SCIENTIFIC COMMITTEE

FIFTEENTH REGULAR SESSION

Pohnpei, Federated States of Micronesia

12-20th August 2019

Analysis of tag seeding data and reporting rates

WCPFC-SC15-2019/SA-IP-06

T. Peatman[1], J. Scutt Phillips[1], F. Roupsard[1], C. Sanchez[1], B. Leroy[1], N. Smith[1]

[1] Oceanic Fisheries Programme (OFP), Pacific Community, Noumea, New Caledonia

Executive summary

Data from tag seeding experiments have been used to estimate prior distributions for reporting rates for use in MULTIFAN CL (MFCL) assessments of tuna stocks in the Western Central Pacific Ocean. These prior distributions are used to minimise bias in assessments resulting from the non-reporting (or detection) of tag recoveries, and as such are a critical input to the assessment models.

Here, we update analyses of tag seeding data to attempt to increase confidence in reporting rate estimates with application to the 2019 skipjack assessment, and provide advice on the design of tag seeding experiments moving forward.

We refined the approach used to generate reporting rate prior estimates for MFCL models. This results in a substantial increase in penalty terms, due to a more appropriate estimation of uncertainty in flag-specific reporting rates. Data from seeding experiments suggests that tag reporting rates were significantly lower in 2015, and potentially 2017. However, it is difficult to explore this in detail due to the low number of seeding experiments in recent years, coupled with the unbalanced nature of the seeding dataset with respect to temporal coverage by flag.

Analyses of tag seeding data by tag type suggested that steel-headed tags may have lower rates of tag shedding than plastic dart tags in a tag seeding context, but no significant tag-type effects were detected (i.e. tagging of dead fish). We recommend that plastic dart tags be used exclusively in the future, as steel-headed tags may actually act to alert tag finders that tag seeding has taken place, given that steel-headed tags have only been used in the PTTP in seeding experiments.

A simulation model was developed to look at the number of seeding experiments required to detect step changes in reporting rates. The simulation model, whilst simplistic in nature, suggests that the current levels of seeding experiments are likely only sufficient to detect substantial and widespread changes in reporting rates. Increases in the numbers of seeding experiments per annum, and the coverage of fleets through time are required if tag seeding is to continue to be an effective component of the PTTP.

Specific recommendations for the tag seeding experiments and analysis are provided below:

- 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;
- Tag seeding experiments should be undertaken across as wide a range of vessels as possible within each fleet;
- Tag seeding experiments should use plastic dart tags, rather than a combination of plastic dart and steel-head tags;
- The number of seeding experiments should be increased from current levels. A minimum target of 32 seeding experiments per annum is suggested based on available information, noting that development of the simulation model would allow more robust advice to be provided in the future on future levels of seeding and how this effort should be distributed throughout the fishery.

1 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. In total, more than 430,000 tuna have been tagged and released through the PTTP, of which nearly 80,000 have been recovered and reported to SPC (SPC-OFP, 2019). Tag seeding experiments have been undertaken as a component of both the RTTP and PTTP, in which observers on purse seiners 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 under-reporting of tag recoveries. Historically, 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. Berger *et al.*, 2014).

Here, we attempt to increase confidence in reporting rate estimates with application to the 2019 skipjack assessment (Vincent et al., 2019a, b). Three analyses are presented:-

1. Updated estimates of reporting rate priors based on tag seeding experiments.
2. Tests for differences in tag-shedding between tag types, i.e. steel head vs. plastic dart tags.
3. A simulation exercise to inform the numbers of tag seeding experiments required to test for temporal changes in reporting rates.

2 Methods

2.1 Updated reporting rates

Reporting rate models were constructed based on the approach of Peatman et al. (2016). Flag-specific reporting rates were estimated based on tag seeding experiments. Random samples were drawn from each flag-specific reporting rate distribution, and these were combined to estimate reporting rate prior distributions for the skipjack assessment regions based on catch-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 2017 inclusive (Table 1). Tag seeding experiments from 2018 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 242 seeding experiments, representing 5,151 tags seeded from which 2,891 recoveries were reported to SPC.

Beta-binomial models of reporting rates, μ_t, were fitted using the R package gamlss (Rigby and Stasinopoulos, 2005)

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

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

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

with: rel_t and rec_t, the total number of seeded tags and reported recoveries, respectively, for seeding experiment t; 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 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 taken place without the knowledge of the crew (‘*not seen*’); and, $abnormalyear_t$, a categorical variable for whether reporting rates were abnormal for the year in question (i.e. ‘*TRUE’* for 2015 and 2017, and ‘*FALSE’* for other years); and, the over-dispersion parameter σ. We note that the $yeartype_t$ covariate was included as exploratory data analyses indicated that reporting rates for seeding experiments in 2015 and 2017 were lower than for other years, with no apparent variation in reporting rates for other years.

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. The $compromised_t$ variable was set to ‘*not seen*’, to give reporting rate estimates for uncompromised seeding experiments. The $abnormalyear_t$ effect was set to ‘*FALSE*’. We note that few tagged skipjack were likely to be available for recovery in 2015 as tagging cruises in 2014 and 2015 were bigeye and yellowfin focussed Central Pacific cruises (e.g. OFP-SPC, 2019), and 2017 PTTP tagging data were not included in the 2019 skipjack assessment (Vincent et al., 2019b). Region-specific reporting rate distributions were obtained by taking catch-weighted averages of the flag-specific reporting rate. The mean and variance of the region-specific reporting rate distributions were then extracted, with the penalty parameter given by $penalty = (2 * variance)^{-1}$.

2.2 Tests for differences in tag shedding between tag types

There has been concern that plastic dart tags may be shed more easily when inserted into dead fish, as the anchoring of the tag in the dorsal fin spines (the pterigyophores) and muscle will not be secured by the healing of the surrounding flesh (Leroy et al., 2009). Steel-headed tags are thought to achieve a more secure attachment in dead fish, and so shedding rates of steel-headed tags in a tag seeding context are thought to be low. However, plastic dart tags are the conventional tag type of choice in tagging cruises, and so the finding of steel-headed tags may act as an indicator that a seeding experiment has taken place. Steel-headed and plastic dart tags have both been used in tag seeding experiments in the PTTP. The usage of the different tag types has varied through time. There have been limited numbers of double tagged fish with one tag of each type. In recent years, observers have been instructed to tag two tuna with a single steel-headed tag, and two tuna with a single plastic dart tag, for each set where seeding is undertaken.

We analysed the tag seeding dataset in a variety of ways to test for differences in tag shedding between plastic dart and steel-headed tags. We used reporting rate as a proxy for tag shedding on the assumption that, all else being equal, any observed differences in reporting rates between tag types was likely to be a result of differences in shedding, and filtered the dataset for comparisons between tag types to reduce the chance that this assumption was violated. We also assumed that, for any tag reported recovery of a double tagged fish, the tag finders would have reported both tags as recovered if both tags were still present. This assumption appears reasonable as rewards are provided per tag reported, not per fish.

First, we fitted the model of reporting rates from Section 2.1 with the addition of a tag type effect, having excluded from the modelled dataset seeding experiments where only one tag type was used. This left data from 215 seeding experiments, with 3,740 seeded tags and 2,085 reported recoveries. Second, we compared tag-shedding rates using the double-tagged dataset from tag seeding experiments. There have been 634 fish that were double tagged with one tag of each type and reported as recovered, and 214 fish that were double tagged with either two plastic dart tags or two steel-headed tags and reported as recovered. We compared the proportions of tag recoveries with reporting of both tags, and fitted a logistic regression model to the data from all double-tagged fish that were reported as recovered to test for a tag type effect on the probability of a tag having been reported.

2.3 Simulation model to inform design of tag seeding experiments

A simulation exercise was undertaken to look at the number of seeding experiments required to detect step changes in reporting rates. We simulated data for two periods: a pre-change period, parameterised to reflect estimated reporting rates from the current tag seeding dataset (see Table 3 for a summary of parameters used); and, the post-change period, parameterised to reflect an assumed step-change in these reporting rates. We simulated 1,000 sets of observations, with varying numbers of tag seeding experiments in the post-change period, fitted simplified models of reporting rate to each set of observations, and extracted the proportion of models which correctly identified an increase or decrease in reporting rate.

The majority of flag-specific reporting rates appear to fall in to two groups: those with reporting rates of approximately 80 %; and those with reporting rates of approximately 50 %. We will refer to these two groupings as type A and type B respectively. There were approximately 100 seeding experiments for type A in the dataset used to fit the models of reporting rates, with 200 seeding experiments for type B. As such, the pre-change period observations were drawn from beta-binomial distributions, with parameters outlined in Table 3. The overdispersion parameter was set at the value estimated from the reporting rate model described in Section 2.1. For the post-change period, we applied either a 15% increase, a 15% decrease or a 30% decrease in reporting rates for both fleet types, and simulated observations for 8, 12, 16, 32, 64 and 128 experiments per fleet type. Issues with model convergence were encountered when attempting to fit models with fewer than 8 experiments in the post-change period, and so these incomplete results are not presented here.

The model fitted to the simulated observations was a simplified version of that specified in Section 2.1, with

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

where $type_t$ is a categorical variable for reporting rate types 'A' and 'B' and $period_t$ is a categorical variable for the pre-change and post-change period. The model can then account for differences in mean reporting rates for each combination of time period and reporting rate type. However, we note that this introduces difficulty in testing for significant differences in reporting rates for fleet types, as there is no single estimated parameter encapsulating the change in reporting rates for both fleet types in the post-change period.

3 Results

3.1 Updated reporting rates

There was strong between-flag variation in reporting rates (Figure 1). There was also relatively strong estimated extra-binomial variation in reporting rates, with an over-dispersion parameter of 0.48. Seeding experiments in 2015 and 2017 were associated with significantly lower reporting rates. Tag seeding experiments considered likely to be compromised were associated with slightly higher reporting rates, but the effect was not significant. There were less than 10 seeding experiments in the modelled dataset for vessels flagged to China (CN), Ecuador (EC), Fiji (FJ), FSM (FM), New Zealand (NZ), the Solomon Islands (SB), El Salvador (SV) and Vanuatu (VU), resulting in low precision in reporting-rate for these flags.

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 (Berger et al. 2014; Peatman et al. 2016). In the absence of empirical data, mean reporting rates for Spain were set to those of Ecuador. Flag-specific reporting rate estimates are provided in Figure 2. The resulting region-specific reporting rate penalty parameters are provided in Table 2. Note that a normal prior is used for reporting rates in MFCL.

3.2 Tests for differences in tag shedding between tag types

The addition of tag-type to the reporting rate model outlined in Section 2.1 suggested that steel-headed tags had slightly higher reporting rates, and so lower shedding rates, than plastic dart tags, but the difference was not significant ($p = 0.160$) (Figure 3).

There were 634 seeded fish that were reported as recovered and were double tagged with one steel-headed and one plastic dart tag. From these 634 fish, 585 steel-headed tags were reported and 577 plastic dart tags were reported. Additionally, there were 214 seeded fish that were reported as recovered and were double tagged with either two steel-headed tags (n = 155) or two plastic dart tags (n = 59). Of the 155 reported double steel-headed tagged fish recoveries, 138 were reported with both tags and 17 were reported with one tag. Of the 59 reported plastic dart tags, 50 were reported with both tags and 9 with one tag. The logistic regression model fitted to the double-tagged dataset (n = 848) suggested that shedding rates of steel-headed tags were lower, but the difference was not significant (p = 0.174). There did not appear to be any difference in tag shedding rates when comparing recoveries that were detected on fishing vessels against recoveries that were detected further down the supply chain.

3.3 Simulation model to inform design of tag seeding experiments

For simulations with a 30% reduction in reporting rates, eight seeding experiments for each of fleet type A and B (i.e. 16 total) were sufficient for the models to estimate decreased reporting rates in the post-change period for more than 90 % of simulated datasets.

For simulations with a 15 % increase in reporting rates, eight seeding experiments for fleet type A were sufficient for the models to estimate increased reporting rates for fleet type A for 90 % of the simulated datasets (Figure 4). However, approximately 32 seeding experiments for fleet type B were required for the models to estimate increased reporting rates for fleet type B for 90 % of the simulated datasets. More seeding experiments were required for fleet type B because a 15 % increase in reporting rates represents a weaker increase on the logit scale compared to fleet type A.

Similarly, for simulations with a 15 % decrease in reporting rates, eight seeding experiments for fleet type A were again sufficient for the models to estimate decreased reporting rates for fleet type A for 90 % of the simulated datasets (Figure 4). Approximately 32 seeding experiments for fleet type B were required for the models to estimate decreased reporting rates for fleet type B for 90 % of the simulated datasets.

4 Discussion

The estimated effects of the reporting rate model reported here are similar to those from Peatman et al. (2016). We included data from all seeding experiments where observers provided information on whether they considered the experiment to have been compromised, and included a model term to account for higher reporting rates for seeding experiments where the crew may have been aware that seeding occurred. This allows data from more seeding experiments to inform between-flag variation in reporting rates, noting that it is relatively rare for all seeded tags to be reported as recovered (< 15% of seeding experiments in the modelled dataset). This differs to the approach used by Peatman et al. (2016), who filtered the modelled dataset for seeding experiments that, based on available information, took place without the knowledge of crew. We note that we also detected weaker between-flag variation in over-dispersion compared to the 2016 analysis, and used a single over-dispersion parameter shared between all flags.

The reporting rate priors estimated here have substantially higher penalty terms than those reported in Peatman et al. (2016). This is primarily a result of a change to the approach used to generate flag-specific reporting rate distributions, and to a lesser extent reflects the increase in number of seeding experiments in the modelled datasets. Peatman et al. (2016) generated flag-specific reporting rate distributions by taking random samples from the beta-binomial distribution defined by the flag specific mean and over-dispersion parameter. The variability in the resulting flag-specific reporting rate distributions thus reflects variability in reporting rates at a seeding-experiment level. Here, we generated flag-specific reporting rate distributions by drawing sets of parameter estimates, and obtaining the flag-specific mean reporting rate for each set of parameter estimates. This ensures more appropriate variability in flag-specific reporting rate distributions and has the added benefit of taking into account the correlation between parameters. It also results in more intuitive penalties for the PTTP in relation to those for the RTTP (penalty = 244; based on Hampton, 1997), which are derived from fewer seeded tags and so should have a weaker penalty. We note that the stronger penalty terms for PTTP reporting rates appears to have helped keep MFCL PTTP reporting rate estimates for purse seine fisheries off the upper bound of 0.9 (Vincent et al., 2019b). This

The reporting rate model detected significantly lower reporting rates from tag seeding experiments in 2015, and potentially 2017. It is difficult to explore the potential causes of the low reporting rate in 2015, due to the imbalanced nature of the seeding experiment dataset and the relatively low numbers of seeding experiments undertaken that year. The low reporting rate in 2017 could be partially explained by delays in the reporting of tag recoveries from the seeding experiments. Any reduction in tag reporting in 2015 would have a limited impact on the 2019 (or future) skipjack assessment given the low number of expected returns from a limited number of skipjack releases made during the Central Pacific cruises of 2014 and 2015 (SPC-OFP, 2019).

We also report here on a variety of analyses that aim to assess the extent to which tag shedding rates differ between steel-head and plastic dart tags in a seeding experiment context (i.e. the tagging of dead fish). All analyses suggested that tag shedding rates may be lower for steel-head tags, but that the tag type effects were not significant. We note that, during the PTTP, steel-head tags have only been used in tag seeding experiments. As such, the presence of steel-head tags may actually act to alert tag finders that a tag seeding experiment has taken place, and subsequently result in upwards bias in estimates of reporting rates. As such, we recommend that plastic dart tags be used in future seeding experiments. There are tag seeding kits with plastic dart and steel-head tags that have already been distributed throughout the region. We recommend that these kits continue to be used until new tag seeding kits have been distributed and tag seeding protocols updated.

The simulation exercise presented here suggests that, at the current rates of tag seeding experiments, following a 15 % increase or decrease in tag reporting rates it would take at least 3 to 5 years of tag seeding experiments to have a reasonable chance of detecting the change. It is important to note that the simulation model is overly simplistic. In reality, any change in reporting rates may occur at a finer scale than explored in the simulation model, e.g. apply to specific fleet(s), or for tuna product that end up in specific canneries. Furthermore, changes in reporting rates may take place over a more extended period of time rather than a step-change. As such, the results of the simulation model should be interpreted as minimum guidelines if accurate estimates of reporting rates are to be obtained. In this context, we suggest 32 seeding experiments per annum as a minimum target in the short-term, targeted to fleets and regions where recoveries are most likely to occur. The current levels of tag seeding are likely only sufficient to detect substantial and widespread changes in reporting rates. Future development of the simulation study presented here would allow the generation of more robust advice on the minimum levels of seeding experiments, and how these should be distributed throughout the fishery. Tag seeding experiments generate valuable data, allowing estimation of flag-specific reporting rates which are then used to generate parameters for reporting rate priors used in MFCL assessment models. The number of seeding experiments has declined by 90% from the peak of 80 experiments per annum in 2012 and 2013. The relatively low number of seeding experiments in 2015 through to 2017, and the unbalanced nature of the dataset with respect to coverage by fleet through time, has already had an impact on our ability to monitor and interpret estimates of reporting rates from seeding experiments. Increasing the number of tag seeding experiments, and improving the coverage of fleets through time, will be required if tag seeding is to continue to be an effective component of the PTTP.

Specific recommendations for the tag seeding experiments and analysis are provided below:

- 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;
- Tag seeding experiments should be undertaken across as wide a range of vessels as possible within each fleet;
- Tag seeding experiments should use plastic dart tags, rather than a combination of plastic dart and steel-head tags;
- The number of seeding experiments should be increased from current levels. A minimum target of 32 seeding experiments per annum is suggested based on available information, noting that development of the simulation model would allow more robust advice to be provided in the future on future levels of seeding and how this effort should be distributed throughout the fishery.

Acknowledgements

This work was supported by the National Fisheries Authority of Papua New Guinea, the New Zealand Agency for International Development, the Australian Centre for International Agricultural Research, the 9th (SciFish project)and 10th (SciCoFish project) European Development Fund, Global Environment Facility (Pacific Islands Oceanic Fisheries Management project), the Republic of Korea, the République Francaise (Fond Pacifique), Republic of China, Heinz Australia, the Western and Central Pacific Fisheries Commission, Lenfest Ocean Program and the University of Hawaii (Pelagic Fisheries Research Programme). We thank Jed Macdonald for his constructive comments and suggestions on an earlier version of the report.

References

Berger A.M., S. McKechnie, F. Abascal, B. Kumasi, T. Usu & 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.

Leroy B., J. Hampton, B. Kumasi, A. Lewis, D. Itano, T. Usu, S. Nicol, V. Allain S. Caillot, 2009. PTTP summary report: review phase 2. WCPFC-SC5-2009/GN-IP-13.

Leroy B., S. Nicol, A. Lewis, J. Hampton, D. Kolody, S. Caillot, S. Hoyle, 2015. Lessons learnt from implementing three large scale tuna tagging programmes in the Western and Central Pacific Ocean. *Fish. Res.*, **163**, 23–33.

SPC-OFP, 2019. Project 42: Pacific Tuna Tagging Project Report and Work-plan for 2019-2022. WCPFC-SC15-2019/RP-PTTP-02.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rigby R.A. & 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, G.M. Pilling, 2019a. Background analyses for the 2019 stock assessment of skipjack tuna. WCPFC-SC15-2019/SA-IP-04.

Vincent, M.T., G.M. Pilling, J. Hampton, 2019b. Stock assessment of skipjack tuna in the western and central Pacific Ocean. WCPFC-SC15-2019/SA-WP-05.

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	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

Table 2 PTTP reporting rate prior distribution parameters for purse seine fisheries (all fleets), for a) the 8-region 2019 skipjack MFCL model and b) the 5-region 2019 assessment model. The penalty term is inversely related to the variance of the distribution. MULTIFAN-CL currently implements normally distributed priors for reporting rates.

a)

		PTTP	
Species	**Region**	**Mean**	**Penalty**
Skipjack	6	0.6841	638
	7	0.5749	362
	8	0.5433	793

b)

		PTTP	
Species	**Region**	**Mean**	**Penalty**
Skipjack	2	0.5749	361
	3	0.5433	793
	5	0.6841	638

Table 3 Summary of parameters used to simulate seeding experiment observations for the pre-change period.

RR type	mu	sigma	Experiments
Type-A	0.8	0.47	100
Type-B	0.5	0.47	200

Figures

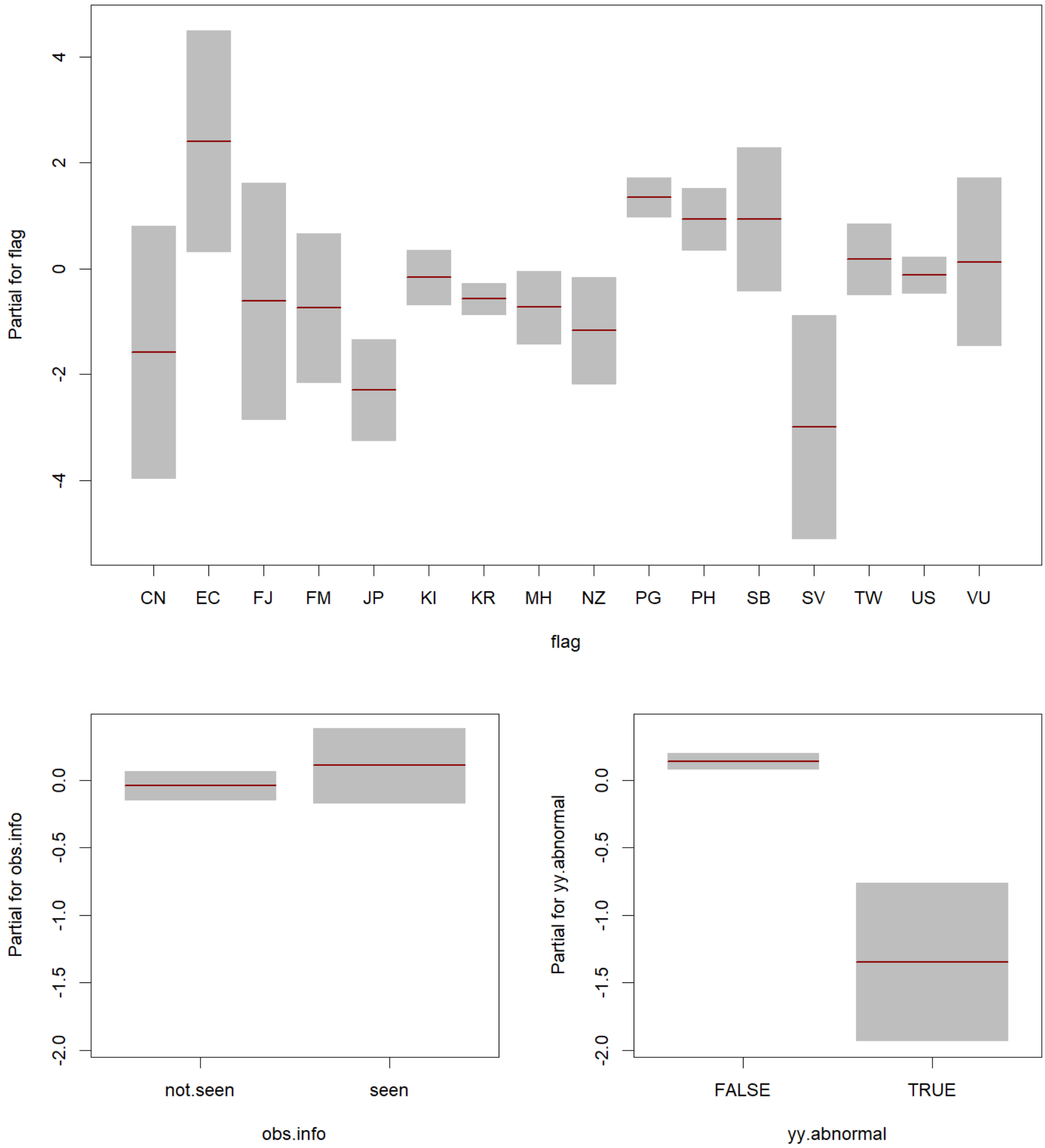


Figure 1 Effect plots for the model of reporting rates: flag (top); whether available information suggested the seeding experiment was compromised ('seen') or not ('not seen') (bottom left); and, whether the seeding experiment took place in 2015 or 2017 ('TRUE'), when reporting rates appear to have been abnormally low. The scale of the y-axis is logit transformed reporting rate.

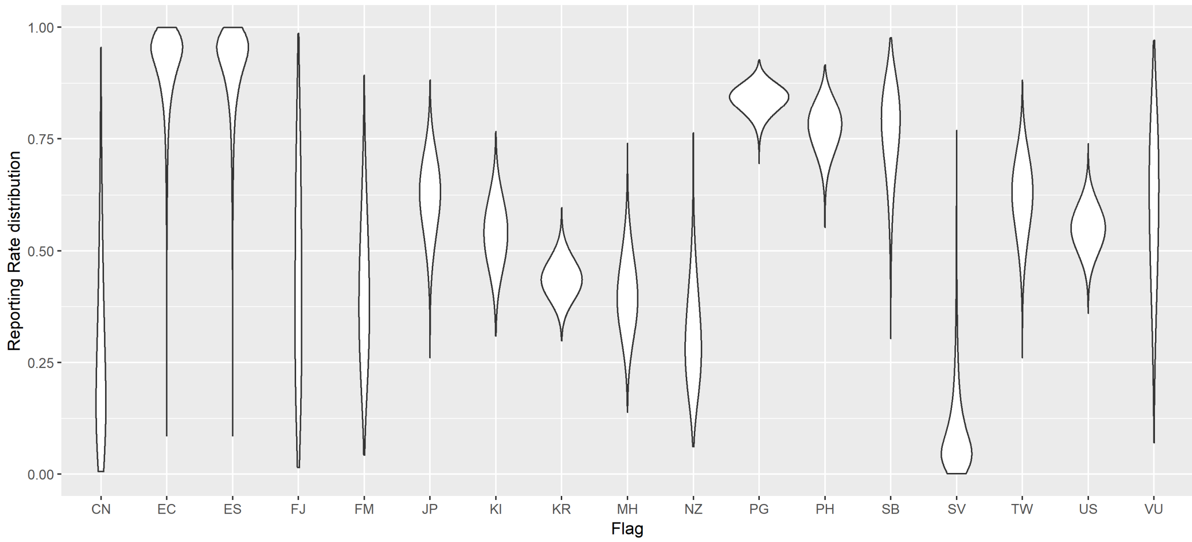


Figure 2 Flag specific reporting rate distributions used to calculate reporting rate prior parameters.

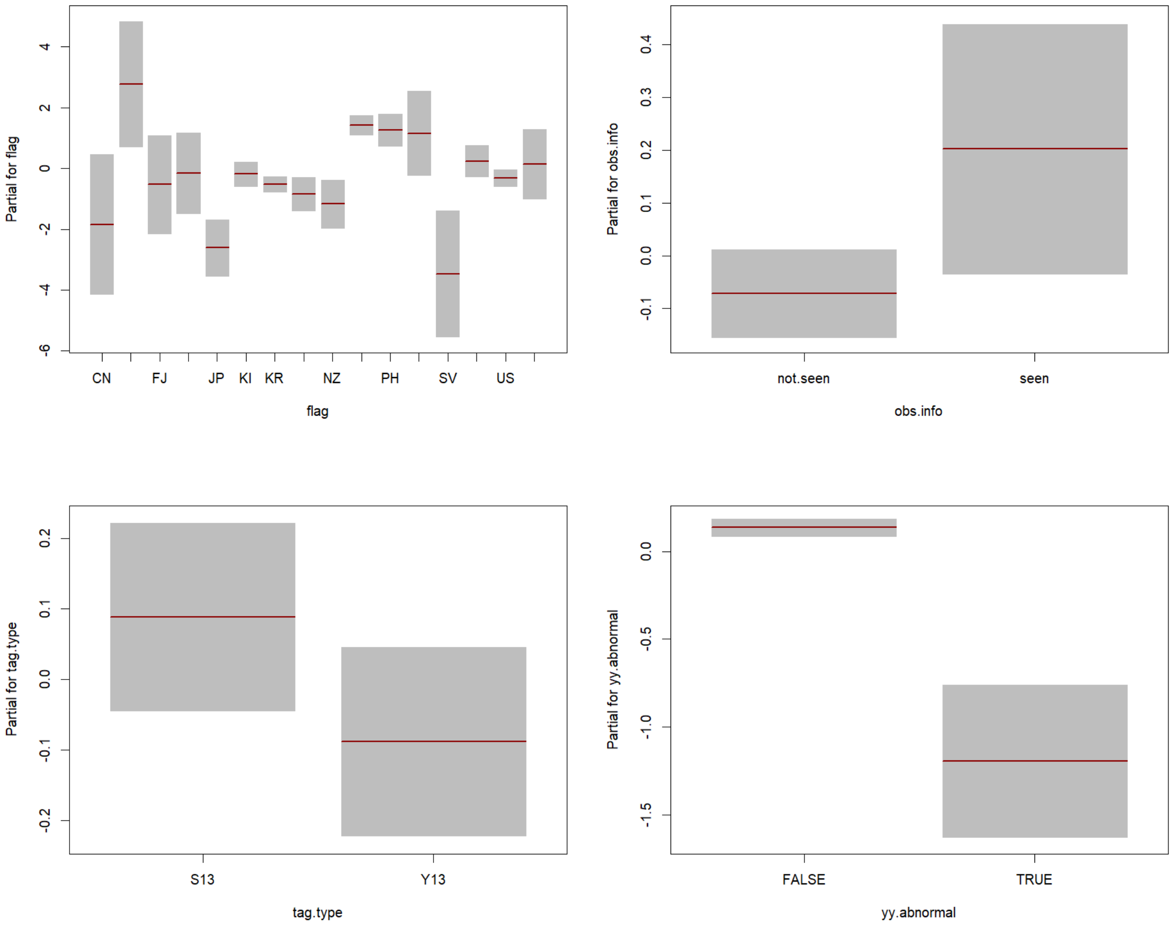
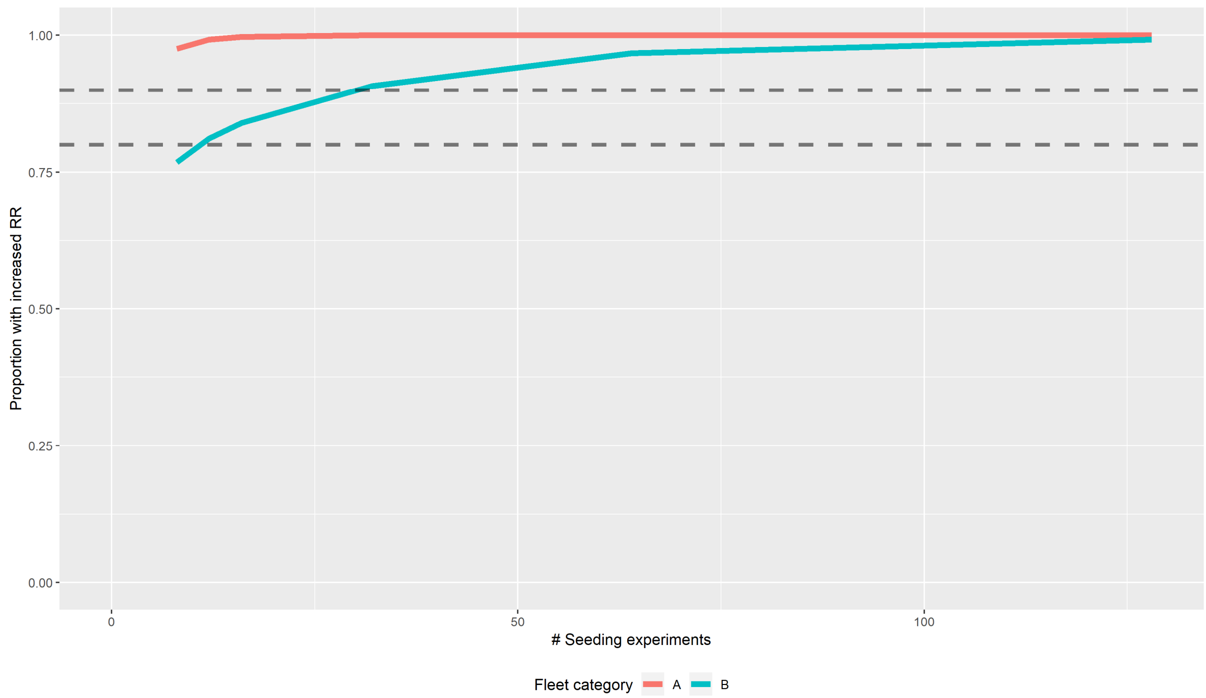


Figure 3 Effect plots for the model of reporting rates used to test for tag type effects: flag (top left); whether available information suggested the seeding experiment was compromised ('seen') or not ('not seen') (top right); tag type (S13 – steel-headed, Y13 – plastic dart) (bottom left); and, whether the seeding experiment took place in 2015 or 2017 ('TRUE'), when reporting rates appear to have been abnormally low (bottom right). The scale of the y-axis is logit transformed reporting rate.

a) 15% increase



b) 15% decrease

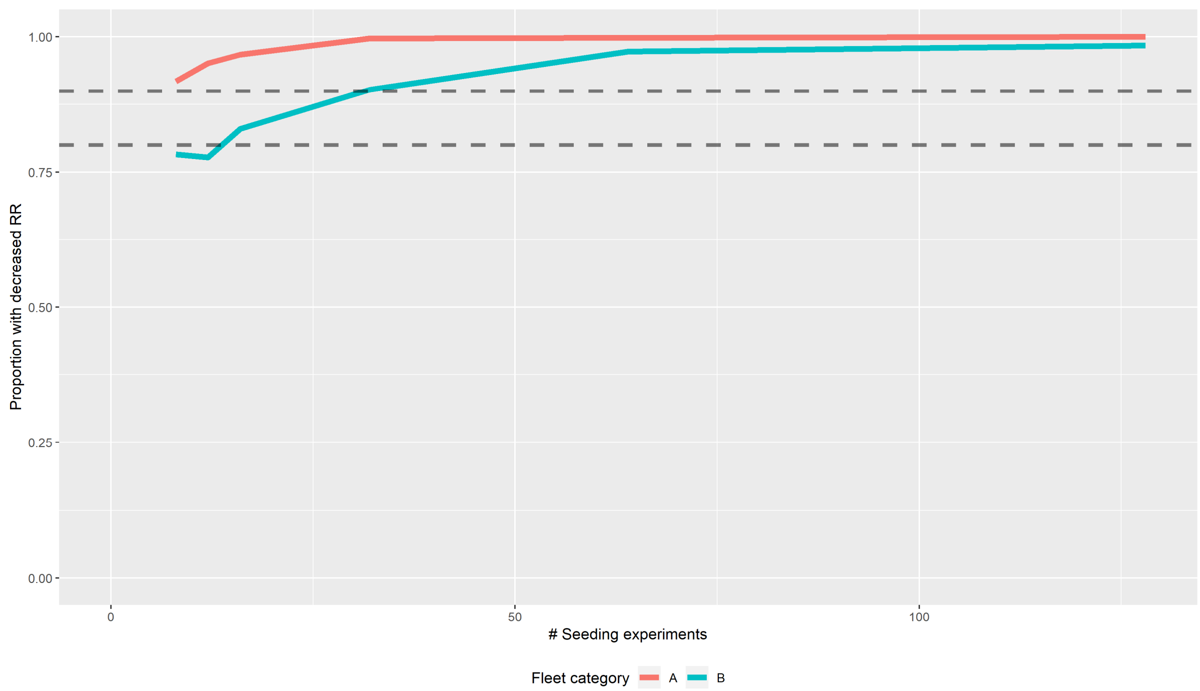


Figure 4 Proportion of the 1,000 simulated datasets for which the reporting rate model estimated increased (a) or decreased (b) reporting rates in the post change period for fleet type A (red) and B (turquoise). Panel a) has a 15 % increase in reporting rates, panel b) a 15% decrease. The y-axis is number of seeding experiments per fleet type.