On the use of tagging data in statistical multispecies multi-area models of marine populations

Sigurdur Hannesson, Audbjorg Jakobsdottir, James Begley, Lorna Taylor, and Gunnar Stefansson


The use of multispecies models of marine stocks along with recognition of the importance of temporal differences in spatial overlap has resulted in migration rates playing an increasingly important role in models of fish stocks. Moreover, traditional estimates of growth based on samples from fishing gear are confounded with the selection pattern, which is exacerbated when multispecies issues are considered. For these and other reasons, there is a need to include explicit tagging data as a component of fisheries models. A statistical multispecies multi-area framework has been extended to predict tag returns and subsequently to incorporate tagging data in likelihood components to be used when estimating migration rates. The information content of such data is not clear a priori, but simulations indicate the point estimates to be quite reasonable. A bootstrap method is proposed, based on bootstrapping entire tagging experiments (rather than individual fish). The resulting bootstrapped uncertainty estimates are generally applicable and are found to be close to the true values in the simulated examples.

Keywords: bootstrapping, migration estimation, multispecies models.

Received 13 December 2007; accepted 28 March 2008; advance access publication 20 August 2008.

Introduction

Recent multispecies models incorporate multiple areas to account for species overlap, an important concern when considering species interactions (Stefansson and Palsson, 1997a, 1998; Begley and Howell, 2004; Taylor et al., 2007). This implies an immediate need to account for migration in such models, and these migrations are typically poorly known. The rates therefore need to be estimated and the errors associated with these estimates need to be determined.

This is all the more important in the light of recent developments in the management of fisheries, where the potential for marine protected areas (MPAs) as management tools is becoming better acknowledged, although still under debate (Hannesson, 1998; Lubchenco et al., 2003). Such management procedures have been criticized for lack of rigour, because they may not have been subject to the same level of testing as traditional management systems (Hilborn et al., 2004), although some such testing has been initiated through simulations (Stefansson and Rosenberg, 2005).

In some cases, it is possible to use such simulations of different migration scenarios to evaluate the effects of MPAs. However, a more common requirement will be to obtain reasonable point estimates of migration rates for a given environment. In particular, this type of information is likely to allow some leeway in how, for example, MPAs are set up. The evaluation of corresponding harvest control rules will require knowledge of the corresponding error rates. These error rates are quite important, because fairly small changes in assumed migration rates can seriously affect simulation results (Stefansson and Rosenberg, 2006). The same concerns will apply when considering multispecies models where, for example, species overlap is an important parameter and is a function of migration rates.

Here, we describe several simulation experiments designed to investigate the effects of using tagging data to estimate migration rates. The experiments are intended to shed some light on the estimability of migration rates in the context of statistical multispecies multi-area models and to provide a means to measure the uncertainty associated with the estimates.

The concept of a substock (or, equivalently, a stock component) is used to denote the fish in a given area and maturity stage. Most of this work uses data simulated quarterly on two areas, with fish migrating between the areas while subject to harvest mortality, natural mortality, and other processes.

Bootstrap methods are designed and used to ascertain the estimation errors associated with maximum likelihood estimation of migration rates in the non-linear population dynamics models. Simulated populations are used to verify the estimation of error rates.

Finally, we present a case study for immature and mature cod (Gadus morhua) around Iceland, using a real dataset.

General framework

General multispecies population dynamics

The following describes in fairly generic terms the population dynamics which can be incorporated in current spatially explicit multispecies models of marine animal populations. All of these
processes are implemented in a computer program, Gadget (see http://www.hafro.is/gadget), designed for generic multispecies modelling (Begley, 2004; Begley and Howell, 2004). If considered important in the parent population (or substock in Gadget terminology), these processes also need to be taken into account when implementing tagged subpopulations.

Denote by \( N_{\text{alsmrt}} \) the number (\( N \)) of animals of species \( s \), age \( a \), length \( l \), and maturity stage \( m \), alive in region \( r \), and time-step \( t \). This group of individuals is a subset of the entire collection of animals of a given species. Such a collection will be referred to as a substock. Mathematically, this is just a collection of numbers, but they are usually unknown and will need to be estimated using statistical techniques.

Normally, there are only two maturity stages within the model, immature and mature animals (\( m = 0,1 \), though for closed life cycles, there is a need to include the egg and larva stages as a separate population. Always, only a fixed area, time-step, and maturity stage are considered, and in this case, such indices are omitted and the notation is simplified to \( N_{\text{alsm}} \).

Several processes can affect such a group of animals and cause a change in numbers in the group. A framework such as Gadget (Taylor et al., 2007) includes the following:

(i) Migration undertaken by a population on a given time-step is described by matrices, \( P = (p_{ij})_{i=1...R,j=1...R} \) containing the proportion \( p_{ij} \) of the population that moves from area \( r_j \) to area \( r_i \), and \( R \) is the number of areas. Hence, if \( u = (u_1, \ldots, u_R) \) denotes abundance numbers by area for an age–length–maturity–species group in a population, and \( P \) is such a matrix, then the area distribution after the migration has taken place is \( Pu \). When the migration matrices vary by age, time, maturity stage, etc., the notation \( P_{\text{alsmrt}} \) may be needed.

(ii) Consumption of a prey by a predator and fishing by a fleet is described through a set of equations which describe the desired food supply of a predator and the suitability of each prey class. This is only used in the context of fishing in the present paper.

(iii) Natural mortality other than predation/fishing mortality is usually taken as a fixed number for each species, but it may of course vary with time, age, or size.

(iv) Growth in length within multispecies population dynamics models is most naturally implemented through growth in weight, because the weight increase is more naturally linked to consumption. Alternatively, growth may be according to a fixed growth schedule in either length or weight. In any case, as the model is based on the frequency in length groups, growth in length needs to correspond to transfers of fish between length groups in such a manner that the predicted mean growth is maintained for the length group. This is implemented through the use of a length-update probability distribution, where the updating distribution is a beta-binomial which has the mean length increment as one parameter, and the other is a free parameter (\( \beta \)) and can be estimated or fixed.

(v) A (mature) population may spawn to generate offspring and lose biomass. This possibly results in spawning mortality.

(vi) Maturation involves shifting fish from a population of immature fish to another population of mature fish. This is done using proportions (as in the migration process), which are designed to mimic the resulting proportion of mature fish in each age–length cell.

(vii) Ageing involves increasing the age by 1 year, which occurs at the last time-step of a year (except for the oldest age group, which is usually a plus group).

The order in which the processes are implemented is important. The above sequence is the order in which these are implemented in Gadget.

Tagged subpopulations

When fish have been tagged, it is useful to think of the tagged fish as a subpopulation of the population (or substock) from which they were taken. A tagged subpopulation can be modelled as a subgroup, \( N_{\text{alsmrt}} \) of the numbers in a population, \( N_{\text{alsm}} \) where \( e \) denotes a given tagging experiment. Each tagging experiment forms a separate subgroup.

Ideally one would like to tag fish and record, at the time of tagging, every possible piece of information, so they can be entered as a completely known population into the model. Naturally, when fish are tagged, they are also length-measured and the position is known, but the age and maturity stage are generally not identified. Therefore, when forming the model population from the tagging data, the tagged fish are allocated to age groups in accordance with the age composition of the substock at the time-step \( t \) of tagging:

\[
N_{\text{alsmrt}} = N_{\text{alsm}} \times N_{\text{alsmrt}} / N_{\text{alsm}},
\]

where \( " \) denotes aggregation over the corresponding index.

In addition to the processes described above, tag loss may need to be taken into account. This is done through a parameter, \( \gamma \), referred to as tag loss. At each time-step \( 1 - e^{-\gamma} \), a proportion of the tagged subpopulation loses its tags at the same time as natural mortality acts on the subpopulation.

Ignoring for the moment this additional process, the new tagged subpopulations have exactly the same dynamics as the original population. Assuming that the number of tagged fish is low compared with the population totals, the tagged fish do not affect the dynamics of the population as a whole. Therefore, the tagged dynamics can be viewed conditionally on the dynamics of the whole population, but from an implementational point of view, these components will normally be simulated simultaneously. It is clear that computational effort will increase as the number of tagging experiments increases.

For technical reasons, it is useful to think not only of the number of tagged fish, but also of the proportion of tags, \( \hat{p}_c = N_{\text{alsm}} / N \), in a given group (with \( a, l, m, s, r \) and \( t \) fixed).

The tagged subpopulations need to be subjected to exactly the same processes as described for the original populations. For practical reasons, the implementation of this is quite important. Consider first the mortality processes that affect the numbers in a given cell through reduction alone. They do not change the proportion of tagged fish, and the values of \( \hat{p}_c \) remain fixed throughout these processes. It follows that these processes can be implemented for the parent population alone, and the proportions \( \hat{p}_c \) can be used after the process to recompute the tagged numbers. This is in contrast to implementing migration, because in this case the proportions can change drastically and are not easy to track until after the migration process has been completed, so here the process needs to be applied to each tagged subpopulation separately. It is therefore clear that
judicious use alternately of numbers and proportions can simplify implementation considerably. Given the computational resources required for a simulation environment such as Gadget, this can be quite important with regard to computational time.

Observation model
As the same population dynamics process applies to the tagged subpopulation, the predicted number of tag returns is just the predicted tagged catch, \( \hat{C}_{\text{almsrfte}} \), which is predicted like the catch of any other substock (Begley, 2004; Begley and Howell, 2004). The corresponding data or observed count, \( C_{\text{almsrfte}} \), should have this expected value if the model is correct. The new subscript, \( f \), refers to fleet, because there may be many fleets in operation.

Likelihood components
The models used in this paper and implemented in Gadget have deterministic forward-projection models as a basic internal simulation mechanism. Any such forward projection is based on assumptions, several of which may be formulated in terms of parameters which are potentially unknown but can be estimated.

A forward projection of the system results in a set of predicted values for all model components, including the catch in numbers of each length and age group at each time-step. This also provides predicted mean lengths-at-age along with predictions of arbitrary aggregations of catches in weight or numbers. Any of these can be used to obtain likelihood components if corresponding data are available. The estimation itself proceeds by minimizing the sum of all negative log-likelihood components, possibly after weighting each appropriately. In addition to proper likelihood components, it is common to add penalties when bounds are exceeded and a special penalty component which enters as the squared difference is added. In Gadget, this can be quite important with regard to computational time.

The mean length-at-age can be entered as sums of squares, and data on length distributions can be entered as Gaussian distributions. The Gadget likelihood components are described in Begley and Howell (2004). To accommodate tag-recapture data, e.g. as needed in this work, a likelihood component has been added to Gadget as described below.

The recapture data are initially taken as Poisson counts, as proposed by Hilborn (1990). As the Poisson parameter is now \( \hat{C}_{\text{almsrfte}} \), the corresponding likelihood contribution becomes

\[
\Pi_P e^{-\hat{P}} \frac{\hat{C}_{\text{almsrfte}}}{\hat{C}_{\text{almsrfte}}!}
\]

This is the ideal situation, applicable when information about length, maturity, age, and location of recaptures is available for each individual fish recaptured. In cases when one or more of these pieces of information are unavailable, aggregation will be needed in one form or another. Normally, the area, period and tagging experiment are available, but one or more of the others may be missing. It is not clear how to combine data with varying levels of information, but it is certainly possible to use only the information which tends to be available always, e.g. using

\[
\Pi_P e^{-\hat{P}} \frac{\hat{C}_{\text{almsrfte}}}{\hat{C}_{\text{almsrfte}}!}
\]

when information is only available on region, fleet, time-step, and recapture (\( r, f, t, e \)), so aggregates over age, length, maturity stage, and substock (\( a, l, m, s \)) are used.

The emphasis in this approach is on predicting the observed counts. It follows that this is mainly designed to resolve issues of numbers, i.e. to provide some added resolution to mortality rates and population numbers.

Simulation, data, and estimation
The level of information available in tagging data is not clear, nor is the required number of tagging experiments or number of fish in each tagging experiment, to obtain a given reduction in variance of migration parameters. It should also be acknowledged that tagging experiments are rarely carried out on fish selected at random, whereas use within a model would generally be to calculate average migration rates. Similarly, it is not at all clear how the non-linear nature of the population models affects the parameter estimates, but in general, some bias is to be expected. A simulation approach has been developed to verify the assumption and estimation accuracy in the parameters of multispecies, multi-area models.

Biological parameters in the simulated stock
To focus on the effects of migration and associated uncertainty, the population models in simulations are kept as simple as possible. The model is run for 20 years with four (quarterly) time-steps each year. A single stock is simulated, but the stock is distributed across two areas, with an age range of 1–10 years and a length range of 4.5–90.5 cm, at 1-cm intervals, giving 86 length groups.

The simulated stock grows according to a von Bertalanffy growth function [Gadget’s “lengthvbsimple” growth function (Begley, 2004)], with parameter values \( L_\infty = 95, k = 0.08 \). The weight-at-length relationship is a simple power relationship, \( w = d^a \), with \( a = 1 \times 10^{-5} \) and \( b = 3 \). The maximum length group growth was 6, and in the beta-binomial, \( \beta \) was set at 12.

Recruitment enters the stock once a year, and for simplicity, the simulated recruitment enters as 1 million recruits in each area, each year. The stock will be assumed to migrate every time-step. To simulate variable spawning, returning and feeding behaviour, there will be three different migration matrices. These will assume the same migration pattern for the second and third time-step each year. These matrices are:

\[
P_i = \begin{pmatrix} 1 & p_i \\ 0 & 1 - p_i \end{pmatrix},
\]

with \( i = 1, 4 \), \( p_1 = 0.4 \), and \( p_4 = 0 \); and

\[
P_j = \begin{pmatrix} p_j & 0 \\ 1 - p_j & 1 \end{pmatrix},
\]

with \( j = 2, 3 \), and \( p_2 = p_3 = 0.4 \).

Notably, two of the migration rates are always fixed to be the same in the simulations. This is of interest, because it will be used to test how consistent the estimates are.
The fishing fleet catches a fixed proportion of the stock each year, implemented as 6.25% of the stock each quarter, corresponding to a constant fishing mortality (as a proportion) of $F = 0.25$, but alternative scenarios are considered. Natural mortality is set to 0.2 for all ages.

In some real situations, the catch information is disaggregated by season and area, and this will provide some bounds on the migration parameters. In the present models, the catch from the stock was aggregated across areas, so those likelihood components did not affect the estimation of migration rates. This is done to avoid confusion over which part of the model provides information on migration.

Simulated stocks and components
A simulation model has been developed in the R statistical package (see http://www.r-project.org), simulating the population dynamics described in the previous section. This simulation model is an extension of Sigurdardottir et al. (2005), with the addition of generating recaptures from tagging experiments. The simulated data are generated separately from the estimation procedure, which is to be verified.

The simulation predicts the expected number of tag returns from the catch of the tagged stock, $C_{\text{almsf}}$. Stochastic simulated data are generated from a Poisson distribution with this parameter, and the generated number is rounded to the nearest integer. These simulated data are then fed into the estimation procedure to give estimated migration rates (or other parameters if needed).

Repeating these simulations provides multiple sets of estimated migration rates, which can be used to estimate the variability in estimated migration rates (in addition to any bias). As these values are obtained by simulating different “realities”, they can be obtained with any desired accuracy simply by repeating the simulations. It should be noted that these then become “true” standard errors of estimated migration rates, i.e. these give $\sigma_p$.

In real situations, one attempts to estimate the errors using a numerical procedure (such as Hessian matrices or bootstrapping), which gives $\sigma_p$. The simulation approach can be used to verify whether the estimates are realistic at least within the bounds of the simulated world.

Real data
A model with two areas and two stock components was considered using biological assumptions detailed elsewhere (Taylor and Stefansson, 2004; Taylor et al., 2007). The model includes immature and mature cod stock components and only the mature fish migrate between the areas. There are 12 time-steps in each year. Many parameters were fixed to the values estimated from a simpler single area model, such as growth and maturity parameters. Migration rates and two sets of initial population parameters and recruitment parameters were estimated using tagging data, survey indices, and age distribution data. This is because migration is linked to the estimation of initial population and recruitment, which strongly affect the biomass and in turn influences the number of recaptures.

In the model, immature cod can be 1–10 years old and occupy 1-cm intervals in the range 4–130 cm. These grow according to a von Bertalanffy growth function (Gadget’s "lengthvsimple" growth function; Begley, 2004) with a maximum growth of ten length groups per time-step. Recruitment occurs in the first time-step each year, and this substock matures into the mature stock. Mature cod are 3–12 years old and 20–140 cm long (1-cm intervals). They grow like immature cod. This substock migrates with the migration pattern similar to that in the simulated model, described below.

There are two fleets, commercial and the March groundfish survey. Further, there are a considerable number of likelihood components, but the most important, here, are the recapture likelihood components. The others are:

(i) bound likelihood component: a penalty is generated when the bounds given for the parameters to be estimated are exceeded;
(ii) migration penalty likelihood component: generates a penalty if there is a negative entry in a migration matrix;
(iii) catch distribution likelihood components: sums of squares of the difference of the proportion of the data and the model sample of fish caught by time, area, and age, to estimate the initial population;
(iv) understocking likelihood component: sums of squares of overconsumption in the model are included to ensure that no more fish are removed from the stock than a certain maximum proportion (Begley and Howell, 2004; Taylor and Stefansson, 2004); and
(v) survey indices likelihood components (Taylor et al., 2007). These likelihood components are described in detail in Begley and Howell (2004).

A basic split of Icelandic and surrounding waters is given in Figure 1. In this Figure, subareas 101–108 basically give waters <500 m deep and are usually considered the main cod areas. Area 102 is a feeding area, subject to considerable migration, whereas the main spawning grounds are in 101 and 108. For testing with real data, the models use two areas: a “southern area”, Area 1 (subareas 101, 102, 107, and 108) and a “northern area”, Area 2 (subareas 103, 104, and 105). These are all based on the definitions set out in Taylor (2003b). This is not completely in accordance with convention, because common assessments include 102 with the northern part.

When analyses have a fine time-scale and with most of the spawning in subareas 101 and 108, together with the main feeding ground in 102, the migration rates are expected to be fairly low. They are therefore scaled to percentages for presentation.

Tagging data
Tag-recapture data for the years 1991–1997 were extracted from a data warehouse (Kupca and Sandbeck, 2003; Taylor, 2003a, b; Kupca, 2004). Tagged fish were aggregated by month, year, and area, giving 13 tagging experiments. All tagged fish >40 cm were classified as mature, because these tagging experiments targeted spawning aggregations. The tagging experiments are summarized in Table 1.

Tag loss was fixed at 0.02 for all 13 tagging experiments, and other parameters were estimated. At this rate, more tagged fish lose tags than die in each time-step, because natural mortality is 0.2 per year but tag loss is 0.02 per month. Note that tagging mortality and non-reporting are not specifically taken into account in Gadget, so the tag-loss parameter also represents these processes.

The weights on the recapture data in the negative log-likelihood function were all set equal. Although alternatives exist (Taylor et al., 2007), the emphasis here is on a generic estimation
method and generic estimation of uncertainty. It is important that this generic uncertainty estimation works also when the weights are uncertain or incorrect, so no attempt is made to adjust these.

Migration
To differentiate between spawning, returning, and feeding behaviour, three migration matrices are implemented. These matrices are as in Equations (1) and (2), with \( p_1 \) and \( p_2 = p_3 \) to be estimated, but \( p_4 = 0 \).

From January to April (inclusive), migration matrix \( P_1 \) is used, from May to August (inclusive) and December, migration matrix \( P_2 \) is used, and from September to November (inclusive), there is no migration (\( P_3 \) is used). It is assumed that no fish will move to areas other than the two areas involved.

Model implementation
The above population dynamics models and estimation procedures have been implemented within Gadget. Estimation is always implemented with an initial global search, the simulated annealing algorithm (Corana et al., 1987), followed by a local search, the Hooke and Jeeves algorithm (Hooke and Jeeves, 1961), both implemented within Gadget (Begley, 2004; Begley and Howell, 2004).

Bootstrap estimation of uncertainty
The bootstrap (Efron, 1979) is a very general method that can be used to obtain estimates of uncertainty. Given the complexity, specialized bootstrap methods have been developed to shed some light on the issue of estimability and accuracy of parameter estimates in typical multi-area fisheries models. Some care needs to be taken when implementing the bootstrap for tagging data, or any fisheries data for that matter. In particular, it is very clear that the fish within a single tagged sample are not randomly sampled from the entire population of wild fish, so the traditional bootstrap (Efron and Tibshirani, 1994) needs to be modified. The correlated nature of these data is commonly termed the “intra-haul correlation” (Pennington and Volstad, 1994). It is taken into account here by taking the sampling units for bootstrapping purposes to be individual tagging experiments rather than individual tagged fish.

For a scenario with a collection of \( N \) tagging experiments, bootstrapping is implemented by sampling \( N \) experiments (with replacement) from this collection, giving a bootstrap sample. Given this
boottstrapped sample, the parameters are estimated giving a single
bootstrap parameter estimate. This procedure is then repeated to
obtain repeated bootstrap estimates of parameters, subsequently
providing estimated standard errors.

It is not known a priori whether the bootstrap procedure will
provide reasonable estimates of uncertainty. As with any proposed
method, the bootstrap needs to be evaluated and this will be done
within the context of a simulated population, where the distribu-
tion of migration estimates around the expected value is known.

Results
Model verification

Initial tests based on data without any error and deterministic
recaptures were used to verify the consistency between the R simu-
lator and the Gadget estimator. When as many as 50 000 fish are
tagged, the estimated values of the parameters are quite close to
the “true” values (Table 2; maximum difference <1.5%). There
will always be some error in the recapture likelihood, because
the real recaptures, which are integers, are compared with the pre-
dicted values from Gadget, which are real (positive) numbers.

Never is there an obvious problem or bias, although there is
some deviation when estimating all three values as opposed to
restricting the rates to be equal (Table 2). When only 500 fish
are tagged, the point estimates deviate more from the true values
than with 50 000 tagged fish (Table 2). The difference in the esti-
mates of \( p_2 \) and \( p_3 \) when all three are estimated freely is now in the
opposite direction from before. Again, there is no obvious indi-
cation of bias.

The optimization might potentially be sensitive to the initial
values of the migration parameters, but repeated optimization
with alternative initial values gave the same estimates as before
(Table 2).

The likelihood score for the migration rates in the domain \([0, 1] \times [0, 1]\) was computed using Gadget. A plot of the likelihood
surface for the migration rates \( p_1 \) and \( p_2 \) shows that, as a function
of these two parameters, the negative log-likelihood surface is
quite well behaved, with a single minimum (Figure 2).

Estimation from simulated stochastic data

In all the following simulations, recaptures are drawn stochasti-
cally. When each tagging experiment is considered an independent
experiment and the number tagged are treated as a single sub-
population throughout each simulation, the computational
requirements increase drastically, because the number of exper-
iments increases. It is therefore of considerable interest to test a
different approach, where all tagged individuals are considered a
part of the same tagged population, and during a simulation
period, tagged individuals are simply added to the existing
tagged population.

The following subsections consider two tagging scenarios. In
the first scenario with 40 tagging experiments, tagging occurs in
time-steps in both areas for 5 years. The second scenario con-
siders eight tagging experiments, where the year of tagging is ran-
domly chosen from 6 potential years, and there is tagging on every
time-step in both areas for that year.

Scenario 1: all steps—all areas—all years

A tagging experiment is simulated in each of two areas every
quarter in each of 5 years, yielding 40 tagging experiments.

Each tagging as a separate experiment: The number of tagged
fish in each experiment is normally distributed (and rounded to
the nearest integer), with a mean of 280 fish and a standard devi-
ation (s.d.) of 30. For such a generated dataset, the migration rates
are estimated. The procedure is repeated 40 times to verify the
variability in the estimated migration rates. The true values of
the parameters were fixed at 0.4, and the resulting estimates
came out with an average of 0.398 (s.d. = 0.018) for \( p_1 \) and 0.406 (s.d. = 0.012) for \( p_2 \), and a correlation between the two is
0.698. Figure 3a–c shows the distribution of the estimate of the
migration rates \( p_1 \) and \( p_2 \), along with a scatterplot of the estimates
of the two parameters.

An aggregated experiment: Consider again the 40 experiments
conducted above. This time, these are considered as a single aggre-
gated experiment, where fish tagged subsequently are added to the
existing tagged stock during the simulation. The advantage of this
approach is that the simulation time is reduced by an order of
magnitude as the number of populations stored in Gadget is
reduced considerably. Again, the true values of the parameters
were fixed at 0.4. The resulting estimates came out with an
average of 0.399 (s.d. = 0.017) for \( p_1 \) and 0.398 (s.d. = 0.012) for
\( p_2 \), and a correlation between \( p_1 \) and \( p_2 \) of 0.719, compared with 0.698 earlier, with the tagging experiments disaggregated.
These results are depicted in Figure 3d–f in the form of histograms
of the distribution of estimated migration rates, along with a
scatterplot of the two.

Bootstrap estimation of uncertainty: The same scenario as before
is considered here. Bootstrapping was done 100 times to obtain
100 estimates of migration rates, providing estimated standard
errors. The “base value” in this case is the migration rate estimated
from the initial simulated data from which the bootstrap samples
were simulated. For \( p_1 \), the resulting estimates came out with an
average of 0.391 (s.d. = 0.011, base value = 0.383) and with
0.383 (s.d. = 0.011, base value = 0.383) for \( p_2 \), and a correlation of
0.518 (Figure 3g–i). Note that the bootstrap mean is similar to
the true value, indicating that there is not a great bias. Also
notable is the fact that the standard errors are \( \approx 0.011 \), slightly

Table 2. Point estimates of migration rates based on 50 000
simulated tags and variable number (500 or 50 000) of stochastic
(Poisson) recaptures, using fixed true values, but different initial
values and using or not using the correct assumption of equal
migration rates.

<table>
<thead>
<tr>
<th></th>
<th>50 000 simulated</th>
<th>500 simulated tags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p_1 )</td>
<td>( p_2 )</td>
</tr>
<tr>
<td>True values</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Initial values</td>
<td>0.50</td>
<td>0.40</td>
</tr>
<tr>
<td>Estimated values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assume ( p_2 = p_3 )</td>
<td>0.403</td>
<td>0.402</td>
</tr>
<tr>
<td>tagging and catch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>data in likelihood,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>low weight on tags</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assume ( p_2 = p_3 )</td>
<td>0.403</td>
<td>0.402</td>
</tr>
<tr>
<td>only tagging data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>in likelihood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different initial values</td>
<td>0.41</td>
<td>0.42</td>
</tr>
<tr>
<td>Estimated values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All rates estimated</td>
<td>0.401</td>
<td>0.409</td>
</tr>
</tbody>
</table>
lower than the simulated values of 0.012–0.018, but certainly of the same order.

**Scenario 2: all steps—all areas—year chosen randomly**

In this scenario, there are eight tagging experiments, two at each time-step from the year, one from each area. The year at tagging is randomly chosen from 6 years for each of the eight tagging experiments. Tagging data were simulated with Poisson error (100 datasets), using R as before. Gadget was then run with optimization, to estimate the two migration rates.

When each tagging experiment is treated as a separate experiment, fixing the true migration rates at 0.4 results in an average of 0.402 (s.d. = 0.035) for $\hat{p}_1$ and 0.410 (s.d. = 0.029) for $\hat{p}_2$. Figure 4a–c shows the distribution of estimated migration rates along with their bivariate distribution migration rates, where the correlation is now 0.671. When the tagging experiments are aggregated, the results obtained are again similar. As before, the true values are set to 0.4, and the average point estimates of $p_1$ and $p_2$ are 0.397 (s.d. = 0.033) and 0.400 (s.d. = 0.024), respectively, with a correlation of 0.554.

Note that the standard deviation of these migration rates is lower in the multiple experiment than it was in the single experiment. The resulting plots are similar to those obtained earlier (not shown).

For the bootstrap, the same scenario with eight tagging experiments is considered again. Bootstrapping these 100 times with a base value of 0.400 for $p_1$ results in a mean value of 0.397 (s.d. = 0.016). For $p_2$, the base value was 0.399 and the bootstrap mean was 0.396 (s.d. = 0.0257). In this setting, the correlation was 0.626 or slightly lower than with the disaggregated tagging experiments. The bootstrap results are seen in Figure 5a–c.

**Changing fleet and fishing mortality**

In all the simulations above, the catch was modelled as a simple proportion of the population (Gadget’s “linearfleet” predation function) with fishing mortality ($F$, but as a proportion of the stock size) set constant across areas, time-steps, and years.

When fishing mortality varies among time-steps (i.e. proportional fishing mortalities of $F = 0.25, 0.375, 0.125, 0.25$ for time-steps 1–4), the estimates of migration parameters are similar to those when fishing mortality is identical for all time-steps.

**Real data**

The model described above was implemented in Gadget, and parameters were estimated. Bootstrapping was subsequently used to obtain confidence statements for the migration rates. Although the model predicted fewer recaptured fish than were actually recaptured, the proportion of recaptured fish by area was similar. The point estimates of the migration rates were 1.559 and 0.997%.

**Bootstrapping**

 Entire tagging experiments were bootstrapped, and parameters were estimated for each bootstrap sample. Within each bootstrapped sample, the experiments were chosen randomly, with the restriction that there had to be at least one tagging experiment from each area (resulting in 97 rather than 100 bootstraps). The results are given in Table 3 and Figure 6.

The point estimate of absolute bias in the percentage migrating is from 0.08 to 1.5 when the correct percentage is 1–1.5%. The 95% confidence intervals based on the bootstrap are (0.246,
3.100%) for migration rate \( p_1 \) and (0.396, 1.750%) for migration rate \( p_2 \). A boxplot of the migration rates from the bootstrap runs is provided in Figure 7.

The bootstrap means are close to the base value (point estimate from the initial data) but the standard deviation is relatively high. In Figure 6, both the histogram (Figure 6a) of the estimated migration rate \( \hat{p}_1 \) and the scatterplot (Figure 6c) show that the points lie in three groups. In the first, \( \hat{p}_1 < 1.1 \), in the second \( 1.1 \leq \hat{p}_1 \leq 2.6 \), and in the third \( \hat{p}_1 > 2.6 \). Of the observed tagging data, one particular experiment was very large (2818 fish tagged) with a high return rate (38.1%), including almost one-third of all recaptures. Optimization runs using a bootstrapped dataset including this experiment never optimize to solutions with \( \hat{p}_1 < 1.1 \). Runs optimizing with \( 1.1 \leq \hat{p}_1 \leq 2.6 \) include this experiment 1–3 times in the bootstrap data, and with \( \hat{p}_1 > 2.6 \) the experiment is present 2–5 times in the bootstrap data.

**Discussion**

Adding tagging data to the existing (multispecies) multi-area models allows the estimation of migration rates, at least in principle. Results here indicate that these migration rates are indeed estimable, at least for few areas and when reliable estimates of some other model parameters have been obtained from other sources. More-general models will typically be fitted to several...
datasets, and in these cases, one should expect the tagging data to resolve the migration rates if other parameters become estimable from the remaining datasets.

The estimators of migration seem to be biased when using real datasets, and this bias appears to be of the order of 10% of the true value. It is very likely that one could fit better models to the dataset in question, e.g. using a reweighting scheme to estimate the weights given to each likelihood component, etc., but this is somewhat outside the scope of this work. On the other hand, by their very nature, tagging experiments cannot be a randomized sample from the underlying population being modelled. It is therefore quite natural for there to be some biases in the estimates, particularly with areas as large as those used here.

In general, a bias may occur for many reasons. Any model including a tagging process will include assumptions or models for tag loss, survival in the population, survival after tagging, the migration process, and growth. A simple example of a failure of such models is the lack of very high spatial resolution, which is the only way to explain some of the microscale fish behaviour or grouping, and this can easily lead to higher observed tag returns than can be explained with large-area models. Finally, although growth is not linked to the tagging data in this particular model, in the general case, fitting a growth submodel will affect the migration estimates.

These initial simulations indicate that repeated tagging experiments may, at least in some cases, be treated as a single experiment, where new tags are added to tagged populations already in the model. This greatly reduces the computations required. It is not known, however, how variable survival rates or other

Table 3. Results from estimating migrating percentages \( p_1 \) and \( p_2 \) using bootstrapping on real data for cod in Icelandic waters.

<table>
<thead>
<tr>
<th>Point estimate (base value)</th>
<th>Bootstrap mean</th>
<th>Bootstrap s.d.</th>
<th>Bias</th>
<th>Relative bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 ) 1.559%</td>
<td>1.356%</td>
<td>0.748%</td>
<td>-0.202%</td>
<td>-0.130%</td>
</tr>
<tr>
<td>( p_2 ) 0.997%</td>
<td>0.916%</td>
<td>0.387%</td>
<td>-0.080%</td>
<td>0.081%</td>
</tr>
</tbody>
</table>

Figure 4. Scenario 2, bootstrapped Gadget estimates based on simulated tagging experiment with true values at \( p_1 = p_2 = 0.4 \). Eight tagging experiments, two per time-step, one from each area in a randomly picked year. Each tagging experiment is treated as an individual experiment within Gadget. Univariate distributions of the estimated migration rates \( \hat{p}_1 \) (a), \( \hat{p}_2 \) (b), and bivariate scatterplots of \( \hat{p}_2 \) vs. \( \hat{p}_1 \) (c).

Figure 5. Bootstrapped Gadget estimates based on simulated tagging experiment with true values at \( p_1 = p_2 = 0.4 \). Base data are chosen from a Poisson simulation, and the bootstrap is used to obtain uncertainty estimates. Histograms indicate the distribution of \( p_1 \)-estimates (a), \( p_2 \)-estimates (b), and bivariate scatterplot of \( \hat{p}_2 \) vs. \( \hat{p}_1 \) (c).
Complications might affect these results. Preliminary results indicate that weights attached to likelihood components may be important. Indications are that a tagging experiment with a large number of tagged fish and recaptures changes the estimate of the migration rates. The weights should depend on the number of comparisons or the number of fish recaptured, but the exact form is not obvious.

In addition to estimating migration rates, an important issue is obtaining the errors of these estimates. This is non-trivial, because the models are highly non-linear, and several earlier simulations indicate that simple approaches to variance estimation in fisheries models provide highly misleading results, even in single-stock, single-area models (Gavaris et al., 2000; Restrepo et al., 2000; Patterson et al., 2001).

Bootstrap-based methods based on resampling entire tagging experiments appear to capture the uncertainty in estimating migration parameters. Although this result may be conditional on the models used, the results obtained appear promising.

Gadget is simple when it comes to reducing the number of tagged fish. Only tag loss in time is considered here, but there is also tagging mortality, which should reduce the initial number of tagged fish. Other issues include non-reporting, which would be modelled by reducing the predicted recaptures by a fraction. These two issues cannot be distinguished by the data considered here and should therefore appear as one parameter to be estimated (Hilborn, 1990).

In addition to this layout, it must be noted that there is information in the recapture data on the growth of each recaptured fish. This information is only used implicitly in the above, and not at all when the likelihood components are aggregated over length and age groups. Gadget is a Markovian model, neglecting the length of fish before the previous time-step. It is therefore not trivial how information on the history of the individual fish should be brought into a new likelihood component, but this is certainly not insoluble and needs to be explored. The easiest way would be to put each length group as a specific tagging experiment, but that would increase the computational time dramatically. It is, however, possible to compare the mean length of recaptured fish from the model and observed recaptures, and this has in fact been implemented within Gadget.

Only one likelihood function for tag-recapture data has been considered here but there are other possibilities, such as a multinomial likelihood (Hilborn, 1990). The exact choice of a likelihood function is likely to affect the precision of results as is true in ordinary multiple linear regression. The important issue is, however, that the bootstrap approach can be used with any such estimation method to obtain uncertainty estimates. Further, it is fairly well established that standard statistical assumptions rarely apply to fisheries data, be these length distributions, landings, age compositions, stomach samples, or tagging experiments (Pennington and Volstad, 1994; Pennington, 1996; Stefansson and Palsson, 1997b; Brynjarsdottir and Stefansson, 2004; Elvarsson, 2004). One can therefore not expect to find “the” correct likelihood function. If one can use a bootstrap method to evaluate the uncertainty even for incorrect likelihoods, then this potentially gives a means to evaluate which likelihoods give less biases and variance.

Two-area examples have been considered here. Adding more areas to a model increases the computing time for single simulations. In the “real data” example, some 30 parameters need to be estimated for each area, in addition to the migration rates whose number would increase too. Although it might be interesting to investigate the effects of more complex geometries, this should probably be done in the light of particular applications.

An interesting result from this analysis is the low estimate of migration of mature fish between the areas chosen here. This is
probably a consequence of classifying subarea 102 with the south. If this result holds under further scrutiny, it may imply that fairly homogeneous (sub)populations are obtained using this split. This should therefore be considered for assessments, because this may lead to more consistent estimates of catches in numbers-at-age, survey indices, etc., than when the area split corresponds to higher migration rates.

Acknowledgements
Much of the work described here was undertaken while the authors were employed at the MRI (Marine Research Institute, Reykjavik) and uses data from the MRI databases. The R simulator was originally written by A. J. Sigurðardóttir and E. Olafsdóttir, whereas the Gadget code has been in development for more than a decade by many programmers at the MRI and IMR (Institute of Marine Research, Bergen). The work was supported in part by EU grantQLK5-CT1999-01609, a grant from The Icelandic Centre for Research (Rannis), and the FMAP project within the Census of Marine Life.

References


doi:10.1093/icesjms/fsn132