IOTC-2008-WPTT-18

Reporting rate analyses for recaptures from Seychelles port for yellowfin, bigeye and skipjack tuna

R. M. Hillary,
Division of Biology,
Imperial College,
London,
UK.

IOTC Secretariat,
IOTC,
Victoria,
Seychelles.

J. J. Areso,
Oficina Espanola de Pesca
Victoria,
Seychelles.

Abstract

This paper details analyses undertaken to estimates reporting rates for Seychelles landed recaptures of yellowfin, bigeye and skipjack tunas. The tag seeding data are used in a binomial form and GLM and Bayesian techniques are employed to explore the information in the various predictors and the uncertainty in the reporting rate estimates. Key factors that are explored are year, quarter, seeder type (skipper or observer), species and commercial size category. A clear and substantial increase in the reporting rate is seen over time – for all species and size categories – which clearly correlates with the onset of the reward and recapture program section of the RTTP and with the increase of the number of fish being recovered, increasing the awareness of the stevedores.The GLM approach was employed to indentify informative factors and potential interactions between them, and then the Bayesian estimator was then used to estimate reporting rate distributions at the GLM-derived aggregation level. In the years and quarters in which the main portion of tag returns exist we observe a very high reporting rate of above 90% in almost all species.

Introduction

Good estimates of reporting rates are absolutely vital when intending to use mark and recapture data to estimate exploitation/survival rates and population abundances – without any information on reporting rates we simply cannot perform such analyses.

Tag reporting rate can be estimated from tag seeding experiment (Youngs, 1974 ; Green *et al.*, 1983 ; Campbell *et al.*, 1992, Hampton, 1997) ; by comparison of the tag return rate for a platform with those of a control group reporting *a priori* all recoveries

– such as observers , or finally by observations of recoveries at different stage of catch handling or processing (Hillborn, 1988).

Within the framework of the IOTTP, the IOTC has conducted since 2004 a tag seeding experiment onboard the Purse-Seine fleet based in Seychelles, during which tuna caught by those Purse-Seiners were discreetly tagged by fisheries observers or voluntary skippers before the fish were placed in the brine wells. The tags used were almost similar to the ones used for the tagging and release of live fish. In order not to alter the detection by the stevedores unloading the boat, or the workers processing the fish, the leader of the tag was strictly the same as for the other tag with the same printing. The only difference resides in the attachment of the tag which differs from the plastic barb used on the normal tags. Seeding tags are implanted on dead fish, as a result the anchorage of the tag within the pterigyophores is not secure by the healing of the fish around the attachment. To avoid rapid shedding, metal attachments were used, securing a better anchorage within the flesh of the fish. The tags were implanted in the same location on the fish, at the basis of the second dorsal fin. During each trip, 15 tags and one applicator was provided to the tagger who was asked to tag the three species in different wells and to spread the tag within different size categories. As a result, since 2004, 1056 SKJ, 1076 YFT and 218 BET have been seeded within the fleet and unloaded in Seychelles. In This paper, only the one reported by stevedores in Seychelles have been accounted for, in order to have an estimation of the reporting rate for the recovery platform “purse-seiner unloaded in Seychelles”. Differences of reporting rate per year, per size category and per tagger - observers and skippers- were investigated through a Generalized Linear Model (Nelder, J. And Wedderbum, R. 1972) for the tags unloaded in Seychelles.

The data were organized in a binomial form: for each of the key data groupings (seeder, quarter, year, species, size category) if a seeded tag was reported it is recorded as the number one, if not reported it appears as a zero giving us success/failure data. There we 2350 tags seeded over the years 2004 to 2008 for two tagger types (observer and skipper), three size categories (commercial size categories), and three species (yellowfin, bigeye and skipjack), that were unloaded in Seychelles by stevedores. To study only the reporting rates from the recovery plateform “Stevedores in Seychelles”, the recoveries from the cannery where removed.

Data and methods

Two approaches were used to explore the information in the tag seeding data: GLM and Bayesian. The GLM analysis was exploratory in the sense that this type of approach makes it somewhat easier to look for the significance or otherwise of a range of effects. The results of the GLM analyses were then used to set the aggregation level used in the Bayesian analysis. The data are in binomial form (numbers of successes and failures of observing a seeded tag) and are available by year, quarter, tagger type, species and commercial size category.

GLM exploratory models

The simplest initial GLM was to have (non-interacting) effects for year, quarter, species, tagger and size category all as categorical variables. For this simple first-attempt model the only significant effects were for the year effects in 2006, 2007 and 2008, and a quarter effects for quarter 4 One would expect a year effect given we see an increase in the nominal reporting rates from 2005 to the present and the quarter 4 effect is likely driven by the sharp increase in reporting from quarter 3 to quarter 4 in 2005. There appears to be no strong evidence of species,size-category or seeder differences in reporting rate.

In terms of what conclusions we can make from the GLM analyses, there appears to be little evidence of species, size category or seeder differences in mean reporting rate. There are some clear year and quarter interaction processes as well as a potentially complex relationship between tagger and year. In reality, there is a rapid increase in the nominal reporting rate in mid-2006 and then a fairly stable reporting rate up to the present making it difficult to see how a linear model – even with interactions – can properly model the process.

Bayesian beta-binomial model

For each year, species, tagger-type and size category we have a recorded number of reporting successes, $N_{y,q,sp,t,s}$, and reporting failures, $M_{y,q,sp,t,s}$. We use a binomial probability model for the reporting probabilities, $\pi_{y,q,sp,t,s}$, so that

$$p(N_{\cdot\cdot\cdot} \mid \pi_{\cdot\cdot\cdot}) = \binom{N_{\cdot\cdot\cdot}+M_{\cdot\cdot\cdot}}{N_{\cdot\cdot\cdot}} \pi_{\cdot\cdot\cdot}^{N_{\cdot\cdot\cdot}} (1-\pi_{\cdot\cdot\cdot})^{M_{\cdot\cdot\cdot}} \quad (1)$$

Henceforth, the three dots seen in Eq. (1) refer to the relevant aggregation scale of the data being analysed. As mentioned, we are using Bayesian estimation which requires the definition of a prior distribution for the reporting probabilities. We assume a beta prior for the reporting probabilities:

$$p(\pi_{\cdot\cdot\cdot}) = B(\alpha,\beta) \quad (2)$$

By letting *α=β=1* we define an effective uninformative prior to be sure that the data supply the information on the reporting rate parameters. Also, assuming a beta prior combines with the binomial likelihood/probability model to give a beta posterior density function:

$$p(\pi_{\cdot\cdot\cdot} \mid N_{\cdot\cdot\cdot}, M_{\cdot\cdot\cdot}) = B(\alpha+N_{\cdot\cdot\cdot}, \beta+M_{\cdot\cdot\cdot}) \quad (3)$$

This choice of a conjugate beta prior makes drawing from the posterior distribution of the parameters a simple task, so we can easily explore the issues of how sample size and not just the other effects affect our reporting rate estimates, as sample size is the main cause of uncertainty in such situations.

Results

From the GLM work there is little evidence of species or size-specific differences in mean reporting rate. We do not look at size-specific seeding information in this work but do look at species-specific information as there is a large difference in the number of seeded tags placed on the three main species and we can look at how this affects the precision of the reporting-rate estimates.

Tagger effects

The first issue we deal with is whether there truly are any tagger effects and to do this we chose the years in which there are enough tags seeded by both tagger types to do a comparative analysis – we chose 2005 and 2006 to do this test. In 2005/2006 there were 148/66 observer seeded tags and 348/606 skipper-seeded tags. The median and 95% probability interval for the observer-seeded and skipper-seeded reporting rates in 2005 were 0.43 (0.36-0.51) and 0.44 (0.39-0.49), respectively. For 2006 they were 0.8 (0.69-0.88) and 0.88 (0.86-0.91), respectively. In 2005 there was no significant difference in the mean or median of these two distributions and the only real difference is that the skipper-seeded estimate is more precise than the observer-seeded estimate given the larger sample sizes of skipper-seeded tags in this year. For 2006 it would seem at first glance that there is perhaps a difference between the two tagger types but this must first be considered in the context of the quarterly change in nominal reporting rate over this year (as the recovery PR program began) and when the tags were seeded. Observer-seeded tags were seeded roughly evenly in quarters 2 to 4; the majority of skipper-seeded tags were placed in quarters 3 and 4 – after the observed increase in nominal reporting rate. If we look at the estimates for observer and skipper-seeded tags for quarters 2 to 4 we have 0.63, 0.77 and 0.92 for observers and 0.84, 0.89 and 0.92 for skippers. The differences are largely seen in quarters 2 and 3 but not in quarter 4. It is difficult to say for sure that there are no differences in the reporting rates from observer and skipper-seeded tags but there is certainly no obvious or consistent trend in the observed differences. For the purposes of establishing reporting rate estimates over the period for which the overwhelming majority of the tag return information exists (2006-2008) we use only the skipper-seeded tags.

Estimates for 2006 to 2008

Using only the skipper seeded tags we now propose to look at how differing sample-size by species affects the precision of the reporting rate estimates. There are far more tags seeded for yellowfin (1076) and skipjack (1056) than for bigeye (218) and while the GLM analyses suggested little species difference in mean reporting rate differences in sample size will manifest as differences in the precision of the reporting rate estimates. Figures 1, 2 and 3 show the posterior distributions of the reporting rates-by-quarter for yellowfin, skipjack and bigeye, respectively.

For yellowfin (Figure 1) we see the rapid increase in reporting rate from quarter 1 to quarter 2 in 2006 which then stayed at a high level (0.9-0.95) until late 2007 and then experienced a gradual downwards trend to around 0.85 in 2008 and then an increase in the second quarter of 2008. For skipjack we have no seeding data in the first quarter of 2006 and we observe a consistently high reporting rate between 0.85 and

0.95 over the whole period with a similar, if slightly gentler, downward trend over 2007 as seen for yellowfin. For bigeye our estimates are clearly more uncertain (given lower sample sizes) and are also missing for the first quarter of 2006 and, apart from lower estimates of around 0.7 in the second quarter of 2006 and of 2008, remain high between around 0.85 and 0.95 for the rest of the time.

The all-species reporting rate can be seen in Figure 4 and is largely dominated by the yellowfin and skipjack trends for reasons of similar trends (increasing in 2006, slightly decreasing in 2007) and sample size. Overall, we observe a strong increase in the reporting rate from 2006 quarter 1 to quarter 2 and a generally high (> 0.85) reporting rate from this point up to the present. In terms of point estimates, given the binomial model and uninformative priors the posterior mode of the reporting rates coincides with what might be considered the 'raw' estimates of reporting rate (tags detected/sample size) and Table 1 details the posterior modal reporting rates for 2006 to 2008 for the all-species data.

Table 1: Posterior modal estimates of reporting rate for the all-species seeding data for the period 2006 to 2008.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
2006	0.61	0.84	0.89	0.91
2007	0.93	0.91	0.91	0.88
2008	0.88	0.91	N/A	N/A

Discussion

This paper details reporting rate estimates for the Stevedore-processed catches coming through Victoria, Seychelles. A mixture of GLM and Bayesian techniques were employed to explore the level of information on factors such as species, size and tagger in these data. It was seen that year and quarter interaction effects are present but that species and size-specific effects were not likely to influence the estimates of mean reporting rate, only the precision of these estimates. A sharp increase in reporting rate was observed from quarters 1 to 2 in 2006 and reporting rates have stayed high (> 0.85) ever since, with perhaps a slight decreasing trend throughout 2007. There are no visible systematic differences in the estimates of reporting rate by species so an all-species reporting rate is perhaps the most sensible to work with as it utilises all of the available data.

Figure 1: Yellowfin reporting rate posteriors (and median) for the skipper-seeded data for 2006-2008.

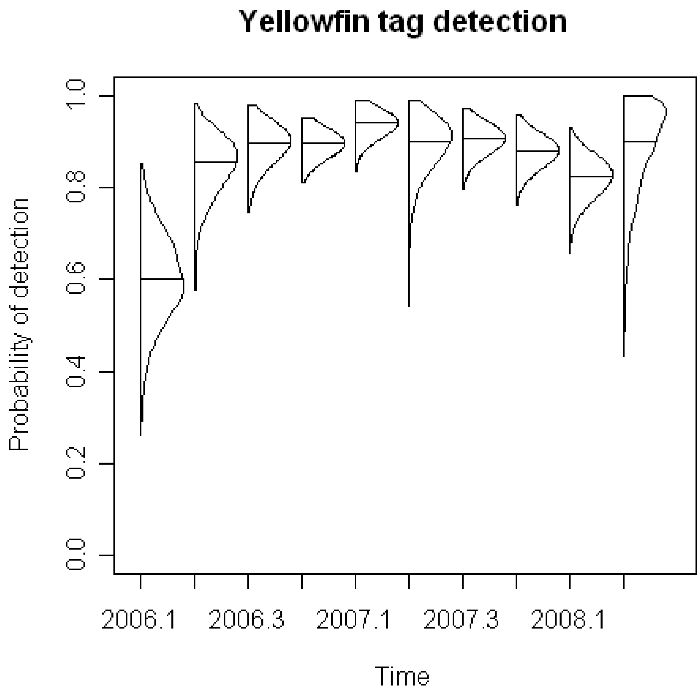


Figure 2: Skipjack reporting rate posteriors (and median) for the skipper-seeded data for 2006-2008.

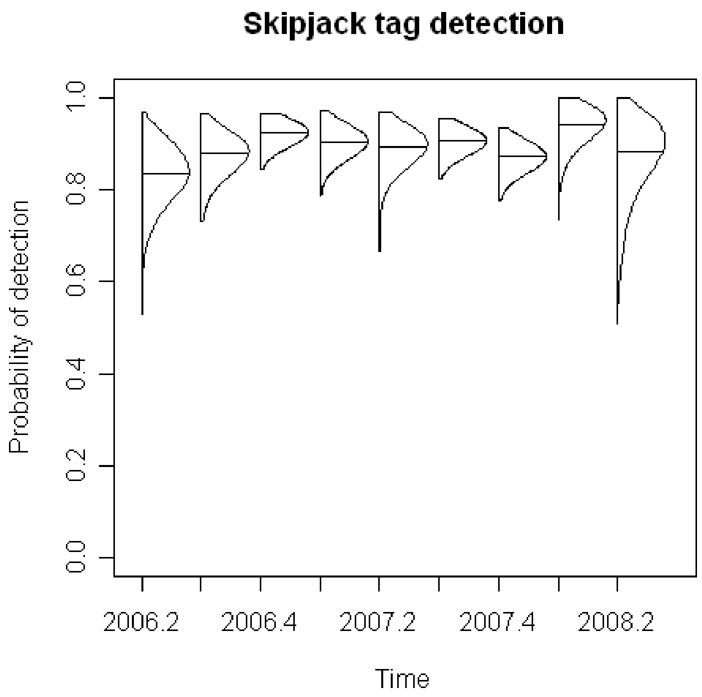


Figure 3: Bigeye reporting rate posteriors (and median) for the skipper-seeded data for 2006-2008.

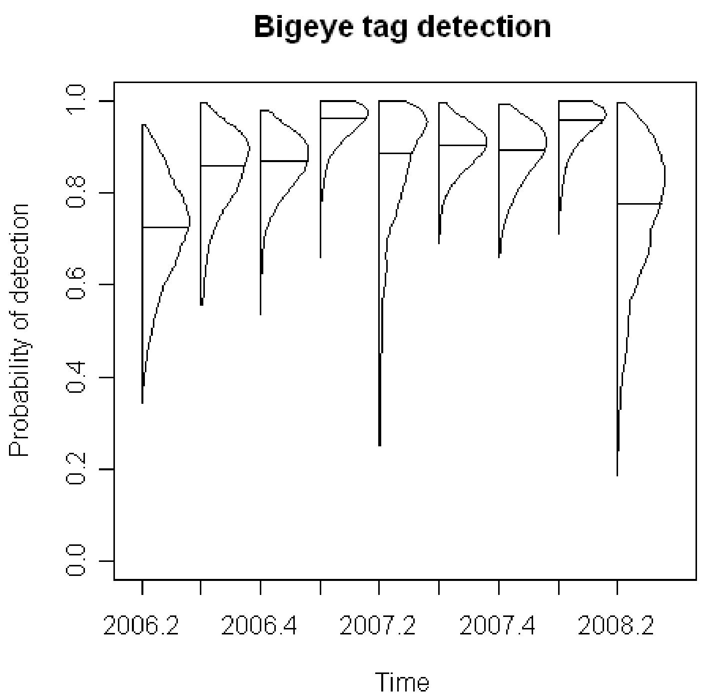
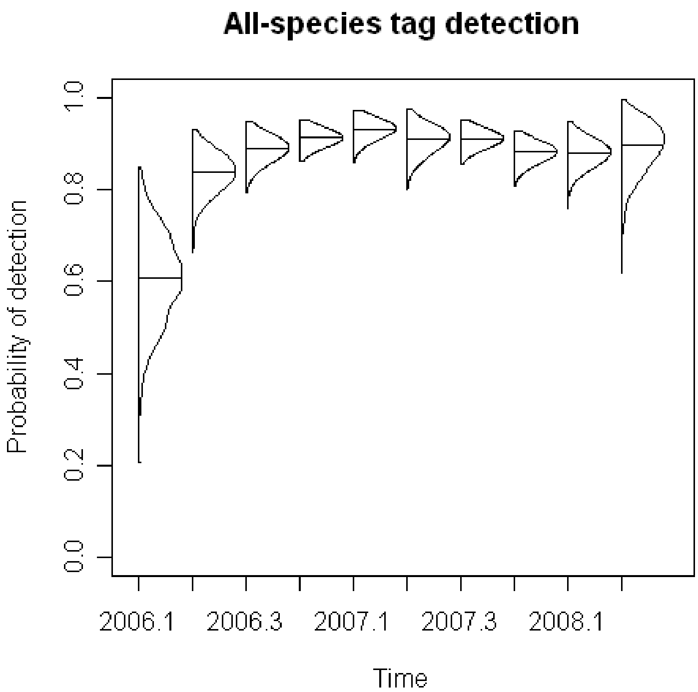


Figure 4: All-species reporting rate posteriors (and median) for the skipper-seeded data for 2006-2008.


All-species tag detection
Probability of detection
0.0
0.2
0.4
0.6
0.8
1.0
2006.1
2006.3
2007.1
2007.3
2008.1
Time