SCIENTIFIC COMMITTEE
SIXTEENTH REGULAR SESSION

ELECTRONIC MEETING
11-20 August 2020

ANALYSIS OF TAG SEEDING DATA AND REPORTING RATES FOR PURSE SEINE FLEETS

WCPFC-SC16-2020/SA-IP-04

T. Peatman[1]

[1] Independent fisheries consultant for the Pacific Community (SPC), Ocean Fisheries Programme (OFP)

Executive summary

Reporting rate models were constructed based on the recommended approach of Peatman et al. (2019). Models were fitted to data from tag seeding experiments on purse seine vessels and used to estimate flag-specific reporting rates.

Model runs with year as a categorical variable suggested an apparent step change in reporting rates, with lower levels of reporting from 2015 onwards. The models detected strong between-flag variation in reporting rates. There were relatively few seeding experiments in the modelled dataset for vessels flagged to China (CN), Ecuador (EC), FSM (FM), New Zealand (NZ), the Solomon Islands (SB), El Salvador (SV) and Vanuatu (VU), resulting in lower precision in effects for these flags. Region-specific reporting rate distributions for the 2014 regional structure were estimated.

Specific recommendations for the tag seeding experiments and analysis are:

- Tag seeding should be continued as long as regular tag recoveries are being received, targeted to fleets and regions where these regular recoveries are most likely;
- The current low levels of tag seeding experiments have compromised the ability to explore in detail what might be driving apparent recent reductions in tag reporting, and to robustly detect the timing and strength of the apparent reductions. This impact of temporal changes in reporting is exacerbated by imbalance of tag seeding data with respect to fleet-specific coverage through time;
- More consistent coverage of tag seeding experiments through time is suggested, with a particular emphasis on fleets that are likely to be recovering tags based on their areas of operation relative to PTTP tag releases;
- More tag seeding experiments are required each year in order to robustly detect temporal changes in reporting. A minimum target of 32 seeding experiments per annum is suggested.

Introduction

SPC have tagged and released tunas in the Western Central Pacific Ocean (WCPO) since 1977, across three tagging programmes: the Skipjack Survey and Assessment Programme (SSAP), 1977 to 1981; the Regional Tuna Tagging Programme (RTTP), 1989 to 1992; and, the current Pacific Tuna Tagging Programme (PTTP), since 2006. Tag seeding experiments have been undertaken as a component of both the RTTP and PTTP, in which observers on purse seiners surreptitiously mark captured tuna with conventional plastic tags, thereby 'seeding' the catch with tagged fish. Throughout the report, 'tag seeding experiment' refers to an observer trip on a specific fishing vessel during which tags were seeded.

The MULTIFAN-CL stock assessments of WCPO tuna stocks account for recovered tags that are not detected and/or reported to SPC using reporting rate parameters, defined by the proportion of recovered tags that are detected and reported. Incorporation of reporting rates addresses systematic under-estimation of fishing mortality rates and over-estimation of stock biomass due to underreporting of tag recoveries. Historically, purse seine tag reporting rates for MULTIFAN-CL assessments have been estimated using tag seeding experiments, using the proportion of seeded tags that are subsequently reported to SPC (e.g. Hampton 1997; Berger et al., 2014).

This information paper estimates reporting rate priors based on tag seeding experiments for application in the 2020 tuna stock assessments.

Methods

Reporting rate models were constructed based on the approach of Peatman et al. (2019). In summary, models were fitted to data from tag seeding experiments and used to estimate flag-specific reporting rates. Random samples were drawn from each flag-specific reporting rate distribution, and these were combined to estimate reporting rate prior distributions for the yellowfin and bigeye assessment regions using weighted averages of flag-specific reporting rates.

Tag release and recovery information were extracted from SPC's master tuna tagging database for all tag seeding experiments undertaken from 2007 to 2018 inclusive (Table 1, Table 2). Tag seeding experiments from 2019 onwards were excluded to ensure sufficient time for seeded tags to be detected and reported to SPC and thus minimise downwards bias in reporting rates. Since 2009, observers have recorded whether they believed that fishing vessel crew had seen the seeding of tags, or whether crew had asked questions that suggested that they were aware that tag seeding had taken place, i.e. whether the tag seeding experiment was likely to have been compromised. Reporting rates on fishing vessels are higher from compromised seeding experiments (Peatman et al., 2016). The dataset for the reporting rate models was filtered to remove seeding experiments where observers did not provide information on whether they considered the experiment to have been compromised. This left data from 249 seeding experiments, representing 5,318 seeded tags from which 3,032 recoveries were reported to SPC.

Beta-binomial models of reporting rates were fitted in R version 3.6.1 (R Core Team, 2019) using the package gamlss (Rigby and Stasinopoulos, 2005). We used the reporting rate model specification from Peatman et al. (2019) as a starting point. We tested alternative approaches to modelling temporal variation in reporting, including inclusion of the year of seeding as a categorical variable, and inclusion of year as a continuous variable with splines to test for non-linear effects (see Results for more information). Reporting rate models were fitted to tag seeding data aggregated across all species. We used AIC to choose the specification of temporal variation in reporting. The final model specification was

$$E[rec_t] = rel_t \mu_t$$

$$\mathrm{Var}(rec_t) = rel_t \mu_t (1 - \mu_t) \left[1 + \frac{\sigma}{1+\sigma}(rel_t - 1)\right]$$

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_0 + flag_t + abnormalyear_t + compromised_t$$

with: μ_t the reporting rate for seeding experiment t; rel_t and rec_t, the total number of seeded tags and reported recoveries, respectively; β_0, the global intercept; $flag_t$, a categorical variable for vessel flag; $compromised_t$, a categorical variable for whether available information suggested that the seeding experiment was likely compromised (‘*seen’* - the crew saw the observer seeding tags, the crew asked the observer questions about the seeding experiment, or the observer was uncertain as to whether they had been seen), or that seeding had likely taken place without the knowledge of the crew (‘*not seen*’); $abnormalyear_t$, a categorical variable for whether reporting rates were abnormal for the year in question (i.e. 2015 and onwards); and, σ an overdispersion parameter.

Flag-specific reporting rate distributions were generated from the fitted model by drawing 10,000 sets of parameters from the multivariate normal distribution $N_k(\boldsymbol{\beta}, \boldsymbol{\Sigma})$, defined by the vector of estimated parameter means $\boldsymbol{\beta}$ and their covariance matrix $\boldsymbol{\Sigma}$, where k is the number of estimated parameters. These parameter sets were then applied to each combination of flag and time-period (i.e. *abnormalyear* factor levels) to generate 10,000 reporting rate estimates for each combination. The $compromised$ variable was set to ‘*not seen*’ in predictions, to give reporting rate estimates for uncompromised seeding experiments. Region-specific reporting rate distributions for the time periods 2006 to 2014 and 2015 to 2018 were obtained by taking species-specific catch-weighted averages of the flag-specific reporting rates across the relevant years. Region-specific reporting rate distributions for the duration of the PTTP were obtained by taking weighted averages of the flag-specific reporting rates for the time periods 2006 to 2014 and 2015 to 2018, weighted by the product of the total species-specific PTTP tag releases and species and flag-specific catches across the relevant years. This approach ensures that flags with higher catches contribute more to region-specific reporting rates, and the same for time periods with more PTTP tag releases. The mean and variance of the region-specific reporting rate distributions were then extracted, with the penalty parameter given by $penalty = (2 * variance)^{-1}$. Flags that did not contribute a minimum of 1% to the total catch for any assessment region were excluded. Note that, as the reporting rate models are not species specific, the differences in reporting rate priors between species result from differences in the catch weightings.

At the time of writing[1], two regional specifications are being considered for the 2020 bigeye and yellowfin assessments: the 9 region model used in the 2014 bigeye and yellowfin assessments with the southern boundary of regions 1 and 2 at 20°N, referred to as the ‘2014 regional structure’ (e.g.

[1] Subsequent discussions at the SPC Pre-assessment Workshop (PAW) recommended the ‘2017 regional structure’ as the sole basis for the 2020 assessments.

see Davies et al., 2014); and, the 9 region model explored in the 2017 bigeye and yellowfin assessments with the southern boundary of regions 1 and 2 at 10°N, referred to as the '2017 regional structure' (e.g. see Tremblay-Boyer et al., 2017). We calculated reporting rate prior parameters for both regional structures, though the priors are insensitive to the change in the southern boundary of regions 1 and 2 given the low levels of purse seine effort between 10°N and 20°N.

Results

Model runs with year as a categorical variable suggested an apparent step change in reporting rates, with lower levels of reporting from 2015 onwards (Figure 1). Including year as a categorical variable increased the AIC relative to a model with no temporal effects (Table 3, ΔAIC = 6.0). Introducing a step-change in reporting rates in 2015 (i.e. defining 'abnormal years' as 2015 onwards) gave the strongest reduction in AIC (Table 3, ΔAIC = -5.7) and so was included in the final model specification used to generate reporting rate prior parameters. We note that the model specification from Peatman et al. (2019), with 'abnormal years' defined as 2015 and 2017, also lead to a similar reduction in AIC (Table 3, ΔAIC = -5.5). No significant continuous temporal changes in reporting rates were detected, though the inclusion of year as a spline did result in a modest reduction in AIC compared to a model with no temporal effects (Table 3, ΔAIC = -0.9). The relatively low numbers of tag seeding experiments from 2015 onwards resulted in relatively low precision in temporal effect estimates (e.g. see Figure 2).

Tag reporting rates from 2015 onwards were lower than for the period 2009 to 2014 (Figure 2, $p = 0.005$). Tag seeding experiments considered likely to be compromised were associated with slightly higher reporting rates, but the effect was not significant (Figure 2, $p = 0.14$). The models detected strong between-flag variation in reporting rates (Figure 2). There were relatively few seeding experiments in the modelled dataset for vessels flagged to China (CN), Ecuador (EC), FSM (FM), New Zealand (NZ), the Solomon Islands (SB), El Salvador (SV) and Vanuatu (VU) (Table 2), resulting in lower precision in effects for these flags (Figure 2).

We also tested reporting rate models with a $compromised_t$ effect in the specification of the overdispersion parameter, i.e. allowing flexibility for the model to have varying mean reporting rates and overdispersion for seeding experiments considered likely to be compromised. The additional parameter resulting in a marginal improvement in AIC (ΔAIC < 0.1), but the effect was not statistically significant. We did not pursue this further, but it could be looked at again in future analyses.

For the estimation of flag-specific reporting rate distributions, the Japanese flag effect was considered unlikely given the reported recoveries from the fleet, and so the Taiwanese flag effect was applied to the Japanese fleet as assumed in previous analyses (e.g. see Berger et al. 2014). In the absence of empirical data, reporting rates for EU Spanish vessels were assumed to be the same as those for flagged to Ecuador (e.g. see Berger et al., 2014). Flag-specific reporting rate distributions are provided in Figure 3. The resulting region-specific reporting rate distributions for the 2014 regional structure are provided in Figure 4 and Figure 5. Reporting rate prior parameters for the 2014 regional and 2017 regional structures are provided in Table 4. Note that a normal prior is currently used for reporting rates in MFCL. The reduction in reporting rates in region 8 from 2015 onwards is weaker than for regions 3 and 4, due to increases in catch proportions in region 8 from flags with high reporting rates.

References

Berger A.M., S. McKechnie, F. Abascal, B. Kumasi, T. Usu and S.J. Nicol (2014). Analysis of tagging data for the 2014 tropical tuna assessments: data quality rules, tagger effects, and reporting rates. WCPFC-SC10-SA-IP-06.

Davies, N., S. Harley, J. Hampton and S. McKechnie (2014). Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC10-2014/SA-WP-04.

Hampton, J. (1997) Estimates of tag-reporting and tag-shedding rates in a large-scale tuna tagging experiment in the western tropical Pacific Ocean. *Fishery Bulletin*, **95**, 68-79

Peatman, T., S. Caillot, B. Leroy, S. McKechnie F. Roupsard, C. Sanchez, S. Nicol and N. Smith (2016). Analysis of tag seeding data and reporting rates. WCPFC-SC12-2016/SA-IP-13.

Peatman, T., J. Scutt Phillips, F. Roupsard, C. Sanchez, B. Leroy and N. Smith (2019). Analysis of tag seeding data and reporting rates. WCPFC-SC15-2019/SA-IP-06.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rigby R.A. and D.M. Stasinopoulos (2005). Generalized additive models for location, scale and shape,(with discussion). *Appl. Statist.*, **54**, 507-554.

Tremblay-Boyer, L., S. McKechnie, G. Pilling and J. Hampton (2017). Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC-SC13-2017/SA-WP-06.

Tables

Table 1 Total tag seeding experiments per year, and tag seeding experiments per year in the dataset used to estimate reporting rate priors.

Year	Total experiments	Experiments in modelled dataset
2007	12	0
2008	14	0
2009	22	2
2010	17	0
2011	47	33
2012	77	73
2013	80	74
2014	30	29
2015	19	18
2016	15	8
2017	9	5
2018	7	7

Table 2 Total tag seeding experiments in the modelled dataset by year and flag.

Year	CN	EC	FM	JP	KI	KR	MH	NZ	PG	PH	SB	SV	TW	US	VU
2009	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
2011	0	0	1	1	0	13	2	0	3	4	3	0	1	5	0
2012	1	1	2	3	8	21	2	1	7	3	1	3	5	13	2
2013	0	0	0	3	5	15	4	3	26	9	0	0	0	9	0
2014	0	2	0	0	0	5	3	1	7	2	1	0	3	5	0
2015	0	0	0	3	4	0	0	0	0	1	1	1	2	6	0
2016	0	0	0	0	2	0	1	0	0	0	0	0	1	4	0
2017	0	0	0	0	3	0	0	0	0	0	0	0	0	2	0
2018	0	0	1	0	0	0	0	1	1	0	0	0	0	4	0
Total	**1**	**3**	**4**	**10**	**22**	**54**	**12**	**6**	**46**	**19**	**6**	**4**	**12**	**48**	**2**

Table 3 AIC for different specifications of temporal effects in the reporting rate model, and the change in AIC (ΔAIC) relative to a model with no temporal effects.

Formula	df	AIC	ΔAIC
$\sim flag + compromised$	17	1367.9	-
$\sim flag + compromised + \text{factor}(year \in \{2015, \dots, 2018\})$	18	1362.2	-5.7
$\sim flag + compromised + \text{factor}(year \in \{2015, 2017\})$	18	1362.4	-5.5
$\sim flag + compromised + \text{pb}(year)$	18.00001	1367.0	-0.9
$\sim flag + compromised + \text{factor}(year)$	25	1373.9	6.0

Table 4 PTTP reporting rate prior distribution parameters for purse seine fisheries (all flags), for the 2017 regional structure. Reporting rate priors are provided for the time period 2006-2014 ('pre-2015'), 2015-2016 ('post-2018'), and the period 2006-2018 ('combined PTTP'). Reporting rate parameters for the 2014 regional structure were identical.

		pre-2015		post-2015		combined PTTP	
Species	**Region**	**Mean**	**Penalty**	**Mean**	**Penalty**	**Mean**	**Penalty**
Bigeye	3	0.5601	378	0.4522	193	0.5517	410
	4	0.6283	433	0.4197	264	0.6120	478
	8	0.6987	684	0.6550	186	0.6952	727
Yellowfin	3	0.5714	392	0.4638	177	0.5677	404
	4	0.5584	849	0.3981	247	0.5529	897
	8	0.7110	832	0.6484	177	0.7089	846

Figures

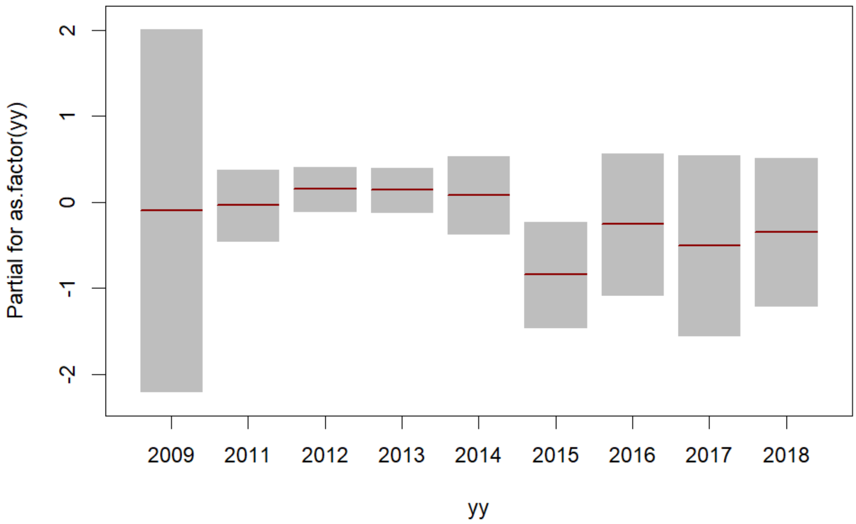


Figure 1 Year effects from the reporting rate model with inclusion of year as a categorical variable (instead of $abnormalyear_t$).

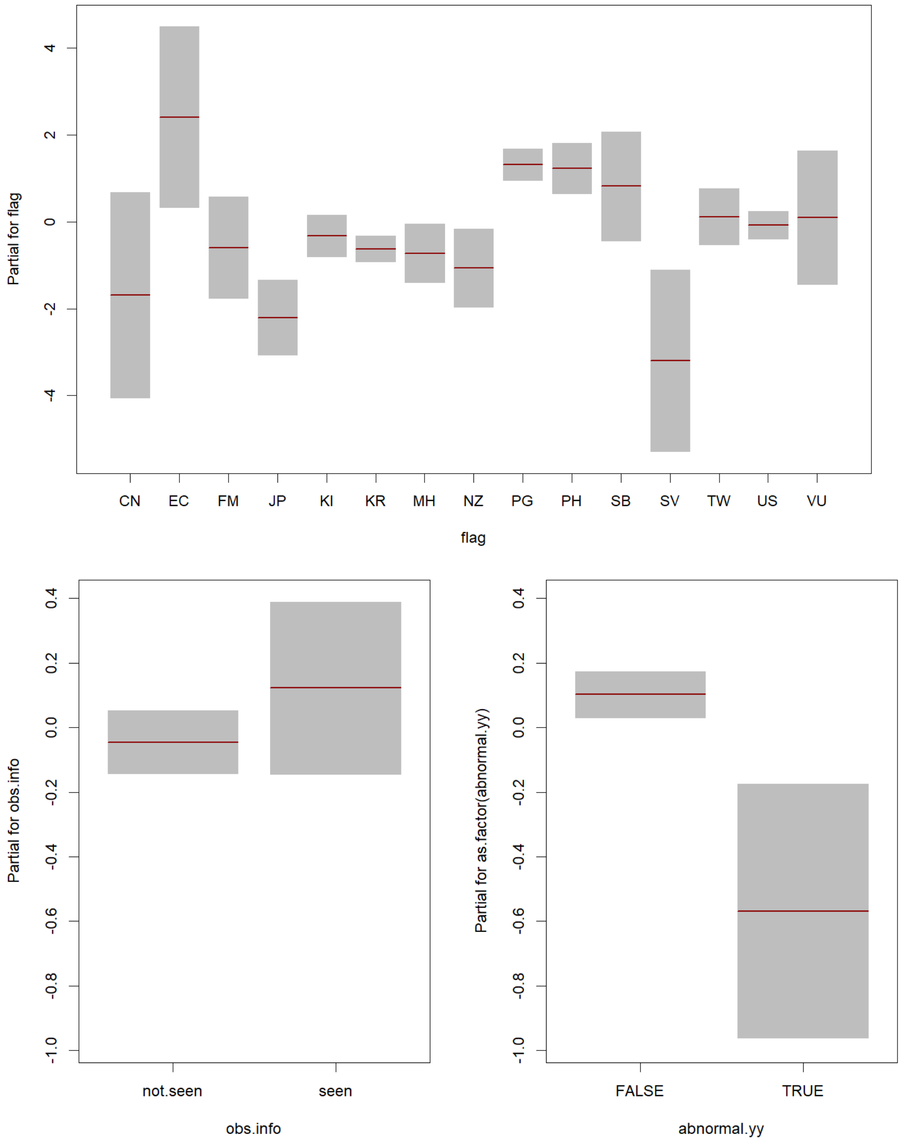
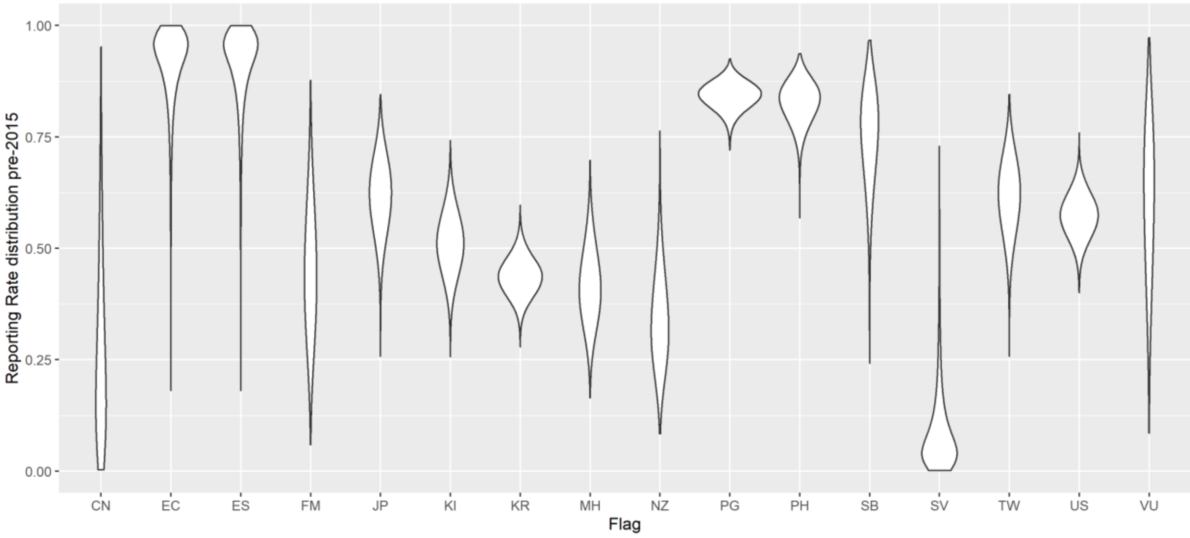


Figure 2 Effect plots for the model of reporting rates: flag (top); whether available information suggested the seeding experiment was compromised ('seen') or not ('not seen') (bottom left); and, whether the seeding experiment took place from 2015 onwards ('TRUE'). The scale of the y-axis is logit transformed reporting rate.

a) 2006-2014



a) 2015-2018

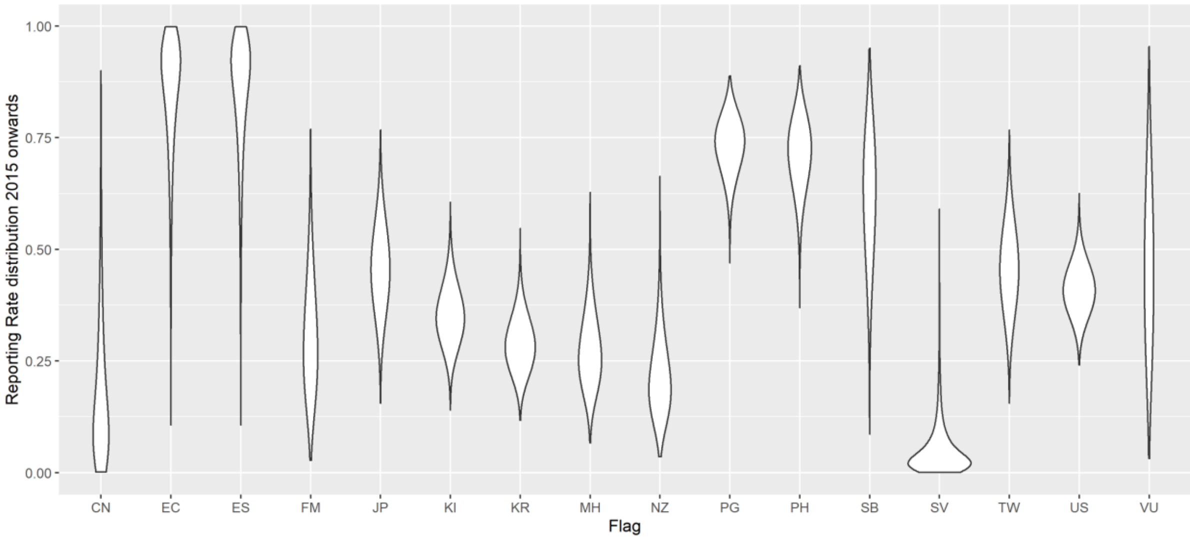
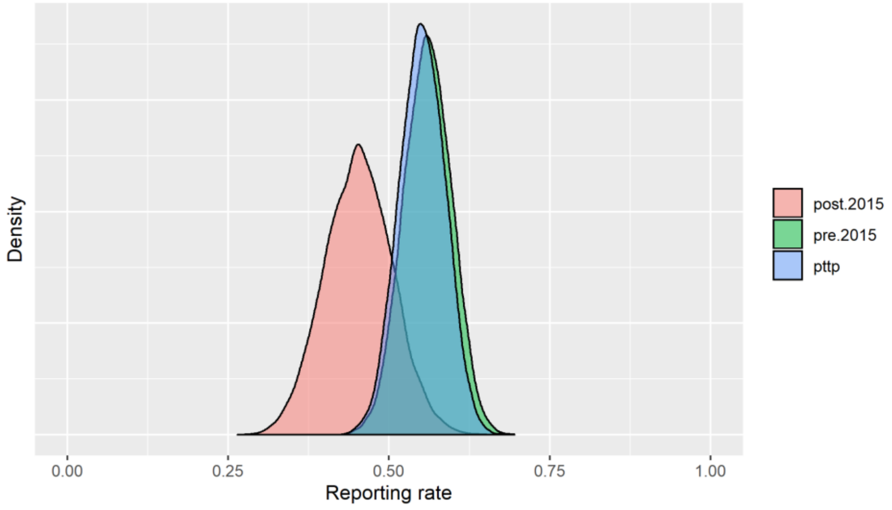
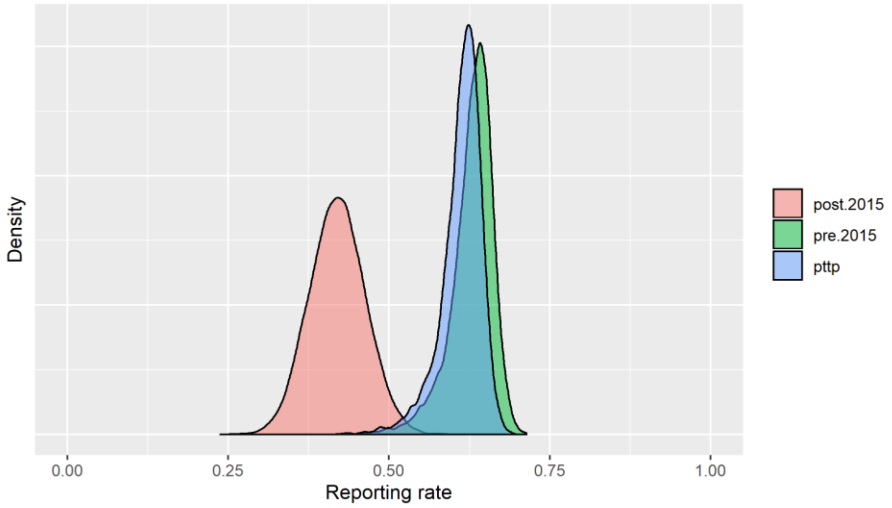


Figure 3 Flag specific reporting rate distributions used to calculate reporting rate prior parameters for a) 2006-2014 and b) 2015-2018.

a) Region 3



b) Region 4



c) Region 8

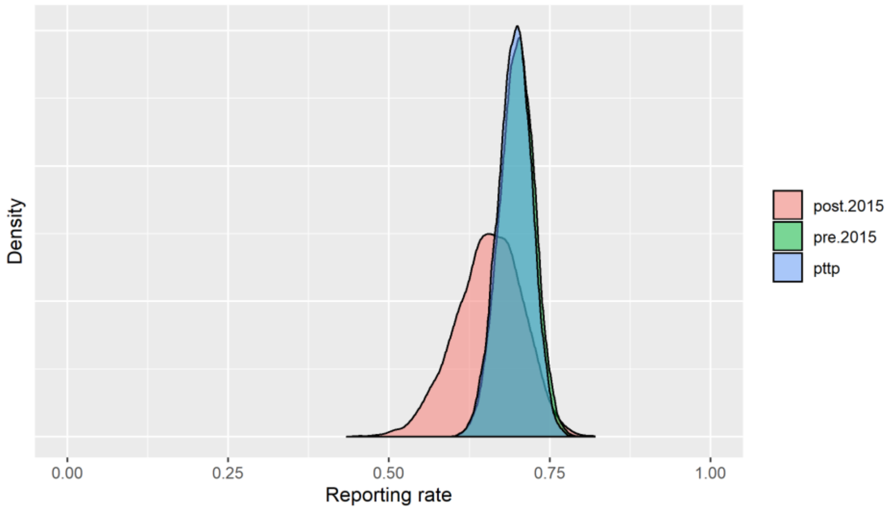
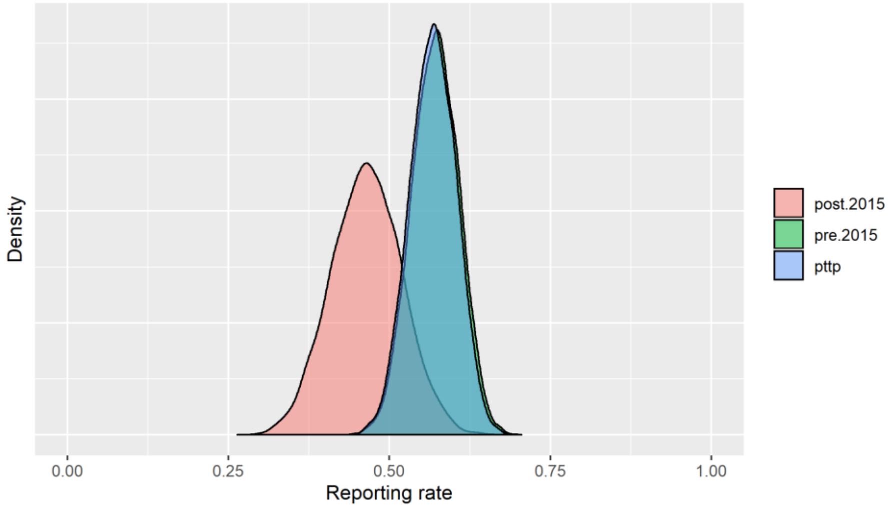
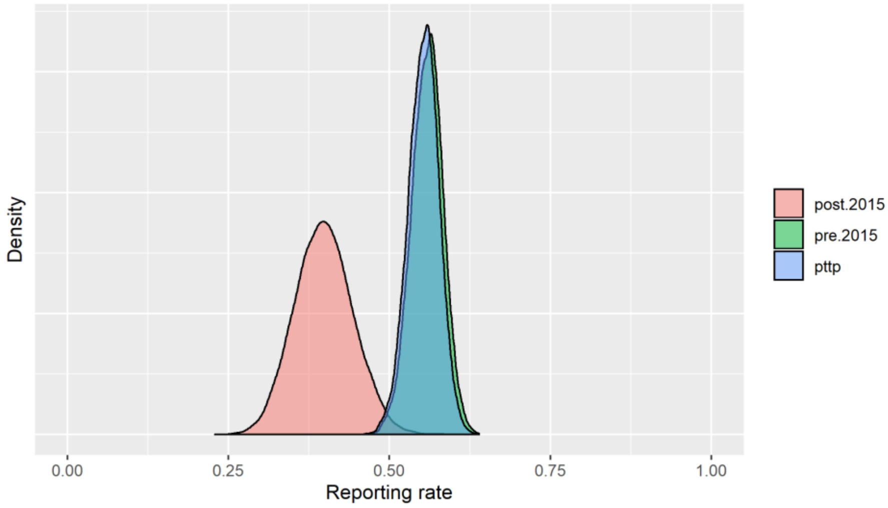


Figure 4 Region-specific reporting rate distributions for bigeye with the 2014 regional structure for a) region 3, b) region 4 and c) region 8.

a) Region 3



b) Region 4



c) Region 8

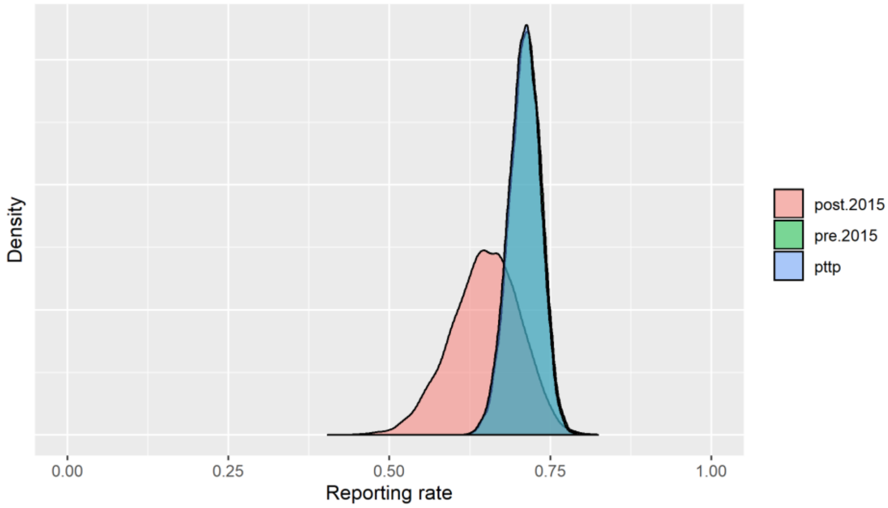


Figure 5 Region-specific reporting rate distributions for yellowfin with the 2014 regional structure for a) region 3, b) region 4 and c) region 8.