SCIENTIFIC COMMITTEE
EIGHTEENTH REGULAR SESSION

ELECTRONIC MEETING
10-18 August 2022

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

WCPFC-SC18-2022/SA-IP-19

T. Peatman[1]

---

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

Executive summary

Reporting rate models were constructed based on the approach of Peatman et al. (2020), fitted to data from tag seeding experiments on purse seine vessels. The reporting rate models were used to estimate flag-specific reporting rates. Flag-specific reporting rates were combined to generate reporting rate distributions for purse seine fisheries in the 2022 skipjack assessment, from which reporting rate prior parameters were extracted.

The analyses presented here provide the strongest evidence yet of a change in tag reporting during the PTTP, with an apparent reduction in reporting rates from 2015 onwards. However there remains considerable uncertainty around the structure, strength and timing of any change in reporting rates due to the limited number of tag seeding experiments conducted from 2015 onwards. The evidence for a temporal change in reporting rates during the PTTP does not appear sufficiently strong to support the inclusion of time-varying reporting rates in the 2022 skipjack assessment, given the additional flexibility that this would give the assessment model. Higher levels of tag seeding experiments are required to enable more robust monitoring of temporal changes in reporting rates in the future, and to provide more confidence that reporting rates are appropriately represented in stock assessment models.

We invite the Scientific Committee to consider the following recommendations for the tag seeding experiments and analysis:

- The Scientific Committee note that the continuing 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 estimate the timing and strength of these apparent reductions. The low level of seeding experiments is exacerbated by the imbalanced nature of the tag seeding data with respect to fleet-specific coverage through time;
- 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;
- A minimum target of 32 seeding experiments per year is recommended (see Peatman et al., 2019);
- More consistent coverage of tag seeding experiments through time is recommended, with a particular emphasis on fleets that are likely to be recovering tags based on their areas of operation relative to PTTP tag releases.

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 caught tuna with conventional plastic tags, thereby 'seeding' the catch with tagged fish. Throughout this 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 fishery and tag programme specific reporting rates, i.e. the proportion of recovered tags that are detected and reported. Incorporation of reporting rates in the assessment models addresses systematic under-estimation of fishing mortality rates and over-estimation of stock biomass due to under-reporting of tag recoveries. Reporting rates are estimated within the assessment model and are constrained by reporting rate prior distributions which are provided as an input, based on either analysis of data external to the assessment model or more subjective determinations of plausible reporting rates. The priors penalise estimated reporting rates that are further away from the mean of the prior distribution, with the strength of the penalisation controlled by a penalty term. Historically, purse seine tag reporting rate prior distributions for MULTIFAN-CL assessments have been estimated using tag seeding experiments, using the proportion of seeded tags that are subsequently detected and 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 2022 skipjack stock assessment, based on the approach of Peatman (2020).

Methods

Tag release and recovery information were extracted from SPC's master tuna tagging database for all tag seeding experiments undertaken from 2007 to 2020 inclusive (Table 1, Table 2). Tag seeding experiments from 2021 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 in recent years. 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. The rates of detection and reporting of tags on fishing vessels are higher from compromised seeding experiments (Peatman et al., 2016). The analysed dataset for the reporting rate models was filtered to remove tag seeding experiments where observers did not provide information required to determine whether a tag seeding experiment was likely to have been compromised. This left data from 261 seeding experiments, representing 5,597 seeded tags from which 3,200 recaptures were reported to SPC.

Beta-binomial models of reporting rates were fitted in R version 4.1.1 (R Core Team, 2021) using the 'gamlss' package (Rigby and Stasinopoulos, 2005). We used the reporting rate model specification from Peatman (2020) as a starting point. All reporting models included categorical variables for vessel flag and whether available information suggested a tag seeding experiment was compromised. We tested alternative approaches to modelling temporal variation in reporting rates, including: no temporal effects, inclusion of the year of seeding as a categorical variable, inclusion of year as a continuous variable as a linear effect (with year standardised by its mean and standard deviation) or as a penalised spline smoother to test for non-linear effects, and step-changes in reporting (see Results for more information). Reporting rate models were fitted to tag seeding data aggregated across all species. Model selection was undertaken using AIC.

The selected model specification was

$$\mathrm{E}(rec_t) = rel_t \mu_t$$

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

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

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 (*TRUE* - the observer was aware that he was seen seeding tags by crew, the crew asked the observer questions about the seeding experiment, or the observer was uncertain as to whether or not they had been seen seeding tags), or that the observer considered it likely that they had seeded tags without the knowledge of the crew (*FALSE*); $period_t$, a categorical variable for the time period of the seeding experiment (i.e. whether the seeding experiment started before 2015, or from 2015 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 to generate 10,000 reporting rate estimates for each combination. The $compromised$ variable was set to '*FALSE*' 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 onwards were obtained by taking skipjack catch-weighted means of the flag-specific reporting rates across the relevant years. Region-specific reporting rate distributions for the duration of the PTTP were obtained by calculating the weighted mean of the flag and time period-specific reporting rates, weighted by the product of flag and time-period specific skipjack catch and time-period specific proportion of total PTTP skipjack tag releases (pre-2015 = 0.856; 2015 onwards = 0.144). 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. Catches of the domestic Indonesian and Philippines purse seine fisheries were also excluded, on the assumption that available tag seeding data are only representative of reporting rates for the distant water fishery in region 5 (see Figure 1 for the region structures).

Results and discussion

Previous analyses of tag seeding data have suggested an apparent reduction in reporting rates from 2015 onwards (e.g. Peatman, 2020). Model runs with year as a categorical variable continue to suggest an apparent step change in reporting rates, with lower levels of reporting from 2015 onwards (Figure 2). Including year as a categorical variable reduced the AIC slightly relative to a model with no temporal effects (Table 3, ΔAIC = 0.6). A number of approaches were considered for modelling temporal variation in reporting rates. Introducing a step-change in reporting rates in 2015, i.e. defining a time period effect as pre-2015 and 2015 onwards, gave the strongest reduction in AIC (Table 3, ΔAIC = 14.3) and so was included in the final model specification used to generate reporting rate prior parameters. Models with a linear and non-linear temporal effect were equivalent, giving a more modest reduction in AIC (ΔAIC = 8.3) with a significant linear reduction in tag reporting rates over time (coefficient = -0.273, p = 0.0013). The relatively low numbers of tag seeding experiments from 2015 onwards resulted in lower precision in temporal effects in the latter part of the time series (e.g. see Figure 2).

The selected reporting rate model estimated significantly lower tag reporting rates from 2015 onwards relative to pre-2015 levels (Figure 3; coefficient = -0.864, p < 0.001). Tag seeding experiments considered likely to be compromised were associated with slightly higher reporting rates, but the effect was not significant (Figure 3; coefficient = 0.052, p = 0.77). The models detected strong between-flag variation in reporting rates (Figure 3). 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 3).

We also tested reporting rate models with a $compromised_t$ effect in the specification of the overdispersion parameter, i.e. allowing a differing level of overdispersion for seeding experiments considered likely to be compromised. Inclusion of the additional overdispersion parameter was not supported by AIC (ΔAIC = -0.8).

The Japanese flag effect (Figure 3) was considered unlikely given the numbers of reported recoveries relative to other flags. The Taiwanese flag effect was applied to Japan when estimating flag-specific reporting rate distributions, as assumed in previous analyses (e.g. see Berger et al. 2014). We note that Japanese vessels unload catches in Japanese ports, in contrast to other purse seine fleets operating in the WCPO. As such reporting rate estimates for Taiwanese purse seiners, or indeed those of other purse seine fleets, may not reflect those for Japanese vessels due to differences in the supply chains of product between the fleets. In the absence of empirical data, reporting rates for EU Spanish vessels were assumed to be the same as those for Ecuadorean flagged vessels (e.g. see Berger et al., 2014), and reporting rates for purse seiners flagged to Nauru and Tuvalu were assumed to be the same as those for vessels flagged to Kiribati.

Flag-specific reporting rate distributions are provided in Figure 4. The resulting region-specific reporting rate distributions for the eight-region structure are provided in Figure 5, with reporting rate prior parameters provided in Table 4. The reduction in reporting rates in region 6 from 2015 onwards is weaker than for regions 7 and 8, due to increases in catch proportions in region 6 from flags with high reporting rates in the recent years.

Reporting rates for region 5 were generated using the approach from the 2019 skipjack assessment (Vincent et al., 2019), i.e. by estimating a reporting rate distribution for regions 7 and 8 combined and applying a 50% reduction to the penalty parameter. The percentage of purse seine catches from Japanese vessels in the distant water fisheries in region 5 are relatively high, which is reflected in the reduction applied to the penalty parameter. However, as noted in the Methods section, available tag seeding data may not be representative of reporting rates for the domestic Indonesia and Philippines fishery in region 5 due to differences in fishing vessel characteristics, product flows of catches through the supply chain etc. As such, we recommend that the reporting rate prior is only used for the distant water fishery in region 5.

In the 2019 assessment, the purse seine fishery in region 3 shared the same PTTP reporting rate as pole and line and longline fisheries. The majority (c. 75%) of purse seine catches in region 3 during the PTTP were from Japanese vessels, with the remainder caught by a range of fleets close to the southern limit of the fishery. Approximately 90% of the Japanese catches in the region were also made between 10 and 15°N. As such, we consider that the PTTP reporting rate parameters generated for the distant water fishery in region 5 could also be used for the purse seine fishery in region 3 of the 8 region structure.

Reporting rate models were fitted to tag seeding data from all three tropical tuna species combined, reflecting the assumption that reporting rates were species invariant. We fitted reporting rate models including a species effect and a random intercept for tag seeding experiment ID, to explore whether this assumption was likely to be violated. This model did not detect significant variation in reporting rates between species, though there was a suggestion of higher reporting rates for bigeye ($p = 0.13$). We note that the numbers of bigeye seeded with tags are relatively low (5% of the total), compared with 76 and 19% for skipjack and yellowfin respectively.

Observers record information that is used to determine whether a tag seeding experiment was likely to have been compromised, i.e. that the crew on the fishing vessel were aware that the observer was tagging fish. Detection and reporting of seeded tags on fishing vessels is significantly higher for tag seeding experiments considered likely to have been compromised (Peatman et al., 2016). However, the reporting rate models reported here did not detect a difference in overall reporting rates for ‘compromised’ seeding experiments, relative to experiments where crew on the fishing vessel were apparently unaware that tag seeding was taking place. The exclusion of tag seeding experiments missing information for the ‘compromised’ variable removed 63 of the 65 tag seeding experiments conducted from 2007 to 2010. The model selection process was repeated using the full tag seeding dataset, without the ‘compromised’ variable, to assess the sensitivity of the reporting rate model to this data filtering. The model with most support had the same temporal effects structure as for the filtered dataset, with a step-change in reporting rates in 2015 ($\Delta AIC = 13.5$ relative to a model with no temporal effect). The strength of the reduction in reporting rates (coefficient = -0.727) was slightly weaker than that from the model fitted to the filtered dataset (-0.864). The model with year included as a categorical variable provided some evidence that reporting rates may have been relatively low in 2008 and 2010 (Figure 6). Tag reporting rates in the earlier years of the PTTP may have been lower, for example if there were lower levels of awareness of the PTTP either on fishing vessels or elsewhere in the supply chain, and so lower rates of detection and reporting of tags. This appears unlikely given the efforts to raise awareness of the PTTP tagging programme in the region. The apparent low rates of reporting in the earlier years could also reflect less awareness of the tag seeding experiments, and so fewer compromised seeding experiments. Equally, the apparent variation in reporting rates in the early years of the PTTP could be due to the relatively low number of tag seeding experiments. Nevertheless, the models suggest that temporal variation in reporting rates during the PTTP may be more complex than simply a reduction in tag reporting rates from 2015 onwards.

The analyses presented here provide the strongest evidence yet of a change in tag reporting during the PTTP. However, the relatively low numbers of tag seeding experiments undertaken in recent years, coupled with the high levels of variation in tag reporting rates between seeding experiments, has compromised our ability to explore these temporal changes in detail, or be confident that a change in reporting rates has actually occurred. As such, there remains considerable uncertainty around the structure, strength and timing of any change in reporting rates. The evidence for a temporal change in reporting rates during the PTTP does not appear sufficiently strong to support the inclusion of time-varying reporting rates in the 2022 skipjack assessment, given the additional flexibility that this would give the assessment model. Instead, we recommend using reporting rate prior parameters calculated for time period 2006-2020, which take account of reduced reporting rates post-2015. Higher levels of tag seeding experiments are required to enable more robust monitoring of temporal changes in reporting rates in the future, and to provide more confidence that reporting rates are appropriately represented in stock assessment models.

Noting the decision to return to 100% purse seine observer coverage at the beginning of 2023 the SSP has implemented new incentives schemes within the national and regional observer programmes to encourage tag seeding experiments. In addition to incentives for observers these include incentives for officers involved in observer placement and debriefing. The SSP has also implemented new schemes for tag finders to encourage the return of found tags (see SC18-RP-PTTP-01 for details).

We invite the Scientific Committee to consider the following recommendations for the tag seeding experiments and analysis:

- The Scientific Committee note that the continuing 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 estimate the timing and strength of these apparent reductions. The low level of seeding experiments is exacerbated by the imbalanced nature of the tag seeding data with respect to fleet-specific coverage through time;
- 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;
- A minimum target of 32 seeding experiments per year is recommended (see Peatman et al., 2019);
- More consistent coverage of tag seeding experiments through time is recommended, with a particular emphasis on fleets that are likely to be recovering tags based on their areas of operation relative to PTTP tag releases.

Acknowledgements

T. Peatman's contribution was supported by the WCPFC and the European Union's "Pacific-European Union Marine Partnership Programme".

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.

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. (2020). Analysis of tag seeding data and reporting rates for purse seine fleets. WCPFC-SC16-2020/SA-IP-04.

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 (2021). 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.

Vincent, M. T., Y. Aoki, H. Kiyofuji, J. Hampton and G. M. Pilling (2019). Background analyses for the 2019 stock assessment of skipjack tuna. WCPFC-SC15-2019/SA-IP-04.

Tables

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

Year	Total experiments	Experiments in modelled dataset
2007	11	0
2008	15	0
2009	22	2
2010	17	0
2011	46	32
2012	78	74
2013	80	74
2014	30	29
2015	19	18
2016	15	8
2017	9	5
2018	7	7
2019	7	7
2020	6	5

Table 2 Tag seeding experiments in the modelled dataset by year and flag, having excluded seeding experiments missing information for the ‘compromised’ variable.

Year	CN	EC	FM	JP	KI	KR	MH	NZ	PG	PH	SB	SV	TW	US	VU	Total
2009	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	**2**
2011	0	0	1	1	0	12	2	0	3	4	3	0	1	5	0	**32**
2012	1	1	2	3	8	22	2	1	7	3	1	3	5	13	2	**74**
2013	0	0	0	3	5	15	4	3	26	9	0	0	0	9	0	**74**
2014	0	2	0	0	0	5	3	1	7	2	1	0	3	5	0	**29**
2015	0	0	0	3	4	0	0	0	0	1	1	1	2	6	0	**18**
2016	0	0	0	0	2	0	1	0	0	0	0	0	1	4	0	**8**
2017	0	0	0	0	3	0	0	0	0	0	0	0	0	2	0	**5**
2018	0	0	1	0	0	0	0	1	1	0	0	0	0	4	0	**7**
2019	0	0	1	1	0	0	0	0	2	1	1	0	0	1	0	**7**
2020	0	0	0	1	0	0	0	0	3	1	0	0	0	0	0	**5**
Total	**1**	**3**	**5**	**12**	**22**	**54**	**12**	**6**	**51**	**21**	**7**	**4**	**12**	**49**	**2**	**261**

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
~ *flag* + *compromised*	17	1437.1	0.0
~ *flag* + *compromised* + *year*	18	1428.8	8.3
~ *flag* + *compromised* + pb(*year*)	18	1428.8	8.3
~ *flag* + *compromised* + factor(*year*)	27	1436.4	0.7
~ *flag* + *compromised* + factor(*year* >= 2015)	18	1422.8	14.3

Table 4 PTTP reporting rate prior distribution parameters for purse seine fisheries (all flags), for the eight region structure. Reporting rate priors are provided for the time period 2006-2014 ('pre-2015'), 2015-2020 ('post-2015'), and the period 2006-2020 ('PTTP overall'). Reporting rate priors for 2006-2020 ('PTTP overall') are recommended for use in the 2022 skipjack assessment.

		PTTP pre-2015		PTTP 2015-onwards		PTTP overall	
Species	**Region**	**Mean**	**Penalty**	**Mean**	**Penalty**	**Mean**	**Penalty**
Skipjack	3	0.587	281	0.373	146	0.567	314
	5	0.587	281	0.373	146	0.567	314
	6	0.679	661	0.551	233	0.667	705
	7	0.589	395	0.399	224	0.571	423
	8	0.579	857	0.362	323	0.559	978

Figures

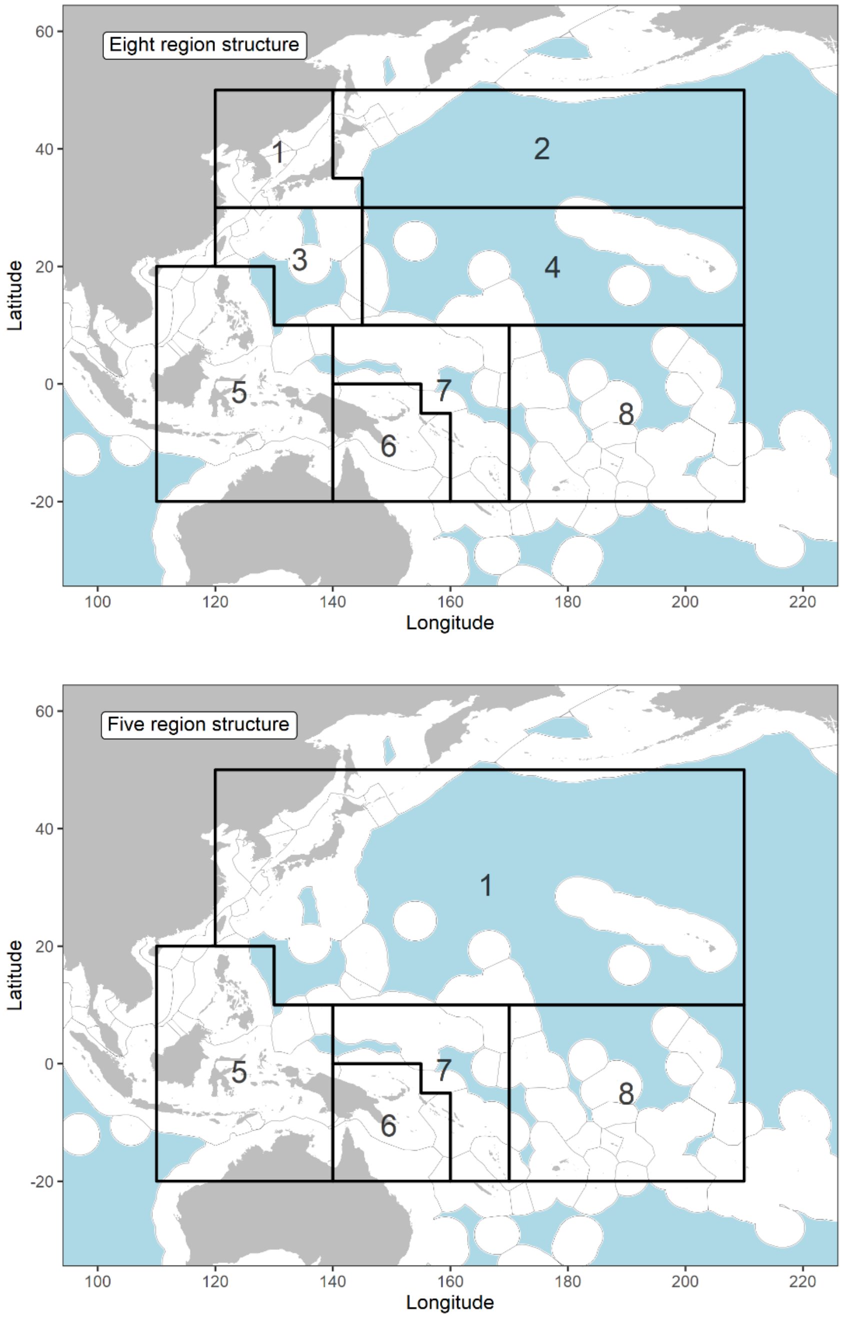


Figure 1 The eight region (top) and five region (bottom) structures used to generate reporting rate priors.

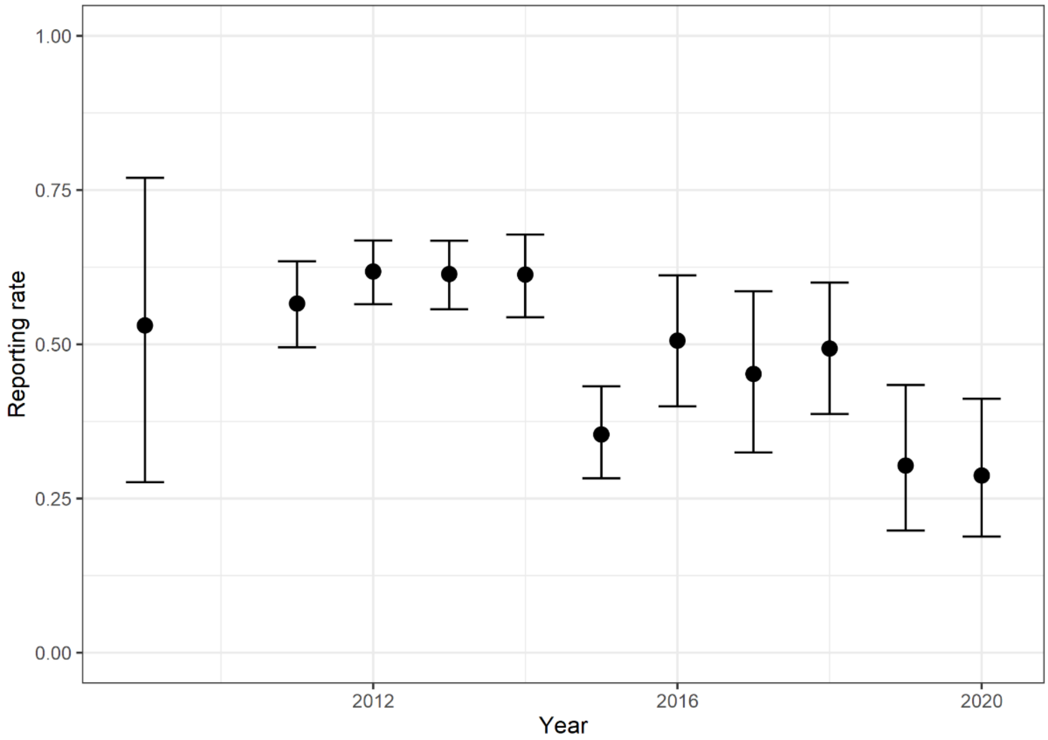


Figure 2 The effect of year (mean ± SE) on reporting rates when included as a categorical variable, for the reporting rate model with categorical variables for year, 'compromised' and flag. 'Compromised' was set to FALSE, with flag set to 'US'.

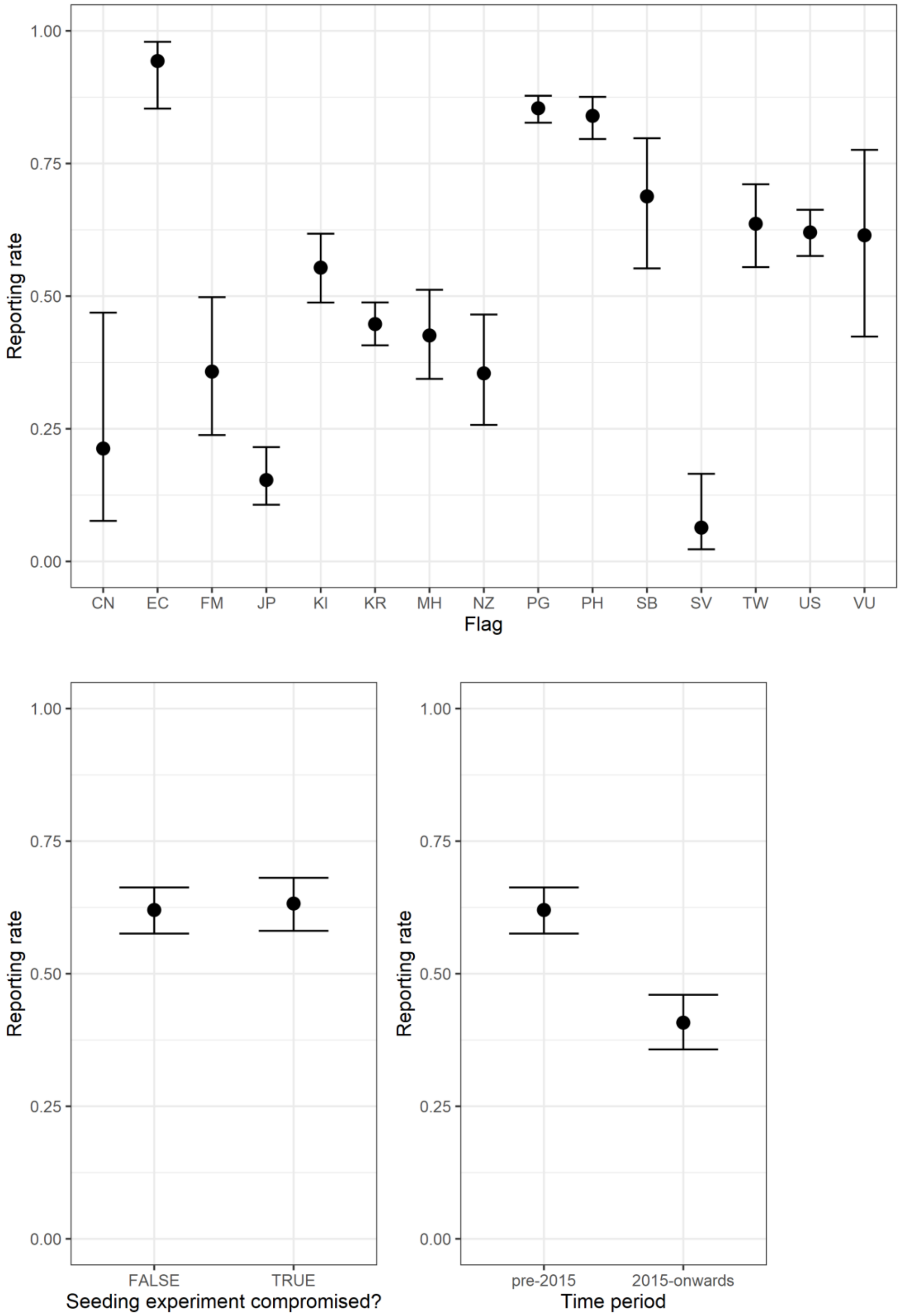
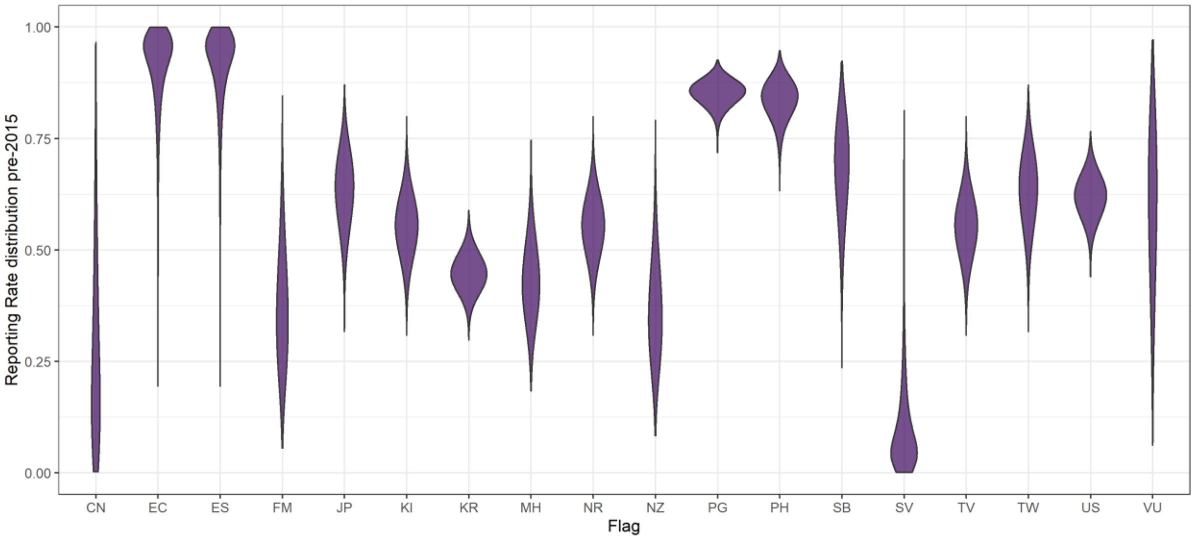


Figure 3 The effect of covariates on reporting rates (mean ± SE) for the selected reporting rate model with effects for flag (top panel), whether available information suggested the seeding experiment was compromised (bottom left), and the time period of the seeding experiment (bottom right). The effect of each covariate was estimated in turn by holding the remaining covariates constant at reference levels (flag = 'US', compromised = 'FALSE', and time period = 'pre-2015').

a) 2006-2014



a) 2015-2020

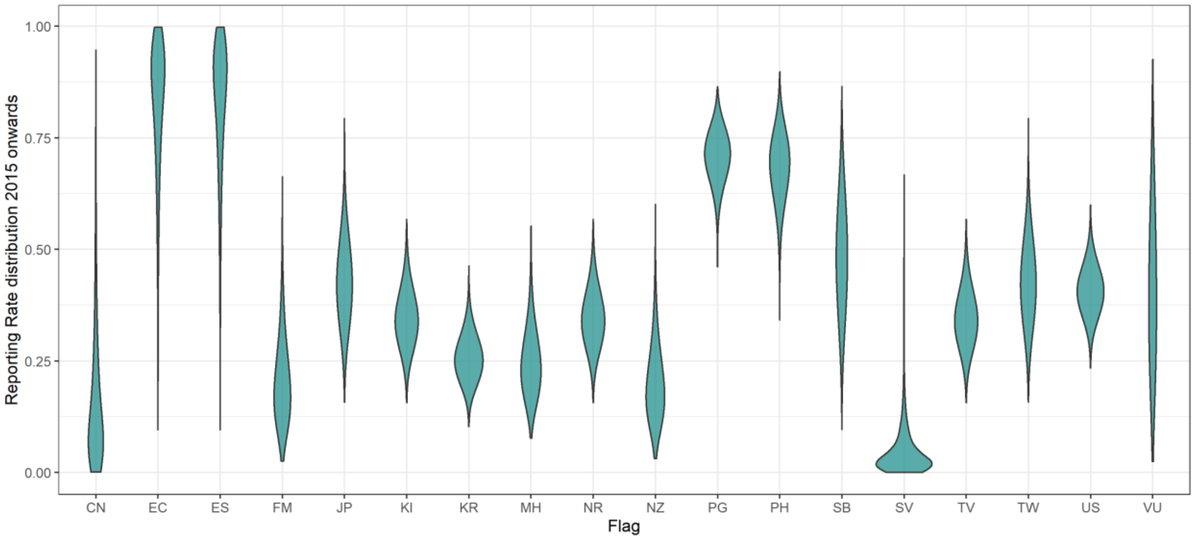
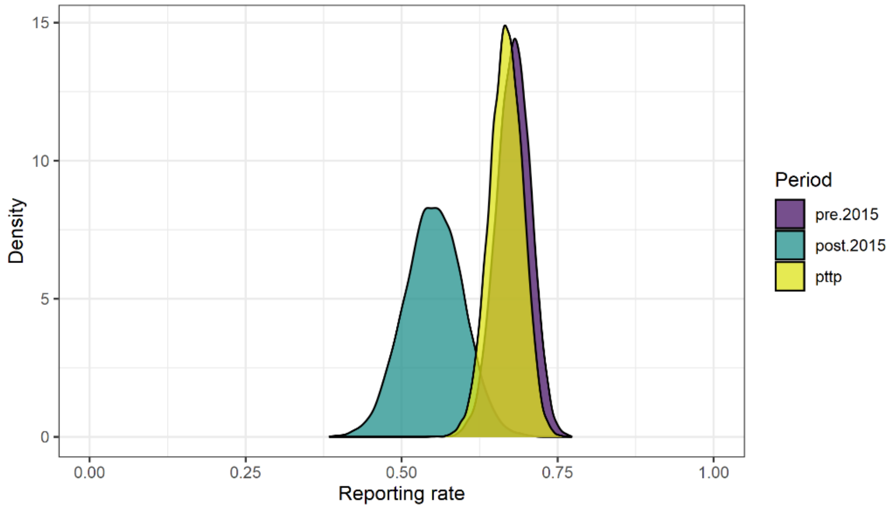
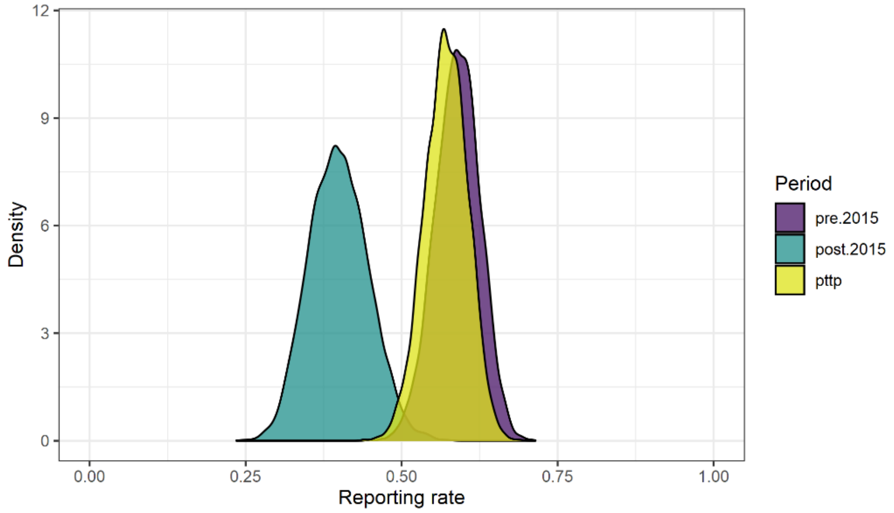


Figure 4 Flag specific reporting rate distributions used to calculate reporting rate prior parameters for a) 2006-2014 and b) 2015 onwards.

a) Region 6



b) Region 7



c) Region 8

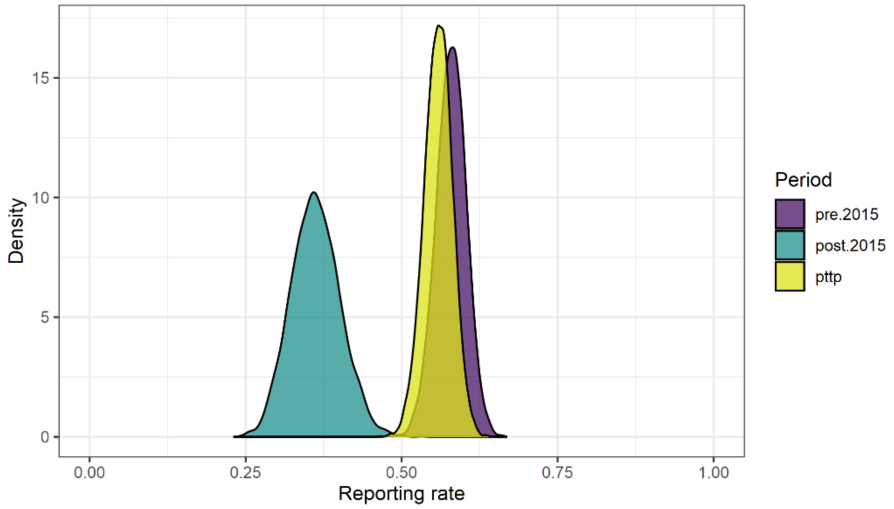


Figure 5 Region-specific reporting rate distributions for skipjack with the eight-region structure for a) region 6, b) region 7 and c) region 8.

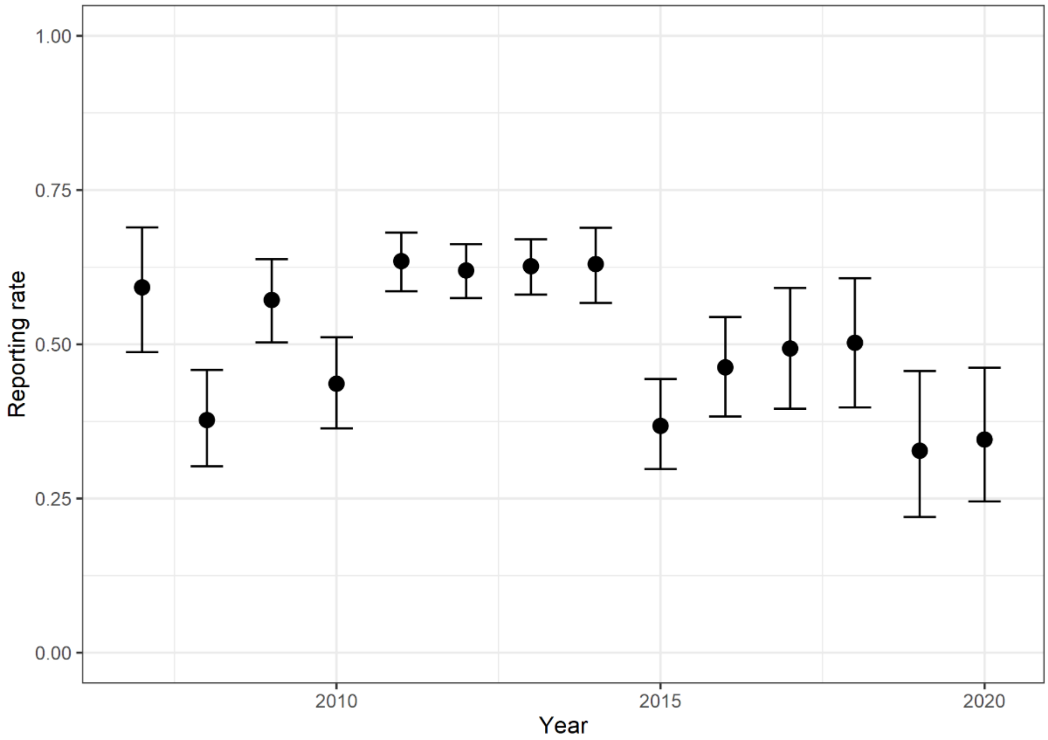


Figure 6 The effect of year (mean ± SE) on reporting rates when included as a categorical variable, for the reporting rate model with categorical variables for year, and flag, fitted to the full tag seeding dataset (i.e. including experiments missing information for the 'compromised' variable). Flag was set to 'US'.