CCSBT-OMMP/1906/04
(OMMP Agenda item 2.5)

Oceans & Atmosphere
www.csiro.au

Changes to SBT OM conditioning code

Rich Hillary, Ann Preece, Campbell Davies

18th June 2019

CSIRO Oceans & Atmosphere
Battery Point, Hobart 7000, Tasmania, Australia.

Copyright and disclaimer

Important disclaimer

Contents

Abstract

The SBT operating models (OMs) are being reconditioned this year for MSE testing of candidate MPs. In addition to updating existing data,we also have two gene tagging estimates for 2016 and 2017 to include for the first time. This paper details the technical specifications of how the gene-tagging data are included in the SBT OM, and the relevant settings and fixed parameters required in the various OM configuration files.

1 Background

This year the OMMP and ESC will be resuming the MSE work begun in 2018 to develop a new MP for the CCSBT. A reconditioning update of the OM is required in 2019 and will include two gene tagging data points for 2016 and 2017 in the conditioning code. These data have already been included in projection code [1, 2] and the same assumptions about the generation of these data in the projections will be mirrored in the conditioning part of the OM.

2 Gene tagging process

The gene tagging data collection process is as follows:

1. In year y, T_y (assumed to be) 2 year old fish are tissue-sampled and re-released off Port Lincoln in South Australia **after** the surface fishery has caught all its fish
2. In year $y+1$, S_{y+1} (assumed to be) 3 year old fish are tissue-sampled in the post-processing facilities in Port Lincoln
3. In year $y+2$, R_{y+2} recaptures are found

We don't go into specifics about the length distribution of tagging and resampling, save that we do this to ensure the maximum chance of tagging 2 year old and resampling 3 year old fish.

3 Likelihood function

In the MP work, we use the simple Petersen estimator for the age 2 abundance in year y, $\widehat{N}_{y,2}$:

$$\widehat{N}_{y,2} = \frac{T_y S_{y+1}}{R_{y+2}},$$

with the Poisson approximation to the variance where the CV in abundance is assumed to be approximated by $1/\sqrt{R_{y+2}}$. For the conditioning of the OM we assume a more flexible distribution: the beta-binomial distribution. The underlying probability of recapturing a biopsied fish is as follows:

$$\pi^r_{y+2} = \frac{T_y}{q^{\mathrm{gt}} N_{y,2}},$$

where q^{gt} represents the fraction of age 2 juveniles available to be tagged in the GAB (default is 1). The other key parameter for the gene tagging likelihood is the over-dispersion coefficient, φ^{gt}: the degree to which the variance in the recaptures exceeds that assumed in the vanilla binomial distribution (i.e $\varphi^{\mathrm{gt}} \geq 1$). With the binomial ($\varphi^{\mathrm{gt}} \equiv 1$), we have the following likelihood:

$$\Lambda^{\mathrm{gt}}\left(R_{y+2} \,|\, S_{y+1}, \pi^r_{y+2}\right) \propto \left(\pi^r_{y+2}\right)^{R_{y+2}} \left(1-\pi^r_{y+2}\right)^{S_{y+1}-R_{y+2}}$$

For the over-dispersed case, $\varphi^{\mathrm{gt}} > 1$, the likelihood is as follows:

$$\alpha^{\mathrm{gt}} = \frac{\left(S_{y+1} - \varphi^{\mathrm{gt}}\right)\pi^{r}_{y+2}}{\left(1-\pi^{r}_{y+2}\right)\left(\pi^{r}_{y+2} + \left(1-\pi^{r}_{y+2}\right)\left(\varphi^{\mathrm{gt}}-1\right)\right)}$$

$$\beta^{\mathrm{gt}} = \frac{\left(S_{y+1} - \varphi^{\mathrm{gt}}\right)\pi^{r}_{y+2}}{\pi^{r}_{y+2} + \left(1-\pi^{r}_{y+2}\right)\left(\varphi^{\mathrm{gt}}-1\right)}$$

$$\Lambda^{\mathrm{gt}}\left(R_{y+2} \,|\, S_{y+1}, \alpha^{\mathrm{gt}}, \beta^{\mathrm{gt}}\right) \propto \frac{\Gamma\left(R_{y+2}+\alpha^{\mathrm{gt}}\right)\Gamma\left(S_{y+1}-R_{y+2}+\beta^{\mathrm{gt}}\right)\Gamma\left(\alpha^{\mathrm{gt}}+\beta^{\mathrm{gt}}\right)}{\Gamma\left(S_{y}+\alpha^{\mathrm{gt}}+\beta^{\mathrm{gt}}\right)\Gamma\left(\alpha^{\mathrm{gt}}\right)\Gamma\left(\beta^{\mathrm{gt}}\right)}$$

and $\Gamma()$ is the gamma function.

4 Settings required in OM configuration files

The data are included as follows in the `sbtdata20XX.dat` file as a table with the following columns: year of release, age of release, year of recapture, number of releases, number of resamples, number of matches. Table 4.1 shows the current data set.

Year of rel.	Age of rel.	Year of recap.	T	S	R
2016	2	2017	2,952	15,389	20
2017	2	2018	6,480	11,932	67

Table 4.1: *Summary of current gene tagging data.*

The remaining control parameters are located in the `sqrt.dat` file:

- `qgt` (q^{gt}): default is set to 1 (and assumed that $q^{\mathrm{gt}} \leq 1$)
- `gtOD` (φ^{gt}): default is set to 1 (and $\varphi^{\mathrm{gt}} \geq 1$)
- `gtsw`: 0/1 switch flag to turn GT data off/on (default set to 1)

5 Fits given reconditioned reference set of OMs

A full diagnostic check of the fits for all updated data sets will be undertaken for the stock assessment in 2020. However, given this is the first time the gene tagging data have been included in the OM, we do summarise how the reconditioned OM fits to these data. The approach taken in the past few years [3] is to simulate a particular data set from its predictive distribution (simulate from the likelihood while integrating across the model ensemble contained in the reference set). If the reference set of OMs was a true posterior, this would be the posterior predictive distribution; given we use the reference set as a proxy for the posterior we refer to it as the predictive distribution.

Figure 5.1 shows the observed and predictive distribution of (in terms of median and 95% credible interval) matches in the 2016 and 2017 gene tagging data (year we denote as year of release/year of abundance estimate). In both cases the median number of matches is slightly below the observed number, indicating a preference for lower age 2 abundance in the gene tagging data, but the credible interval easily encapsulates the data in both cases.

It might seem odd that these data are not fitted effectively perfectly, given there are no other data sets that currently observe these year-classes at the present time. There is, however, a reason-

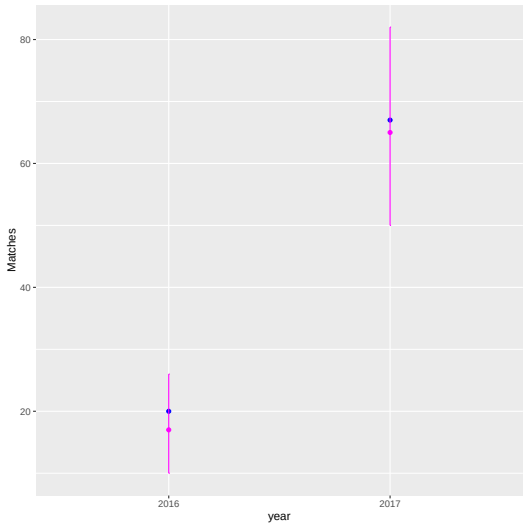

Figure 5.1: *Observed (blue) and predictive median and 95% credible interval (magenta) for the 2016 and 2017 gene tagging recaptures.*

ably informative prior on the year-class strength deviations in the OM, and with auto-correlation built in. The estimates of recruitment prior to 2016 were well above average (especially age 2 abundance in 2015), so built in to the recruitment deviation prior in 2016 and 2017 is a preference for above-average recruitment deviations. This is why the effect looks more obvious for 2016 (which follows the highest recruitment estimate for decades) than for 2017 (as the 2016 age 2 abundance was estimated closer to the expected level). The summary though would be that:

- The conditioning part of the OM has been modified to incorporate the gene tagging data using a flexible beta-binomial likelihood and is implemented as the data are simulated in projection part of the OM
- The data from 2016 and 2017 are fitted well by the reconditioned OM, but suggesting *slightly* lower 2016 and 2017 estimates of age 2 abundance coming from the previous OM and the recruitment deviation prior
- While being cautious about infering too much from only 2 estimates, the gene tagging data does seem to suggest that the previous run of above-average recruitment might be over

6 Acknowledgements

This work was funded by CSIRO and the Australian Fisheries Management Authority.

References

[1] R. M Hillary, A. Preece, and C. R. Davies (2016) Methods for data generation in projections. *CCSBT–OMMP/1609/07*

[2] R. M Hillary, A. Preece, and C. R. Davies (2017) Updates required for new data sources and reconditioning of the CCSBT OM. *CCSBT–OMMP/1706/04*.

[3] R. M. Hillary *et al.* (2017) Reconditioning of the CCSBT Operating Model in 2017. *CCSBT–ESC/1708/14*.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australias catalyst for innovation.

WE IMAGINE. WE COLLABORATE. WE INNOVATE.