Impact of spatial tagging rates and mark-recapture likelihood structures on key stock assessment parameters for Macquarie Island toothfish

R. M. Hillary & J. Day,

CSIRO Oceans & Atmosphere,

Castray Esplanade,

Hobart,

TAS 7004,

Australia.

Abstract

The mark-recapture program underpins the estimation of absolute abundance and migration rates in the Macquarie Island toothfish stock assessment. The current spatial configuration of the assessment is with two regions - northern and southern - with movement between them and variable fishing pressure over time. In terms of tagging rates (per tonne caught), conditional on an overall rate at 2 tags per tonne, we explored three area-specific options. In terms of the mark-recapture likelihood functions, we explored the current stock assessment formulation (in the Stock Synthesis package), and the so-called Brownie formulation. The focus was on the interaction of these alternatives (tagging rate and likelihood function) with a range of plausible true spatial mean recruitment and migration rates, in terms of the uncertainty in resultant estimates of spawning abundance and the key spatial parameters. An information theoretic approach is employed which permits us to avoid the need for either simulation or estimation in producing the uncertainty estimates, using previous work around bias levels for these types of estimators to deal with the bias/variance question. In terms of spatial tagging rates, even when faced with spatially asymmetric migration and fishing pressure, a more even tagging rate seemed to result in more precise estimates of overall abundance and migration.In terms of the expected precision of overall spawning abundance, the current and Brownie likelihoods were broadly similar across all the different spatial tagging rates. In terms of the precision of the migration estimates, the Brownie likelihood clearly outperformed the current likelihood - particularly in terms of avoiding certain scenarios resulting in very high estimates of uncertainty for the current likelihood.

Introduction

We explore how factors we can potentially control (spatial tagging rates per tonne and overall spatial catch splits, effective length sample sizes) interact with things we cannot control (spatial recruitment and migration rates) when we estimate abundance, migration and selectivity in the Macquarie Island stock assessment model.

Methods

The modelling ideas behind the work can be broadly split into three main sections: one dealing with the conditioning of the operating model used as the basis for the population dynamics; another specifying each of the two mark-recapture likelihood functions; and the final dealing with the statistical principles used to construct the uncertainty estimates, given the underlying "true" population dynamics and the specifics of the sampling regimes for the fishery (in terms of both tagging rates and length sampling).

Spatial population dynamics

The Operating Model (OM) is both age and spatially structured. An equilibrium structure is assumed with an average total recruitment level, $\bar{R}$, which is distributed across regions $r = 1, \cdots, R$: $N_{1,r} = \eta_r \bar{R}$, where $\sum \eta_r = 1$. The stock-recruit relationship (assumed to be Beverton-Holt) is factored in at estimation time and we detail that process later on when discussing the conditioning of the OM. For ages $a = 1, \cdots, A - 1$ we have the following:

$$N_{a+1,r} = \sum_{s=1}^{R} \pi_{s,r} N_{a,s} e^{-M} \left(1 - h_s s_{a,s}\right), \tag{1}$$

$$C_{a,r} = N_{a,r} h_r s_{a,r}, \tag{2}$$

where $N_{a,r}$ are the numbers-at-age; $C_{a,r}$ is the catch numbers-at-age; M is the rate of natural mortality; h_r and $s_{a,r}$ the spatial harvest rate and selectivity-at-age, respectively; and $\pi_{s,r}$ the probability that an animal currently in area s will move into area r in the following year (and $\sum_s \pi_{s,r} = 1$). No movement is assumed in the plus group, A, and it is dealt with in the usual way given the equilibrium assumption. Spawning stock biomass and total catch biomass (both region-specific and in total) are intuitively defined:

$$S_r = \sum_a N_{a,r} w_a m_a, \tag{3}$$

$$C_r = \sum_a C_{a,r} w_a, \tag{4}$$

$$S = \sum_s S_s, \tag{5}$$

$$C = \sum_s C_s, \tag{6}$$

where w_a and m_a are the weight and maturity-at-age vectors, respectively. The dynamics of the tagged population are basically the same as those for the actual population with two minor tweaks:

1. There is an additional tag-shedding rate, ν, (if required) and this is "added" to the natural mortality term: $\exp\left(-M - \nu^2\right)$ (squared because we double tag)

2. There is an initial tag survival probability, π^{tag}, that applies only in the year of tagging, and is just used to instantly adjust the total number of tags released into the population.

The spatial split in total catch biomass will be a useful variable - both for conditioning the OM to a given scenario and in defined spatial tag releases - and is obviously defined: $\delta_r = C_r/C$. A total tag release rate-per-tonne, ψ, is assumed to be fixed with some freedom in the spatial tagging rates-per-tonne, ψ_r, where $\psi = \sum_r \delta_r \psi_r$. The tag release numbers-by-region are then simply given by $T_r = \psi_r C_r$.

Another key factor in the conditioning of the OM is spawner biomass depletion, γ (both spatially and in total). The stock assessment assumed that at a spatially aggregated spawning biomass produces the mean recruitment numbers, and we take the same approach. An efficient short-cut is used to calculate γ, given the key parameters $\boldsymbol{\theta} = \{\bar{R}, \eta_r, \pi_{s,r}\}$. This is done first via the total spawner biomass-per-recruit: ρ. For given population parameters, harvest rates and selectivity functions ρ is calculated; it is then calculated for zero harvest rates (no fishing). The ratio of the fished-to-unfished spawner biomass-per-recruit, $\tilde{\gamma}$, is then calculated. The *actual* depletion level (now accounting for the stock-recruit relationship) is calculated as follows:

$$\gamma = \frac{4h\tilde{\gamma} + h - 1}{5h - 1}, \tag{7}$$

where h is the steepness of the stock-recruit relationship (a steepness of 1 implies $\gamma \equiv \tilde{\gamma}$). Spatial SSB depletion levels are then simple to define:

$$\gamma_r = \tilde{\gamma}_r \left(\frac{\gamma}{\tilde{\gamma}}\right). \tag{8}$$

The key release covariates are age and region of release; for recaptures, it is age and region of capture. We first intend to mimic the tag recapture likelihood as it is currently implemented in SS3 Methot and Wetzel (2013), which is used for the current assessment. It takes a two-step account of the tag recapture process as follows:

- For a given release region and age, subsequent recaptures are aggregated across all regions and the assumed likelihood is negative binomial (with associated over-dispersion parameter). For our purposes we assume the base likelihood to be an over-dispersed binomial (for reasons made clear later on). This likelihood will be the dominant source of information on absolute abundance.

- Spatial recaptures are treated as multinomial, with the proportions of recaptures across the various regions the multinomial probability, and the sample size is the total number of recaptures across the regions. This likelihood will be the dominant source of information on migration rates and spatial relative recruitment.

4

The final aspect we consider is the estimation of selectivity (primarily via the length frequency information of the catches by region). We would be under-estimating the uncertainty in both abundance and migration if we simply assumed we knew selectivity without error, so including it makes sense. We first outline the length-based selectivity model, then how this is translated to age, then how the predicted length frequencies are calculated. Region-specific selectivity is assumed to be logistic in nature:

$$s_{l,r} = \left(1 + 19^{-(l-\lambda_{r,50})/(\lambda_{r,95}-\lambda_{r,50})}\right)^{-1} \tag{9}$$

where $\lambda_{r,50}$ and $\lambda_{r,95}$ parameters are the lengths at 50% and 95% selection, respectively. The selectivity-at-age is derived from the selectivity-at-length via the distribution of length-at-age, $\pi_{l|a}$:

$$s_{a,r} = \int s_{l,r} \pi_{l|a} \mathrm{d}l. \tag{10}$$

Tag recapture likelihood models

For the tag likelihood utilised in the current stock assessment, the probability of recapturing a tag released in area r and at age a across **all** the spatial regions τ years after release (with $\tau \geq 1$ given the current non-mixing period of 1 season) is given by the following:

$$p_{a+\tau,|\,r,a}^{\mathrm{rec}} = \xi^{\mathrm{rep}} \sum_{s=1}^{R} \omega_{s,a+\tau\,|\,r,a} h_s s_{a+\tau,s}, \tag{11}$$

where ξ^{rep} is the reporting rate, and $\omega_{s,a+\tau\,|\,r,a}$ is the proportion of tagged fish in area s at age $a + \tau$, given their release in area r at age a. The total number of recaptures is then a binomial variable, with underlying probability as defined in (11) and sample size $T_{a+\tau\,|\,r,a}$, which is the total number of tags still both alive and possessing at least one tag at age $a + \tau$, given their release in area r at age a. The expected number of recaptures is as follows:

$$R_{a+\tau\,|\,r,a} = T_{a+\tau\,|\,r,a}\, p_{a+\tau,\,|\,r,a}^{\mathrm{rec}}. \tag{12}$$

The spatial aspect of the mark-recapture likelihood relates to the proportions of recaptures of a given release event across all regions, via a multinomial distribution. The spatial proportions of recaptures, conditional on a given release age a and region r, are given by the following:

$$\zeta_{s,a+\tau\,|\,r,a} = \frac{p_{a+\tau,s|\,r,a}^{\mathrm{rec}}}{\sum_{k=1}^{R} p_{a+\tau,k|\,r,a}^{\mathrm{rec}}}, \tag{13}$$

where the associated sample size for this multinomial distribution is the total number of recaptures across all spatial regions given in (12). These two likelihood functions, relating first to total recaptures and then to the spatial distribution thereof, fully describe how these data are treated in the current stock assessment model.

The original Brownie model (Brownie *et al.*, 1985) was age-structured and non-spatial focussing on following the recapture histories of animals tagged in a given cohort and recaptured over a

number of subsequent years. Since then it has been extended to spatial (Eveson *et al.*, 2012) and length-structured (Hillary & Eveson, 2015) mark-recapture models. Its utility in permitting the joint estimation of natural and fishing mortality is well known (Eveson *et al.*, 2012; Hillary & Eveson, 2015), but in this context it permits us to encapsulate all the key information (exploitation rate, natural mortality, and migration) in a single likelihood function. The release covariates are the area and age of release; the key recapture covariates are the area and age of recapture. The main three processes that need to be characterised are: (i) the probability of a tagged fish surviving given τ years-at-liberty; (ii) the probability that the surviving tagged fish are in the area of recapture, s; and (iii) the probability that those surviving tagged fish in area s are recaptured. The probability of a fish tagged at age a and relased in area r surviving for τ years-at-liberty is given by the following:

$$\pi^{\text{surv}}_{a+\tau|r,a} = \pi^{\text{tag}} \exp(-(\tau-1)(M+\nu^2)) \prod_{i=1}^{\tau} \left(\sum_{j=1}^{R} \omega_{j,a+i-1|r,a}(1 - h_j s_{i,j}) \right), \tag{14}$$

and the probability of that tag being recaptured in region s at age $a + \tau$ is given by

$$p^{\text{rec}}_{s,a+\tau|r,a} = \xi^{\text{rep}} \pi^{\text{surv}}_{a+\tau|r,a} \omega_{s,a+\tau|r,a} h_s s_{a+\tau,s}. \tag{15}$$

It is worth contrasting the slightly different features of (15) and the spatially averaged recapture probability in (11). For the more Petersen-like formulation in (11) it relates only to the year of recapture, in terms of the spatial distribution of tags and the spatial probability of recapture. For the Brownie formulation in (15) we see that the mortality and migratory history are explicitly included (via the survival probability) with the region-specific recapture probabilities in the year of recapture. To finish construction of the multinomial likelihood that underpins the Brownie formulation we must sum over all the possible years and regions of recapture, given the release age and region. This yields the probability of *never* recapturing a tag, with sample size the number of tag releases minus the total number of recaptures. This ensures all our probabilities sum to one and the multinomial likelihood is properly defined.

To incorporate the uncertainty in the estimated selectivity parameters, via the length frequency data, we assume a multinomial distribution for these data (as is currently done in the assessment). The predicted spatial catch-at-age (2) is converted to a predicted distribution of catch-at-length as follows:

$$C_{r,l} = \sum_a C_{r,a} \pi_{l|a}, \tag{16}$$

$$p_{r,l} = \frac{C_{r,l}}{\sum_k C_{r,k}}. \tag{17}$$

The term $p_{r,l}$ define the length frequency in region r, and for the likelihood model we assume a given effective sample size n^{eff}

Construction of the uncertainty estimates

Traditionally, analyses of this kind first simulate the data, then back-estimate the key parameters for the given models used to simulate them. If workable, this can yield both the potential bias

and parameter uncertainty properties for the pre-agreed data collection regime. This can be both difficult to do in practice (getting models to both converge and obtain uncertainty estimates for all simulated data sets) and not always necessary if previous work has explored the bias properties associated with these types of models. In this case, the predominant issue is obtaining uncertainty estimates, if either the bias properties are known (and can be accommodated), or the models are essentially unbiased for the kinds of sampling regimes under consideration.

An idea originally proposed for exploring the uncertainty of adult abundance and mortality information via close-kin mark-recapture (CKMR) was outlined in Bravington *et al.* (2016). These models generally have a Bernoulli/binomial likelihood function as they relate to pairwise comparisons of animals and whether they are related in a specific way or not - e.g. parent-offspring. The general idea is to construct the Hessian matrix for the key estimable parameters (which defines the approximate parameter covariance matrix) via an approximation to the Fisher information (FI) at the base information level for each data source. For the total tag recaptures these are region and age at release/recapture; for the spatial tag recaptures these are both region and age at release *and* recapture; for the length frequency data this is the region and length at capture.

The Fisher information is defined in two ways. The first is in terms of the expected value of the square of the score (gradient of log-likelihood) at the maximum likelihood estimate:

$$\mathcal{I}(\boldsymbol{\theta}) = \int \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2 \ell(X \mid \boldsymbol{\theta}) \mathrm{d}X = \mathbb{E}^X \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2 \right], \tag{18}$$

where $\ell(X \mid \widehat{\boldsymbol{\theta}})$ is the likelihood function. The Fisher information (see Appendix) can also be written in terms of the second derivative of the log-likelihood (the Hessian):

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X \left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \mid \boldsymbol{\theta}) \right] \tag{19}$$

For a Bernoulli process with probability p, the FI is given by $1/(p(1-p))$ (Appendix); for a binomial process with sample size n this becomes $n/(p(1-p))$. For the Bernoulli process a conservative approximation to the FI is simply $1/p$ - for things such as tag recapture data when the probability of recapture is small this basically is a minor under-estimation (and associated minor over-estimate of uncertainty) in the FI. For multinomial data this approximation is actually exact (as long as we remember to sum across all the possibilities in the process).

In our case, the probability of occurrence is actually a function of the parameters we are really interested in, $p(\boldsymbol{\theta})$. The Fisher information of the parameter vector (see Appendix) is defined as follows:

$$\mathcal{I}(\boldsymbol{\theta}) = \mathcal{I}(p) \left(\frac{dp}{d\boldsymbol{\theta}} \right)^2 = \frac{1}{p} \left(\frac{dp}{d\boldsymbol{\theta}} \right)^2, \tag{20}$$

which, to improve numerical stability given small probabilities, can be shown (see Appendix) to be equivalent to the following:

$$\mathcal{I}(\boldsymbol{\theta}) = 4 \left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}} \right)^{\dagger} \left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}} \right) \tag{21}$$

This gives us the (approximate) Hessian matrix for the data at their base level covariates, $\mathbf{z}$. A very useful result is that the overall Hessian matrix, H, can be defined in terms of the sum over

all the unique covariate groups of the individual Hessian matrices $H(\mathbf{z})$ as defined in (21):

$$H = \sum_{\mathbf{z} \in \mathcal{Z}} n^{\mathbf{z}} H(\mathbf{z}),\tag{22}$$

where $\mathcal{Z}$ is the set of possible covariates, and $n^{\mathbf{z}}$ is the number of observations of that covariate type in the data set. The assumed independence between the different data sets (total recaptures, spatial recapture distribution, and length frequency) means the overall Hessian matrix, from which we can form the approximate covariance of the parameter vector as a whole, is the sum of the data set specific Hessian matrices as defined in (22). The parameter covariance matrix, $\Sigma_{\boldsymbol{\theta}}$, is the inverse of this overall Hessian matrix. Approximate estimates of the variance of parameters *derived* from $\boldsymbol{\theta}$ (such as spawning stock biomass, mean recruitment level and so on) via some differentiable function $g(\boldsymbol{\theta})$ can be derived via the delta method:

$$\mathbb{V}(g(\boldsymbol{\theta})) \approx \left(\frac{dg}{d\boldsymbol{\theta}}\right)^{\dagger} \Sigma_{\boldsymbol{\theta}} \left(\frac{dg}{d\boldsymbol{\theta}}\right).\tag{23}$$

If there is evidence of over-dispersion in any of the key processes (such as for the total tag recapture model (Day *et al.*, 2015)) this is simple to account for in the construction of the overall Hessian matrix for the parameters. With a given estimate of the variance inflation factor, φ, then we simply divide the Hessian matrices by this variance inflation factor, which will inflate the overall variance appropriately.

To make sure all these calculations are accurately done, we use automatic differentiation (AD) - specifically the `TMB` package in R - so that all gradients and derived quantities are calculated to machine precision. This is a fairly complex derivation and perhaps a little opaque given it leans quite heavily on statistical theory, but once done it makes it *very* quick and simple (relative to both simulating data and back-estimating parameter) to obtain the uncertainty estimates of interest, given the data collection settings.

Results

Conditioning the OMs and scenarios explored

The key control parameters are mean recruitment, the spatial recruitment proportion and the migration rates between regions. In a two region model (such as that considered in the Macquarie Island toothfish assessment), this is 4 parameters. We are working in an equilibrium paradigm and so we need to decide what are the key control factors we will impose on the model to parameterise it, and what parameters we estimate to condition the OM.

The four major factors we impose to condition the OM are:

1. Total SSB depletion to be 0.5 (the current target)

2. Total catch to be 440t (the long-term TAC likely to meet the target)

3. The catch split by region (25:75 and 50:50 North and South)

To meet those criteria, once imposed, we allow the following three factors to vary:

1. $\bar{R}$ the mean total recruitment

2. h_1 and h_2 the region-specific harvest rates

The assessment parameters we keep fixed are:

1. $\eta_1 \in \{0.2, 0.4, 0.6\}$ the spatial recruitment proportion

2. $\pi_{1,2} \in \{0.01, 0.05\}$ the movement from N to S

3. $\pi_{2,1} \in \{0.05, 0.025\}$ the movement from S to N

The scenarios around spatial recruitment fraction, regional catch split, and the movement parameters result in 24 distinct scenarios. The naming convention around the scenarios is in terms of $ijkl$ where: i is one of the three η_1 options; j is one of the two N to S movement scenarios; k is one of the two S to N movement scenarios; and l is one of the two spatial catch split scenarios. So *1111* would be $\{\eta_1, \pi_{12}, \pi_{21}, \delta_1\} = \{0.4, 0.01, 0.05, 0.25\}$. Given recent estimates of selectivity given the spatial length data we assumed that : $\lambda_{1,50} = 65$, $\lambda_{1,95} = 80$, $\lambda_{2,50} = 75$, $\lambda_{2,95} = 100$.

The conditioning code actually works by solving the system of non-linear equations that hit the target SSB depletion, total catch and relative catch split by region, by changing $\bar{R}$, h_1 and h_2. All the 24 scenarios were solved so that the targets were attained exactly.

In terms of total and spatial tagging rates per tonne we assumed that total tagging rate per tonne, ψ, was kept at 2 in all cases. Given this overall rate, we explored $\psi_1 = \{1.5, 2, 3\}$ for the Northern area, which obviously interacts with ψ and the spatial catch split to implicitly define the tagging rate for area 2 the Southern area.

Uncertainty estimates of key assessment variables

The key uncertainty estimates we focus on are for equilibrium spawning stock biomass (SSB), and the movement parameters $\pi_{1,2}$ and $\pi_{2,1}$ to see how well the different tagging strategies compare when estimating overall spawning stock size and the movement between areas. To deal with with recruitment variability, and it's contribution to overall uncertainty in equilibrium SSB we used the following two relationships for the unfished SSB:

$$B_0 = R_0 \sum_a e^{-M(a-1)} w_a m_a,$$

$$\mathbb{V}(B_0) = R_0^2 \left(e^{\sigma_r^2} - 1\right) \sum_a \left(e^{-M(a-1)} w_a m_a\right)^2.$$

There is another similar, albeit somewhat more complicated formula to calculate the variance in equilibrium SSB in the exploited state, but we omit that here. The current level of $\sigma_r \approx 0.3$ is assumed in these runs, and the resultant CV in equilibrium SSB ranges between a CV of 0.05 to 0.06. We assume (probably safely) that stochastic recruitment variation will alter the variation in
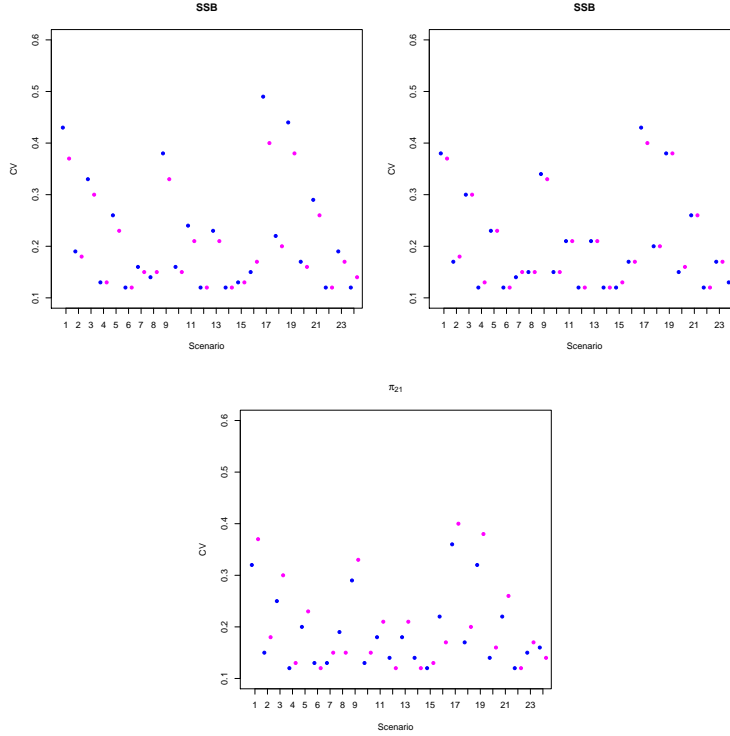
Figure 1: *The predicted CV in overall SSB across all 24 scenarios for a tagging rate of 1.5 (left), 2 (middle), and 3 (right) tags-per-tonne in the Northern area.*

overall abundance, but not migration, and is independent of the other factors. We can then simply used the additive nature of independent variance components to calculate the overall variation in SSB given both the tagging based estimates and recruitment variation.

Figure 1 shows the predicted coefficient of variation (CV) in the overall SSB for each of the three spatial tagging rate scenarios $\psi_1 \in \{1.5, 2, 3\}$ contingent on an overall tagging rate of 2 tags per tonne. The range of CVs for the $\psi_1 = 1.5$ (less tagging in the North) is between 0.12–0.5 but the key features causing estimates of the CV to be above 0.3 are (i) 25% of the catch coming from the Northern area; (ii) higher levels of relative recruitment to the Northern area; and (iii) lower levels of migration from the Southern to the Northern area. There is a clear general trend that - overall - as we gradually increase the number of tags released in the Northern area $\psi_1 \to 1.5 \to 2 \to 3$ the estimates of the CV in overall SSB decrease, particularly the largest CV scenarios.

Figure 2 summarises the median and approximate 95% CI for the North to South movement rate, $\pi_{1,2}$. For lower levels of releases in the northern region ($\psi_1 = 1.5$) if the true rate of migration between the Northern and Southern regions is low (0.01) the estimates gradually get worse as the catch share moves from being 25/75 to 50/50 (i.e. more catch in the Northern area), and as the rate of movement between the Southern and Northern area decreases (from 0.05 to 0.025). This trend also follows for the 2 and 3 tag per tonne release rates in the North. Similarly to the SSB CV estimates, the movement estimates get more accurate as the number of releases in the Northern
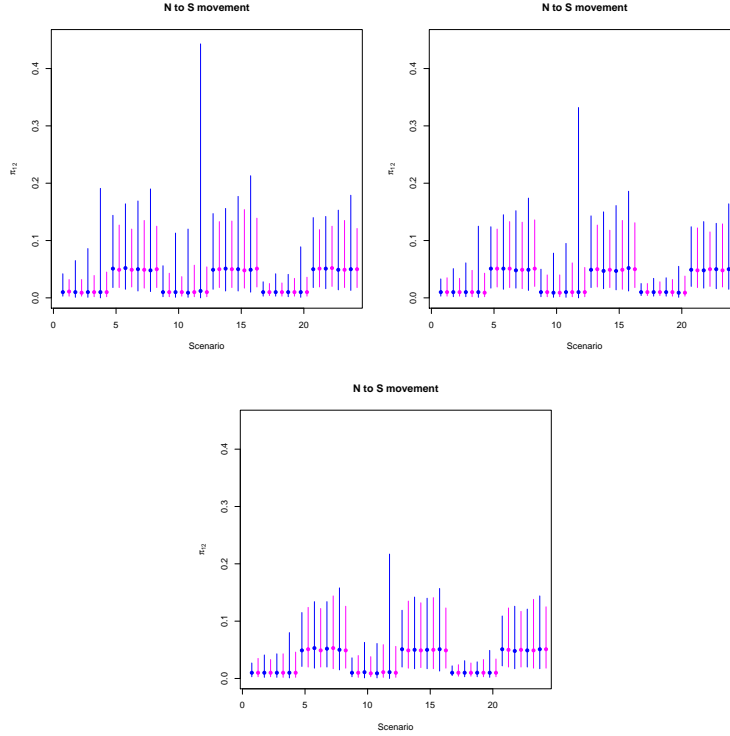
Figure 2: *The median (blue circle) and approximate 95% CI (blue line) for the North to South movement rate, $\pi_{1,2}$, across all 24 scenarios for a tagging rate of 1.5 (left), 2 (middle), and 3 (right) tags-per-tonne in the Northern area.*

region increases.

Figure 3 the same result for the South to North movement rate, $\pi_{2,1}$. The trends for these estimates - across all spatial release scenarios - is less clear. More accurate estimates seem to be associated with scenarios where the spatial catch share is 50/50, movement between the South and the North is lower not higher, or where the only change from the base case is that the proportion of recruitment to the Southern area is at its lowest level ($\eta_1 = 0.6$). In contrast to the North to South movement parameter, across the board as the spatial tagging rate becomes focussed more on the North, the South to North movement estimates generally get more uncertain.

Estimates of selectivity for the Northern area for the length at 50% selection have a range of CVs from 0.2 to 0.3, and for the length at 95% selection the CVs range between 0.4 and 0.6. Both are better when the TAC split is 50/50 across the regions (i.e. where samples are actually taken in the Northern area). Estimates of selectivity for the Southern area for the length at 50% selection have a range of CVs from 0.08 to 0.16, and for the length at 95% selection the CVs range between 0.16 to 0.31. The estimates are also better for when there is more catch taken in the South ($\delta_2 = 0.75$) but not by much, relative to how much the Northern estimates improve when the catch split is more like 50/50.
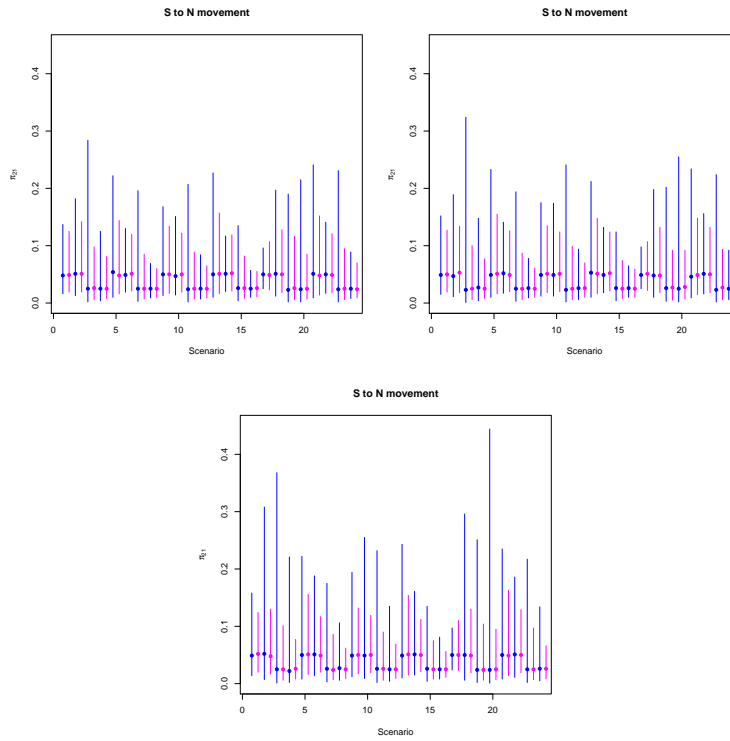
Figure 3: *The median (blue circle) and approximate 95% CI (blue line) for the South to North movement rate, $\pi_{2,1}$, across all 24 scenarios for a tagging rate of 1.5 (left), 2 (middle), and 3 (right) tags-per-tonne in the Northern area.*

Discussion

What constitutes the "best" spatial tagging design has been a question of interest for a number of years for this fishery and the stock assessment used to drive its management advice. In this paper we have tried to systematically approach how the spatial catch distribution, the underlying population dynamic parameter (both scale and spatial), and spatial tagging regimes interact in relation to the uncertainty estimates in the key assessment parameters driven by the tagging data (SSB and migration).

An information theoretic approach, originally suggested in (Bravington *et al.*, 2016) in relation to close-kin mark-recapture models and extended to the conventional mark-recapture and length frequency data we collect routinely, was used. This avoids the need for both simulation and back-estimation of the key parameters. We refer interested readers to Eveson *et al.* (2012) for factors that are known to bias these kinds of estimators. Given that, while spatial tagging rates varied, there was a balanced strategy with releases and recaptures in all spatial areas - and with this and the overall number of recaptures in general - estimation variance will be a much more influential factor than bias.

For estimates of overall SSB, the overarching result was that - across all other factors - having higher rates of tag release in the Northern regions, relative to the South, resulted in more accurate estimates of SSB. More specifically, all the CVs were between 0.1 and 0.35 and this regime avoided the appearance of high estimates (above 0.4 and approaching 0.5). At a more detailed level, lower levels of catch taken in the Northern area, higher levels of relative recruitment to the Northern area, and lower levels of migration from the South to the North result in consistently high CVs for SSB.

When estimating movement between the regions, at the highest level there is a clear trade-off between the accuracy of North to South/South to North movement as we shift the spatial tagging rates. Perhaps unsurprisingly, the more tags-per-tonne we release in the Northern region, the more precise the movement estimates from the North to the South, and the less precise the movement rates from the South to the North. As we move the releases more to the Southern region this trend generally reverses. For the North to South estimates, the 50/50 catch split and lower South to North movement rates seemed to result in less accurate estimates. The contrast with the South to North estimates they generally seemed to associate with more the 50/50 spatial catch split, and with less proportional recruitment to the Southern region ($\eta_1 = 0.6$).

Estimates of selectivity are generally good - more so for the Southern area relative to the North, especially at the older ages. The accuracy of Northern parameters was clearly associated with a 50/50 catch ratio; this effect was less clear for the Southern region, but still associated with higher catch levels in the South.

There are a number of clear influencing factors such as spatial catch split, spatial tagging rate per tonne, south to north migration rates and spatial recruitment factors. Some of these we can control - like spatial catch split and tagging rates - but the others we simply have to live with. Additionally, not all levers that we can pull are going to be beneficial across the board. For example, increasing the tagging rate in the Northern area seems to improve the estimates of SSB and the North to South movement parameters, but the South to North movement estimates

become more uncertain. A 50/50 spatial catch ratio (i.e. not 25/75 North/South) improves the estimates of SSB, South to North movement and Southern selectivity, but did the opposite to North to South movement.

It is not, perhaps, surprising that there is no unequivocally optimal spatial tagging strategy and/or spatial catch distribution. There are clear trade-offs between key assessment outputs. If the primary focus is SSB then higher release rates in the North *and* enforcing a 50/50 spatial catch split quite clearly results in more accurate estimates, and avoids overly large CVs. The 50/50 catch ratio does seem to have negative effects on some of the movement parameters (North to South). Without a full MSE analysis it is hard to know how these trade-offs would play out in terms of fishery performance and management outcomes. What we can perhaps say is that having higher rates of tagging in the Northern area appear to have generally positive outcomes - in particular for the key management variable of SSB. Also, these are implicitly long term rates for all key data collection and management levers - these ignore the potential for additional information that can be gained from time-varying levels of spatial catch and tagging rates.

References

Bravington, M. V., Skaug, H. J., and Anderson, E. A. 2016. Close-kin mark recapture. Stat. Sci.

Brownie, C., Anderson, D.R., Burnham, K.P., Robson, D.S. 1985. Statistical Inference From Band Recovery Data: A Handbook **156** (US FIsh Wildl. Resour. Publ.)

Day, J., Haddon, M., and Hillary, R.M. 2014. Stock Assessment of the Macquarie Island fishery for Patagonian toothfish (*Dissostichus eleginoides*) using data up to and including August 2014. SARAG.

Day, J., Haddon, M., and Hillary, R.M. 2015. Stock Assessment of the Macquarie Island fishery for Patagonian toothfish (*Dissostichus eleginoides*) using data up to and including August 2015. SARAG.

Eveson, J. P., Basson, M., and Hobday, A. J. 2012. Using electronic data to improve mortality and movement estimates in a tag-based spatial fisheries assessment model. Can. J. Fish. Aquat. Sci. **69**: 1–15.

Fay, G. 2011. Stock assessment of the Macquarie Island fishery for Patagonian toothfish (*Dissostichus eleginoides*) using data up to and including June 2010.

Hillary, R. M., Wayte, S., Day, J., and Haddon, M. 2014. Appropriate growth models for Macquarie Island toothfish. SARAG.

Hillary, R. M. and Eveson, J. P. 2015. Length-based Brownie mark-recapture models: Derivation and application to Indian Ocean skipjack tuna **163**: 141–151.

Methot., R.D., and Wetzel, C.R. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**: 86–99.

Appendix

Calculating the Fisher information via the second derivative option requires far more calculation than does the option via the square of the score (i.e. needing only first derivatives). The two methods defined in the main text are equivalent, and we show now why is this the case. Consider first the second derivative of the log-likeilihood:

$$\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \mid \boldsymbol{\theta}) = \frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \mid \boldsymbol{\theta})}{\ell(X \mid \boldsymbol{\theta})} - \left(\frac{\frac{\partial}{\partial \boldsymbol{\theta}} \ell(X \mid \boldsymbol{\theta})}{\ell(X \mid \boldsymbol{\theta})} \right)^2 = \frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \mid \boldsymbol{\theta})}{\ell(X \mid \boldsymbol{\theta})} - \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2$$

Taking the first part of the above equation and calculating its expectation over the data:

$$\mathbb{E}^X \left[\frac{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \mid \boldsymbol{\theta})}{\ell(X \mid \boldsymbol{\theta})} \right] = \int \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ell(X \mid \boldsymbol{\theta}) \mathrm{d}X = \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \int \ell(X \mid \boldsymbol{\theta}) \mathrm{d}X = 0.$$

The additive nature of the expectation, and the result in the previous equation, give rise the the equivalent definitions of the Fisher information:

$$\mathcal{I}(\boldsymbol{\theta}) = \mathbb{E}^X \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln \ell(X \mid \boldsymbol{\theta}) \right)^2 \right] = -\mathbb{E}^X \left[\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \mid \boldsymbol{\theta}) \right].$$

For the Bernoulli process, with probability p, the likelihood (and log-likelihood) of a binary outcome $X = \{0, 1\}$ is

$$\ell(X \mid p) = p^X (1 - p)^{X-1},$$
$$\ln \ell(X \mid p) = X \ln p + (1 - X) \ln(1 - p),$$

and the second derivative of the log-likelihood is given by

$$\frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln \ell(X \mid p) = \frac{X}{p^2} + \frac{1 - X}{(1 - p)^2}.$$

For the Bernoulli distribution we have that $\mathbb{E}(X) = p$, so the expected value of the second derivative of the log-likelihood (i.e. FI) is:

$$\mathcal{I}(p) = \frac{p}{p^2} + \frac{1 - p}{(1 - p)^2} = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{p(1 - p)} > \frac{1}{p},$$

which would make p^{-1} a conservative approximation to the Fisher information for Bernoulli processes (such as the total tag recapture model).

Here we detail exactly how we calculate the Fisher information matrix for the estimated parameters, given the original Fisher information is defined in terms of the underlying probability (which is a function of the estimated parameters). Assuming $p(\boldsymbol{\theta})$ is at least two-times differentiable in terms of $\boldsymbol{\theta}$ then, via the chain rule, we have that

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X \left[\frac{\partial^2}{\partial p^2} \ln \ell(X \mid p) \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right)^2 + \frac{\partial}{\partial p} \ln \ell(X \mid p) \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2} \right].$$

Taking the second part of the expression in the above expectation we have that

$$\mathbb{E}^X \left[\frac{\partial}{\partial p} \ln \ell(X \mid p) \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2} \right] = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2} \mathbb{E}^X \left[\ln \ell(X \mid p) \right] = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2} \frac{\partial}{\partial p} \left(\int \ell(X \mid p) \mathrm{d}X \right) = \frac{\partial^2 p}{\partial \boldsymbol{\theta}^2} \times \frac{\partial}{\partial p} (1) = 0.$$

So the second part of the term over which we take the expectation to obtain the parameter-focussed Fisher information disappears, leaving only the following:

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}^X \left[\frac{\partial^2}{\partial p^2} \ln \ell(X \mid p) \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right)^2 \right] = - \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right)^2 \mathbb{E}^X \left[\frac{\partial^2}{\partial p^2} \ln \ell(X \mid p) \right] = \mathcal{I}(p) \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right)^2.$$

The base level Fisher information for the Bernoulli process is

$$\mathcal{I}(\boldsymbol{\theta}) = \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right)^\dagger \frac{1}{p} \left(\frac{\partial p}{\partial \boldsymbol{\theta}} \right),$$

but with small probabilities (such as those in mark-recapture models) this can scale very poorly numerically. Consider $f = \sqrt{p}$, then

$$\frac{df}{d\boldsymbol{\theta}} = \frac{1}{2\sqrt{p}} \frac{dp}{d\boldsymbol{\theta}},$$

and

$$\left(\frac{df}{d\boldsymbol{\theta}}\right)^{\dagger}\left(\frac{df}{d\boldsymbol{\theta}}\right) = \frac{1}{4p}\left(\frac{dp}{d\boldsymbol{\theta}}\right)^{\dagger}\left(\frac{dp}{d\boldsymbol{\theta}}\right).$$

Given this it follows that:

$$\frac{1}{p}\left(\frac{dp}{d\boldsymbol{\theta}}\right)^{\dagger}\left(\frac{dp}{d\boldsymbol{\theta}}\right) = 4\left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}}\right)^{\dagger}\left(\frac{d\sqrt{p}}{d\boldsymbol{\theta}}\right),$$

which is more numerically stable and easier to deal with.

What can we do if p is not sufficiently small, and ignoring the $(1-p)^{-1}$ begins to increasingly under-estimate the information value in the data? This *can* happen in large-scale tagging events with higher exploitation rates and very good reporting rates. In many tuna applications the probability of recapture can be as high as 20-25% and this begins to be a hard to ignore short coming in the simple Fisher information construction algorithm defined herein. Going back to the true FI for the Bernoulli type process, it would be actuall defined as follows:

$$\mathcal{I}(\boldsymbol{\theta}) = \left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^{\dagger}\frac{1}{p(1-p)}\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right) = \left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right)^{\dagger}\left(\frac{1}{p} + \frac{1}{1-p}\right)\left(\frac{\partial p}{\partial \boldsymbol{\theta}}\right).$$

If your actual probabilities of recapture are large enough that these terms would make a difference then it is likely that the revised FI definition above will be numerically well behaved enough (especially using AD to calculate the derivatives) to be able to be used in place of the approximate version, as applied to cases where p is very small.