	0.24	0.573

Option 3:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.985	0.967	0.875	0.358	0.899	0.224	0.984
2	0.739	0.545	0.766	0.403	0.74	0.457	0.94
3	0.702	0.61	0.654	0.679	0.76	0.582	0.895
4	0.211	0.35	0.46	0.558	0.434	0.485	0.791
5	0.195	0.274	0.4	0.384	0.241	0.265	0.693
6	0.159	0.246	0.379	0.324	0.171	0.25	0.608
7	0.175	0.24	0.382	0.329	0.162	0.251	0.606
8	0.195	0.264	0.392	0.337	0.167	0.235	0.602

Option 4:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.985	0.967	0.875	0.358	0.899	0.221	0.982
2	0.735	0.544	0.758	0.397	0.74	0.456	0.936
3	0.693	0.599	0.634	0.665	0.759	0.581	0.883
4	0.154	0.288	0.402	0.54	0.426	0.472	0.723
5	0.138	0.199	0.304	0.357	0.223	0.2	0.557
6	0.119	0.202	0.282	0.291	0.154	0.205	0.445
7	0.125	0.2	0.301	0.295	0.139	0.208	0.462
8	0.129	0.212	0.312	0.304	0.144	0.185	0.479

Option 5:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.932	0.879	0.88	0.456	0.739	0.317	0.806
2	0.621	0.488	0.686	0.503	0.624	0.386	0.778
3	0.583	0.509	0.593	0.63	0.625	0.474	0.741
4	0.205	0.297	0.473	0.576	0.392	0.426	0.717
5	0.197	0.235	0.453	0.516	0.268	0.332	0.751
6	0.161	0.213	0.433	0.479	0.211	0.315	0.72
7	0.177	0.208	0.434	0.484	0.209	0.32	0.726
8	0.198	0.226	0.446	0.497	0.216	0.312	0.726

Option 6:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.932	0.879	0.88	0.456	0.739	0.312	0.804
2	0.617	0.487	0.676	0.489	0.624	0.386	0.773
3	0.574	0.499	0.569	0.599	0.625	0.474	0.726
4	0.148	0.247	0.401	0.536	0.378	0.407	0.633
5	0.139	0.175	0.334	0.459	0.241	0.238	0.582
6	0.12	0.178	0.314	0.406	0.185	0.25	0.519
7	0.126	0.175	0.333	0.412	0.172	0.258	0.548
8	0.131	0.185	0.346	0.426	0.18	0.24	0.573

Option 7:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.932	0.882	0.875	0.358	0.736	0.224	0.802
2	0.62	0.498	0.678	0.39	0.614	0.376	0.769
3	0.582	0.516	0.576	0.577	0.622	0.473	0.733
4	0.203	0.326	0.436	0.501	0.375	0.407	0.676
5	0.195	0.273	0.398	0.376	0.232	0.259	0.653
6	0.159	0.246	0.378	0.323	0.17	0.25	0.604
7	0.175	0.24	0.382	0.329	0.162	0.251	0.606
8	0.195	0.264	0.392	0.337	0.167	0.235	0.602

Option 8:

Age	1991	1992	1993	1994	1995	1996	1997
1	0.932	0.881	0.875	0.358	0.736	0.221	0.8
2	0.616	0.497	0.67	0.383	0.613	0.375	0.765
3	0.573	0.504	0.556	0.562	0.621	0.472	0.721
4	0.147	0.264	0.377	0.483	0.366	0.394	0.608
5	0.138	0.199	0.301	0.35	0.214	0.194	0.516
6	0.119	0.202	0.282	0.29	0.153	0.205	0.441
7	0.125	0.2	0.301	0.295	0.139	0.208	0.462
8	0.129	0.212	0.312	0.304	0.144	0.185	0.479

Appendix 20:

Review of historical (pre-1990) SBT tagging data for potential application of Brownie mortality estimation models

Tom Polacheck and J. Paige Eveson

FRDC Project 2002/015

Introduction

Extensive tagging experiments for southern bluefin tuna (SBT) have been conducted beginning in the 1960s. These tagging experiments can be divided into three basic groups based on the methods used and the objective/design for the experiments. These are pre-1980, 1983-84 and 1990s. The tagging experiments conducted in the 1990s were designed to tag multiple cohorts at three consecutive ages with the intention that the release and return data could be analyses within a Brownie estimation framework (Polacheck et al. 1996, 1997, 1998) and are examined further in Appendices 15 and 16. However, the pre-1990 tagging experiments were not designed within the context of this Brownie estimation framework. The stated objectives of the pre-1980 tagging experiments were to delineate stock boundaries, show migration paths, confirm growth rates obtained by other methods, and hopefully to assist in the stock assessments. The focus of the 1983-84 experiments was more quantitative with a focus on estimating mortality rates, fishery interaction and local populations. Improved tagging methods were also used in the post-1980 experiments, particularly with the introduction of vinyl tagging cradles (Caton 1991).

Over 50,000 conventional tags were released prior to 1990 and nearly all of these are potentially usable for quantitative mortality and/or population estimation[1]. Extensive use of these data have been used for the quantitative estimation of growth and tag shedding (e.g. Hearn 1986; Hampton 1989; Hampton and Kirkwood 1989; Hearn et al. 1991; Polacheck et al. 2003, 2004; Hearn and Polacheck 2003) as well as some qualitative examination of with respect to migration/movements (Hynd, 1969; Murphy, 1977; Murphy and Majkowski, 1981). However, with some limited exceptions (e.g. Lucas 1974; Hampton 1991; and Hearn et al. 1987) there has been relatively limited use of these data for estimation of mortality rates or for quantitative use in the stock assessment process. In addition, since these previous analyses, there have been extensive developments in estimation methods (see earlier appendices of this report) as well as substantial changes in our understanding of SBT age and growth (Gunn et al. 1996; Hearn and Polacheck 2003; Polacheck et al. 2003, 2004). The latter is essential for the estimation of the age of release of tagged fish, which is a critical input for many of the quantitative estimation approaches. Also, there is now a better understanding of the need to address some of the key issues with respect to incomplete mixing, tag shedding and reporting rates if such quantitative estimation is to be undertaken. In the current Appendix, the potential use of the data from these earlier tagging experiments for providing quantitative mortality rate estimates within the context of the generalized Brownie-Petersen framework developed in this report (Appendices 5 and 7) are examined.

Release Data

Table 1 provides the number of tags released by cohort and age of release for all tagging experiments prior to 1990[2]. Tag releases were screened with respect to quality as described in Appendix 4. However, releases by all taggers (i.e., fishermen as well as scientists) have been included in this table. In additions releases from troll caught

---

[1] A fraction of the releases are considered unusable for quantitative analyses because of noted concerns at the time of release (e.g. injuries, tag placement, etc. – see Appendix 4).

[2] Note that cohort and age of release are estimated quantities based on the release date and the length of the fish at release.

as well as pole and line caught fish have been included. It is clear from Table 1 that for all of the cohorts spawned since 1968, tagging was not conducted in sufficient sequential years and/or over a wide enough age range of fish to provide substantial numbers of releases at more than a single age for any cohort. Since tagging of the same cohort in multiple years is a basic requirement for the application of the Brownie model[3] or the generalizations developed within this report, the lack of releases for multiple ages means that it is not feasible to apply such estimation models to the data from these cohorts. In contrast, simply based on the number of releases by age, there would appear to be sufficient data to consider application of a Brownie or Brownie-Petersen model to the data from the 1960-68 cohorts. However, there are several issues that need to be considered prior to any such application including: (1) who did the tagging, (2) potential incomplete mixing, (3) tag shedding and (4) tag reporting rates.

Tagger

While all of the juvenile SBT tagging experiments were organized by CSIRO, a large fraction of the tagging prior to 1980 was not actually conducted by CSIRO staff or dedicated "tagging technicians" (e.g. WA Fisheries staff) but was conducted by contracted commercial fishermen, often without any direct, on-board supervision by CSIRO staff. Stanley and Hearn (1999) note that "pre-1990 some taggers would have had no or little experience in tagging, especially commercial fishermen". The actual documentation of who did the tagging is incomplete and for a fraction of the releases, the tagger or his training/affiliation is simply unknown. Most cases where the tagger is unknown were for releases prior to 1966 and during the 1970s (Table 2). It should also be noted that for all the pre-1990 tagging the actual tagger was not included in the original computer records for the tagging data. The tagger was subsequently entered based on an examination of the written tagging data sheets. In some cases, there is uncertainty as to whether the tagger has been identified correctly (see Stanley and Hearn 1999). However, this is not likely to be a substantial problem, particularly with determination of whether tagging was done by a commercial fishermen or not (e.g. most of the fish identified as tagged by commercial fishermen occurred when there was no CSIRO or other tagger aboard).

Concerns have been expressed about the reliability of the tagging done by commercial fishermen (Hearn, personal communications[4]). These concerns relate to (1) the quality of the tagging, (2) whether in fact the tags were released and (3) the reliability of the release information (especially length). Attempts to evaluate these concerns are confounded by the fact that years with substantial numbers of releases by commercial fishermen do not tend to overlap with years of substantial numbers of releases by scientists. Thus, scientist releases dominate the tagging in the early 1960s while fishermen releases dominated tagging in 1967-69 (Table 3). Differences also exist in the month when fish were tagged (Figure 1), which would be expected to affect the number of short-term returns (e.g., higher returns would be expected for fish tagged nearer the beginning of the fishing season). There were also differences in the size range of fish tagged; these changed over time and return rates varied with the size at

---

[3] The Brownie model as originally developed can be applied to tagged data pooled over age when there are not significant differences with age in fishing mortality rates and in recruitment among cohorts - neither of which is applicable in the case of the SBT tagging data.

[4] Dr. W. Hearn, CSIRO Marine Research, PO Box 20, North Beach, Western Australia 6020, Australia

release (Figure 2, 3 and 4). Thus, for example, overall 28% of the fish that were less than 80cm when tagged were recaptured compared to only 9% for those greater or equal to 80cm. In addition, catches of SBT increased substantially during the 1960s and early 1970s (Figure 6), which means return rates would have also been expected to have been increasing during this period. The evaluation of fishermen releases in the 1967-69 period is further confounded by the fact that most of these releases were by a single vessel in an area within any given year, in particular in 1968 and 1969.

The concern about the quality of tagging relates to the fact that the commercial fishermen were not trained in either tagging or general scientific/biological techniques. As such, there is a possibility that tag shedding or tagging mortality could be abnormally high and that the number of returns from these releases would be significantly too low. If this were the case, then mortality rates estimates from such releases would be substantially biased unless data were available that would allow these effects to be estimated. However, comparison of the return rates from fishermen and scientist released fish do not support this concern. In fact, overall return rates from fishermen tagged fish exceed scientist tagged fish by nearly a factor of five (Table 3). However, the interpretation of these differences is confounded by the factors discussed above. For example, for the releases in the early 1960s (1959-66), return rates are greater for scientist tagged fish when the data are broken down into area and size of release (Table 3). In contrast for releases between 1967-69, return rates from releases off South Australia (SA) are nearly equivalent for scientist and fishermen releases, while for releases off New South Wales (NSW) return rates for fishermen releases exceed those from scientist releases by over a factor of five when releases less than 80cm are considered (Table 3). Note that almost all of the fishermen releases were less than 80cm (Figure 3). Thus, the return rates from the 1959-1966 period lend some support to concerns about the relative quality/reliability of fishermen releases, while those from the 1967-69 period do not.

The return rates from the 1967-69 fishermen releases are very high and are among the highest seen for any of the large batches of SBT tags ever released. These very high return rates have raised concerns about whether in fact all of the tags were actually released. Instead some tags may never have been released and may have simply been returned and reported as having been retrieved from recaptured fish. While such concerns are hard to evaluate, there is no direct evidence which supports such concerns. Moreover, there are a number of factors that would suggest that the release and return data are valid. These include:
1. The returns came from a large number of fishermen/vessels (Table 4) and reported recapture dates were spread out over time (Figure 5);
2. The reported size range of fish, both when released and recaptured, do not suggest any substantive inconsistencies (Figure 4);
3. Catches in NSW doubled between 1967 and 1968 and again between 1968 and 1969, with 1969 having the largest catches in number ever recorded for NSW. Catches remained near this peak level in 1970 and declined by 19% in 1971 (Figure 6). As such, large increases in the return rates from releases in 1968 and 1969 would be expected;
4. CSIRO internal correspondence at the time the contracts were let stated that the fishermen undertaking the tagging were highly regarded in terms of tagging skill and reliability.

As such, other than the high recovery rate, nothing we have been able to ascertain, either internal to the data or from the limited accounts at the time, suggest concerns with these 1967-69 releases.

Incomplete mixing

Table 5 compares the number of tags returned by recapture location for tags released off WA (Western Australia), SA and NSW in the 1960s for recaptures that were at liberty for over 270 days. Tables 6-8 provide similar comparisons except broken down into tags released by scientists, fishermen or unknown taggers. Evident in these tables is a very strong tendency for fish tagged in NSW to be recaptured almost exclusively in NSW (e.g. 93% overall), while fish tagged in WA or SA tended to be more evenly distributed between SA and NSW (e.g. 40% of the SA releases were recaptured in SA and 28% in NSW). In addition, a significantly lower percent of the recaptures from NSW releases came from longline vessels compared to those from WA or SA (4% compare to 22% and 32% respectively). The spatial location of the returns from longline vessels also differed for releases from NSW and those from WA and SA (Table 9). Thus, a high proportion of the NSW longline returns were from east of 145°E longitude (i.e. the Tasman Sea/New Zealand area) with very few returns from the Indian Ocean. In contrast the longline returns for WA and SA releases were more evenly spread out from the Indian Ocean to the Tasman Sea (Table 9).

If the tagged fish were completely mixed, then the proportion of returns from the different locations for the different release locations should be the same for tagged fish of similar ages released in the same time period. The extent to which these proportions differ would constitute a measure of incomplete mixing. Simple interpretation of the results in Tables 5-8 is confounded by the fact that the releases took place over a decade, while the number and spatial distribution of both releases and commercial catches vary over time. Thus, there are differences in the relative patterns of return locations for tags released prior to 1966 and those released post 1966 (Table 10). This is not surprising as the catches off NSW were increasing in the late 1960s (Figure 6). The differences in the temporal distribution of releases for scientists and fishermen (Table 3) may also explain the large differences in the proportion of their WA and SA releases that were recaptured in NSW. However, any comparison is confounded to some extent by the small number of scientist released tags in the late 1960s.

Although the temporal changes in the distribution of releases and commercial catches complicates the interpretation of the spatial pattern of the return data, the differential pattern of return location by release location provides strong evidences of incomplete mixing of fish among the release locations. In particular, it suggests that once fish entered the NSW fishing area they tended to remain in this area. They did not move or return to the SA region as juveniles. When they moved into more off-shore waters and became vulnerable to longline gear, they remained primarily in the Tasman Sea/New Zealand region. This contrasts strongly with the recaptures from releases from WA and SA, which indicate a much wider mixing among SBT fishing areas. Thus, there was substantial movement of fish from WA and SA into the waters off NSW, but also a large proportion that remained or returned to the SA region. When fish from these areas moved into off-shore waters, they spread out into the Indian Ocean as well as the Tasman Sea.

Tag Shedding

Hampton (1989) and Hampton and Kirkwood (1989) analysed tag shedding rates for the pre-1990 SBT tagging experiments. These analyses were performed on releases pooled over a number of years and taggers but with separate analyses according to the area of release and for fishermen contracted releases. For the 1960s releases, the estimates of shedding rates obtained were reasonably high (Figure 7). Thus, in all cases the estimated probability of a double-tagged fish losing both its tags within three years always exceeded 17% and was as high as 37%. It should be noted that these analyses did not consider individual differences among taggers within the groups of releases that were analysed. As discussed in Appendix 14, this will underestimate the actual shedding rates if in fact heterogeneity in shedding rates exists among taggers. In addition, in the initial few years (1959-1963) of SBT tagging only a small fraction of the releases were double tagged (i.e. 0% prior to 1962 and 22% in 1963). For these single-tagged fish, the probability of recovering a tagged fish without any tag is much higher than for fish that were double tagged.

Reporting Rates

For tags released in the 1960s, no data exists that can be used to directly estimate reporting rates. Thus, no tag seeding experiments were undertaken for the surface fishery and no observers were present in any components of the SBT fishery. Any analysis of the 1960s tag return data would be conditional on assumptions about the tag reporting rates in these years. It should be noted that there was substantial publicity about these tagging experiments. Moreover, they were not undertaken to estimate fishing mortality rates because there was little concern at the time about overfishing of the stock and no limits on catch existed or were being contemplated. As such, one of main factors that can make fishermen reluctant to return tags did not exist (i.e. the perception that returning tags will contribute to reduction in quotas). All of this combined with high tag return rates from NSW in the late 1960s (if these were not an artefact) would suggest that reporting rates may have been relatively high during the 1960s period (at least in the latter years). Nevertheless, it is unlikely that reporting rates were 100%.

Conclusions

The examination of the historic (pre-1990) SBT tagging data presented here suggests that there are substantial problems with using these data for quantitative estimation of mortality rates and population sizes. The data are not directly amendable to the Brownie-Peterson estimation framework developed within this report. For the post 1960s releases, there was simply not enough multiple tagging of the same cohorts at consecutive ages to permit the application of a Brownie-type estimator. For the 1960s releases, potentially sufficient multiple tagging of some cohorts was conducted to allow for a Brownie-type estimator to be used. However, the differential return rates for tags released in different areas (e.g. Tables 5-9) provide strong evidence for substantial incomplete mixing among releases from NSW with those from WA and SA. This indicates that the non-spatial models of Appendices 7 and 9 would not be appropriate and that a spatially explicit estimation model would be required to obtain reliable estimates. The spatially explicit models developed for use with the 1990s SBT tagging experiments (see Appendices 11 and 16) cannot be used for these 1960s data because these models make no allowance for a NSW juvenile component of the stock since this component disappeared in the early 1980s (Caton, 1991). The fully generic

spatial model of Appendix 10 is also not directly applicable because of the lack of releases and fishing in all areas (particularly the longline areas) and the differing seasonality among the fisheries. A spatially explicit model that incorporated hypotheses with respect to the possible movement dynamics along the lines of those in Appendix 11 but tailored for the stock and fisheries of the 1960s could be developed. However, development of such a model was beyond the scope of the current study, particularly given the lack of data for estimating reporting rates, the high shedding rates and concerns associated with the fishermen releases (e.g. the results would be highly dependent upon assumptions about reporting rates and whether or not fishermen releases were included). Nevertheless, further analyses of the 1960s tagging experiments could be informative. In particular, such analyses might be extremely informative with respect to the NSW surface component of the stock and its disappearance in the early 1980s. They might also provide additional estimates of juvenile natural mortality that would be useful for comparison with those from the 1990s experiments.

Even without a quantitative estimation model, the examination of the return and recapture data from the 1960s presented here suggests that historically a large degree of spatial structuring and spatial heterogeneity existed among the juvenile SBT found within Australian waters. The high return rates from NSW and high proportion of NSW returns from NSW releases combined with increased catches in the late 1960s suggests the possibility of a substantial degree of ecological structuring and separation among juvenile SBT between NSW and other areas (WA and SA in particular). This further suggests a high degree of vulnerability to over-exploitation and localized depletion for the NSW component, which could have implications for the rebuilding of the SBT stock.

Literature Cited

Caton, A.E. 1991. Review of aspects of southern bluefin tuna: biology, population and fisheries. Inter-Amer. Trop. Tuna Comm., Spec. Rep. 7: 181-357.

Gunn, J.S., N.P. Clear, A.J. Rees, C.A. Stanley, J.H. Farley, and T.I. Carter. 1996. The direct estimation of age and growth in southern bluefin tuna. FRDC Final Report 92/42.

Hampton, J. 1989. Population dynamics, stock assessment and fishery management of the southern bluefin tuna (*Thunnus maccoyii*). Ph.D. Thesis. University of New South Wales, Kensington Australia. 274pp.

Hampton, J. and G.P. Kirkwood. 1989. Tag shedding by southern bluefin tuna *Thunnus maccoyii*. Fish. Bull. 88: 313-321.

Hampton, J. 1991. Estimation of southern bluefin tuna *Thunnus maccoyii* growth parameters from tagging data, using von Bertalanffy models incorporating individual variation. Fish. Bull. 89: 577-590.

Hearn, W.S. 1986. Mathematical methods for evaluating marine fisheries. Ph.D. Thesis. University of New South Wales, Kensington Australia. 195pp.

Hearn, W.S., R.L. Sandland and J. Hampton. 1987. Robust estimation of the natural mortality rate in a complete tagging experiment with variable fishing intensity. J. Cons. Int. Explor. Mer. 43:107-117.

Hearn, W.S., G.M. Leigh and R.J.H. Beverton. 1991. An examination of a tag-shedding assumption with application to southern bluefin tuna. ICES Journal of Marine Science 48: 41-51.

Hearn, W.S. and T. Polacheck. 2003. Estimating long-term growth-rate changes of southern bluefin tuna (*Thunnus maccoyii*) from two periods of tag-return data. Fish. Bull. 101: 58-74.

Hynd, J.S. 1969. New evidence on southern bluefin tuna stocks and migrations. Australian Fisheries 28:26-30.

Lucas, C. 1974. Working paper on southern bluefin tuna population dynamics. ICCAT on Tuna. Nantes, WTFD-Nantes/74/10.

Murphy, G.I. 1977. New understanding of southern bluefin tuna. Australian Fisheries 36:2-6.

Murphy, G.I. and J. Majkowski. 1981. Sate of the southern bluefin tuna population L: fully exploited. Australian Fisheries 40:20-29.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1996. Estimation of natural and fishing mortality for juvenile southern bluefin tuna based on multi-year tagging of cohorts. CCSBT/SC/96/11.

Polacheck, T., W. Hearn, C. Millar, W. Whitelaw and C. Stanley. 1997. Updated estimates of mortality rates for juvenile SBT from multi-year tagging of cohorts 1997. CCSBT/SC/9707/26.

Polacheck, T., W. Hearn, C. Millar and C. Stanley. 1998. Updated estimates of mortality rates for juvenile SBT from multi-year tagging cohorts. 1998. CCSBT-SC/9808/20.

Polacheck, T., G.M. Laslett, and J.P. Eveson. 2003. An integrated analysis of the growth rates of southern bluefin tuna for use in estimating the catch at age matrix in the stock assessment. Final Report. FRDC Project 1999/104. ISBN 1 876996 38 2.

Polacheck, T., J.P. Eveson, and G.M. Laslett. 2004. Increase in growth rates of southern bluefin tuna (*Thunnus maccoyii*) over four decades: 1960 to 2000. Can. J. Fish. Aquat. Sci. 61: 307-322.

Stanley, C. and Hearn, W.S. 1999. Southern bluefin tuna tag data base: contents, knowledge and pitfalls. CSIRO Marine Research Internal Report. April 1999. Hobart, Australia.

Table 1: Number of tags released by cohort and estimated age for tagging experiments conducted between 1960 and 1984.

Cohort	Age of release				
	0	1	2	3	4
1957	0	0	23	28	4
1958	0	34	72	30	1
1959	0	1	57	3	41
1960	0	169	886	884	151
1961	113	4456	465	611	11
1962	18	5954	2365	235	33
1963	25	5054	825	944	72
1964	276	1937	522	82	0
1965	75	857	293	18	0
1966	196	8150	198	19	0
1967	471	4273	242	9	0
1968	0	2870	494	0	0
1969	1	1376	0	0	0
1970	266	14	0	0	0
1971	0	137	3	0	0
1972	16	864	3	0	0
1973	221	382	1	0	0
1974	359	781	0	0	0
1975	103	3	49	0	0
1976	0	1361	0	0	0
1977	0	30	0	0	0
1978	1000	0	72	0	0
1979	0	481	0	0	0
1980	0	0	0	7	14
1981	0	0	461	62	0
1982	0	5528	712	0	0
1983	871	2333	0	0	0

Table 2: Number and percent of tag releases in each year where the tagger is unknown for tags released prior to 1990.

Year	Total number releases	Number releases by unknown taggers	Percent releases by unknown taggers
1959	57	57	100.0
1960	104	104	100.0
1961	375	227	60.5
1962	5369	351	6.5
1963	7408	5732	77.4
1964	8502	5229	61.5
1965	3099	2162	69.8
1966	2593	22	0.8
1967	9145	82	0.9
1968	4552	0	0.0
1969	3150	1	0.0
1970	2154	1328	61.7
1971	14	14	100.0
1972	154	154	100.0
1973	1088	1088	100.0
1974	745	745	100.0
1975	885	885	100.0
1976	3	3	100.0
1977	1414	506	35.8
1978	1037	1037	100.0
1980	555	1	0.2
1983	6885	0	0.0
1984	3135	0	0.0

Table 3: Number of SBT tagged, number of tagged fish returned, and return rate (number returned divided by number released) by tagger type for various partitions of the tag data from the 1960s. (SA = South Australia, NSW = New South Wales, WA = Western Australia.)

Release Years	Release Length	Days at Liberty	Release Location		Fisher	Scientist	Unknown tagger
1959-1969	All	All	All	Number released	17616	12771	13967
				Number returned	5172	786	603
				Return rate	0.294	0.062	0.043
		<270	All	Number returned	3297	307	173
				Return rate	0.187	0.024	0.012
		>270	All	Number returned	1875	479	430
				Return rate	0.106	0.038	0.031
1959-1966	All	All	All	Number released	1338	12285	13884
				Number returned	134	706	592
				Return rate	0.100	0.057	0.043
			SA	Number released	578	4851	462
				Number returned	16	255	37
				Return rate	0.028	0.053	0.080
			NSW	Number released	743	1592	1319
				Number returned	115	285	112
				Return rate	0.155	0.179	0.085
	<80cm	All	SA	Number released	2932	5079	110
				Number returned	9	87	7
				Return rate	0.003	0.017	0.064
			NSW	Number released	612	1149	917
				Number returned	97	263	91
				Return rate	0.158	0.229	0.099
			WA	Number released	0	5787	11994
				Number returned	0	157	426
				Return rate	-	0.027	0.036
1967-1969	All	All	All	Number released	16278	486	83
				Number returned	5038	80	11
				Return rate	0.309	0.165	0.133
	<80cm	All	All	Number released	16072	292	78
				Number returned	4958	69	10
				Return rate	0.308	0.236	0.128
			NSW	Number released	8414	117	73
				Number returned	3899	10	9
				Return rate	0.463	0.085	0.123
			SA	Number released	2645	175	1
				Number returned	888	59	1
				Return rate	0.336	0.337	1.000

Table 4: The number of vessels that returned tags from releases by scientist, fishermen or unknown tagger.

Tagger	Release Years 1959-1966	1967-1969
Scientist	175	41
Fishermen	47	170
Unknown	148	10

Table 5: The number of surface tag returns by recapture location (SA, NSW, TAS) and the number of longline tag returns for tags released in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania)

Release Location	Surface returns SA	NSW	TAS	Longline returns
WA	278	166	1	128
SA	179	128	0	144
NSW	47	1622	1	77
TAS	3	1	0	3

Table 6: The number of surface tag returns by recapture location (SA, NSW, TAS) and the number of longline tag returns for tags released by scientists in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania)

Release Location	Surface returns SA	NSW	TAS	Longline returns
WA	1	0	0	1
SA	95	13	0	103
NSW	15	110	0	23
TAS	0	0	0	2

Table 7: The number of surface tag returns by recapture location (SA, NSW, TAS) and the number of longline tag returns for tags released by fishermen in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania)

Release Location	Surface returns SA	NSW	TAS	Longline returns
WA	56	43	0	34
SA	63	112	0	20
NSW	28	1464	1	49
TAS	0	0	0	1

Table 8: The number of surface tag returns by recapture location (SA, NSW, TAS) and the number of longline tag returns for tags released by unknown taggers in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania)

Release Location	Surface returns			Longline returns
	SA	NSW	TAS	
WA	160	87	1	84
SA	14	3	0	18
NSW	4	48	0	5
TAS	3	1	0	0

Table 9: The number of longline tag returns by longitude (°E) for tags released in WA, SA, NSW and TAS in the 1960s, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania)

Release Location	Recapture Longitude		
	<115	115-145	>145
WA	66	27	35
SA	61	25	58
NSW	12	7	58
TAS	0	0	3

Table 10: Comparison of the number of tag returns by recapture location (SA, NSW, TAS, or high seas) for tags released in WA, SA, NSW and TAS between 1959-66 and between 1967-69, using only recaptures that were at liberty more than 270 days. (SA = South Australia, NSW = New South Wales, WA = Western Australia, TAS = Tasmania. – these are all surface recaptures)

Release Years	Release Location	Recapture Location			
		SA	NSW	TAS	High Seas
1959-66	WA	222	123	1	94
	SA	116	17	0	124
	NSW	20	182	0	34
	TAS	3	1	0	2
1967-69	WA	56	43	0	34
	SA	63	111	0	20
	NSW	27	1440	1	43
	TAS	0	0	0	1

Figure 1: The number of tags released in the 1960s by month by scientist, fishermen and unknown tagger.
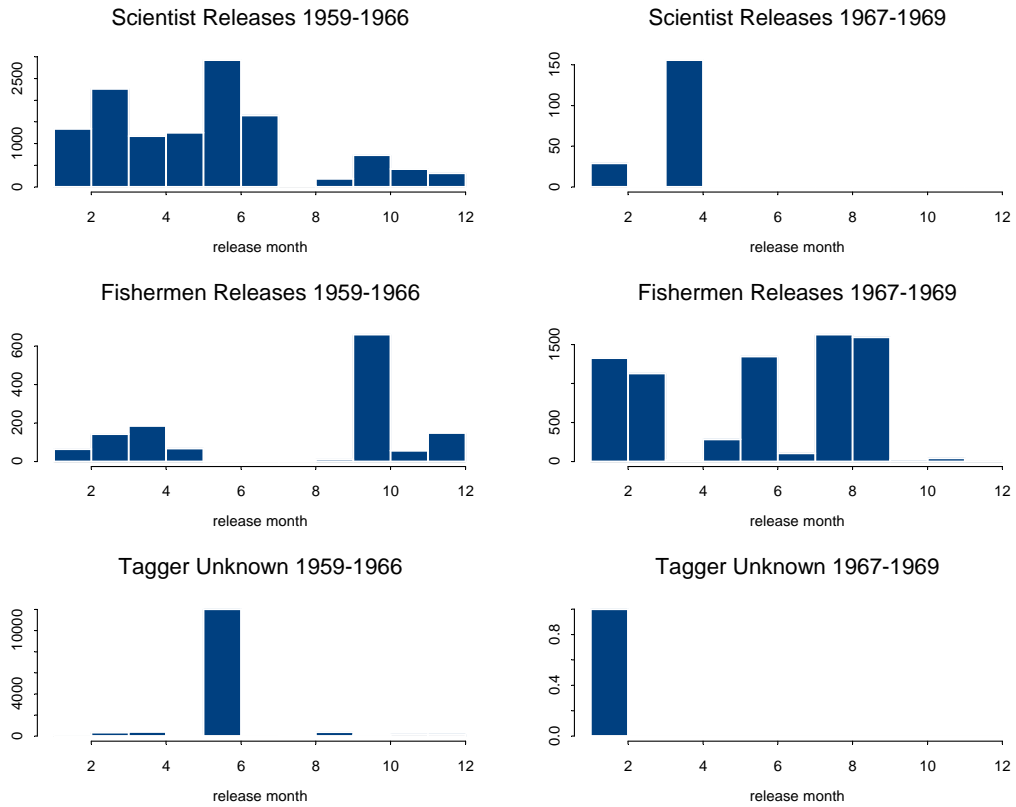
Figure 2: The size distribution of fish tagged in the 1960s by scientist, fishermen and unknown tagger and the release size distribution for those fish that were subsequently recaptured.
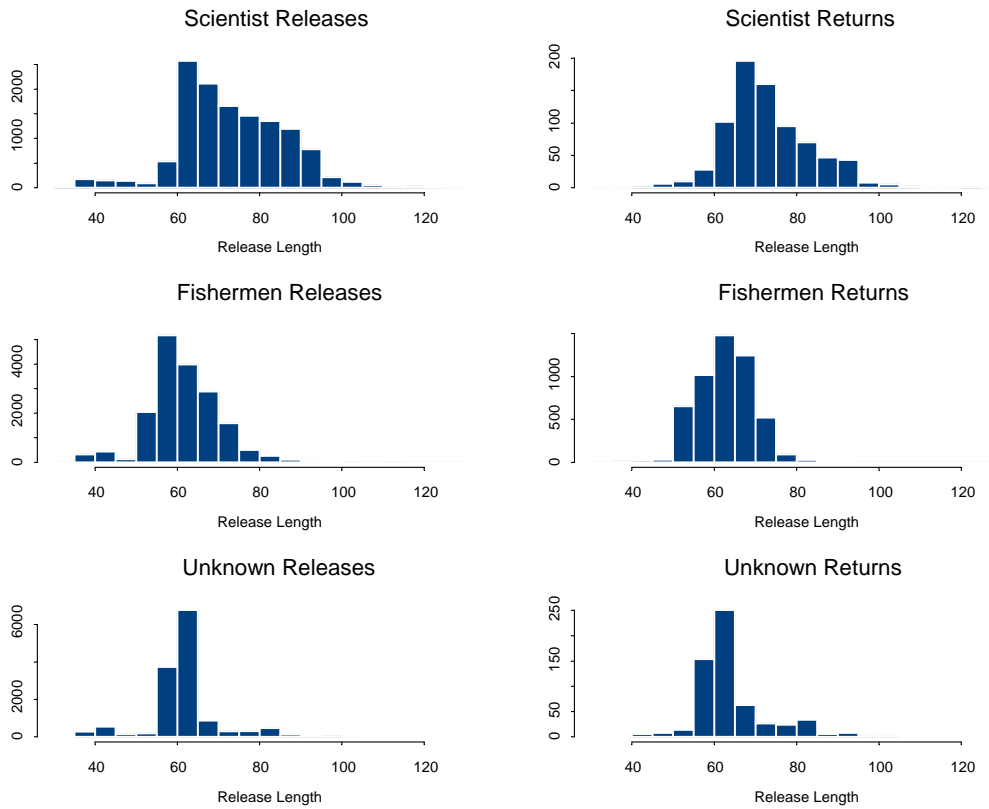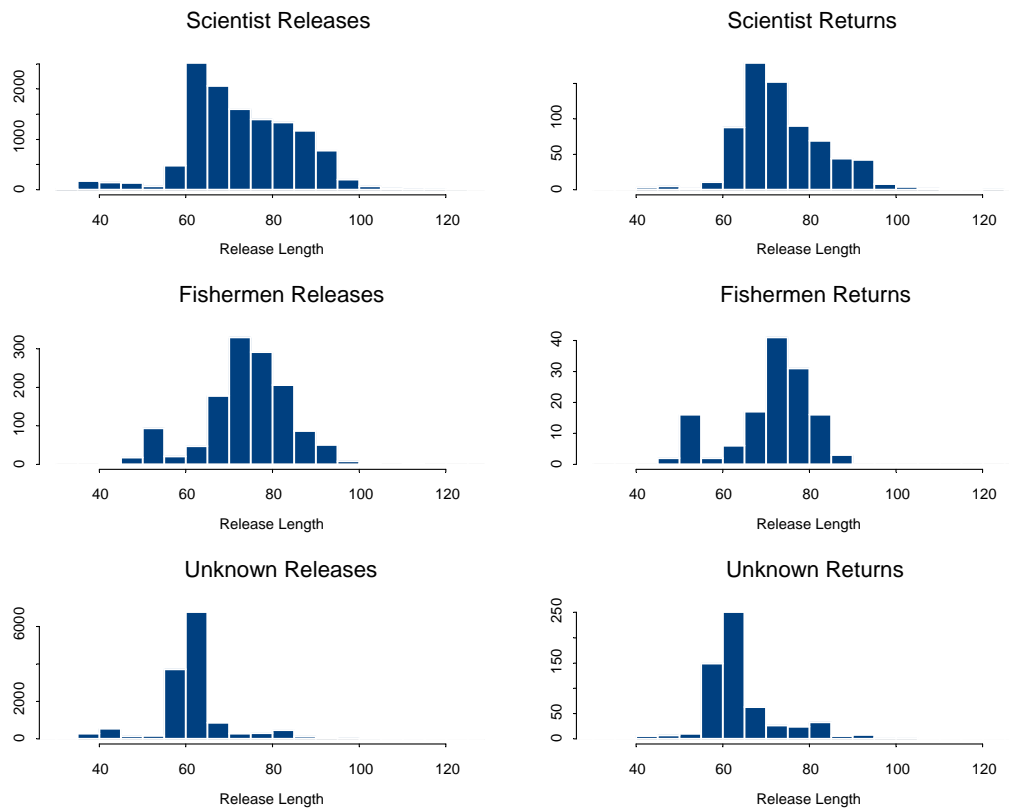
Figure 3: The size distribution of fish tagged in the period from 1960-1966 by scientist, fishermen and unknown tagger and the release size distribution for those fish that were subsequently recaptured.
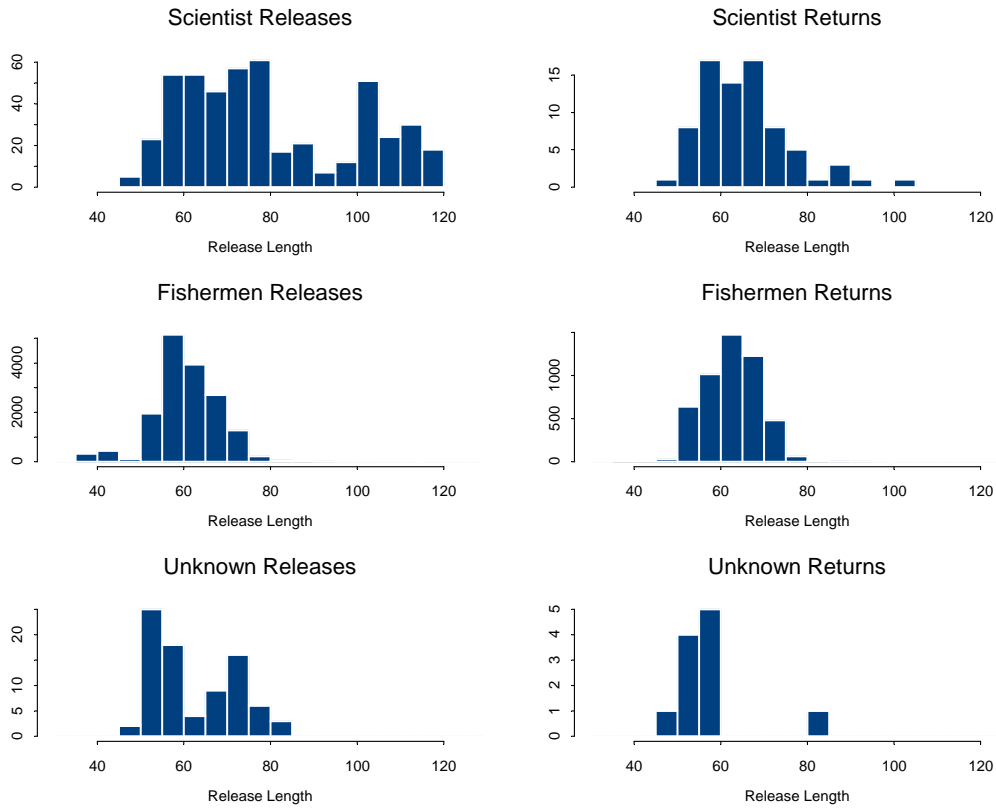
Figure 4: The size distribution of fish tagged in the period from 1967-1969 by scientist, fishermen and unknown tagger and the release size distribution for those fish that were subsequently recaptured.

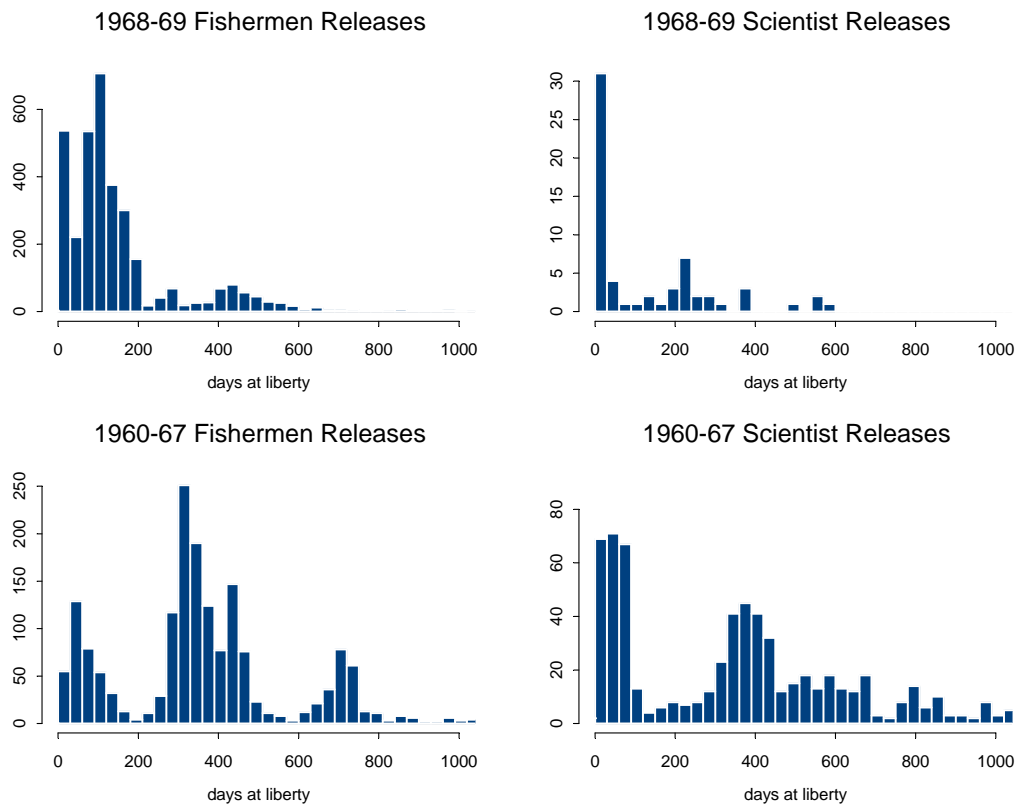Figure 5: Comparison of the days at liberty for tags recovered from releases by fishermen and scientists.

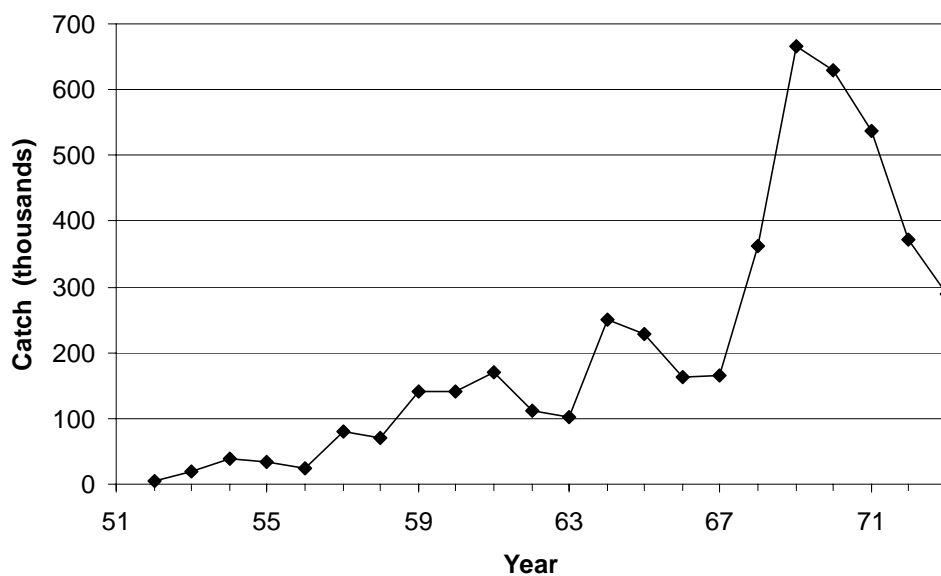Figure 6: The number of SBT caught in New South Wales by quota year (from Caton 1991).

Figure 7: Estimated probability of a double-tagged fish having lost both of its tags for SBT tagging experiments conducted in the 1960s (estimates taken from Hampton and Kirkwood 1989). Note experiment 1 corresponds to releases in NSW by CSIRO in 1963-70; experiment 2 corresponds to releases in NSW by fishermen in 1963-70; experiment 3 corresponds to releases in SA by CSIRO in 1964-69, and experiment 5 corresponds to releases in WA by CSIRO in 1963-67.