SCIENTIFIC COMMITTEE
TWELFTH REGULAR SESSION

Bali, Indonesia
3-11 August 2016

Analysis of tag seeding data and reporting rates

WCPFC-SC12-2016/ SA-IP-13

Tom Peatman[1], Sylvain Caillot[1], Bruno Leroy[1], Sam McKechnie[1] Francois Roupsard[1], Caroline Sanchez[1], Simon Nicol[2], Neville Smith[1].

[1] Oceanic Fisheries Programme (OFP) Pacific Community, Noumea, New Caledonia
[2] Current address: Applied Ecology Centre, University of Canberra, Canberra, ACT, Australia

Analysis of tag seeding data and reporting rates

Executive summary

Data from tag seeding experiments have been used to estimate prior distributions for reporting rates for use in MULTIFAN CL 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 MULTIFAN-CL models.

The methodology used to estimate reporting rates requires the implicit assumption that tags seeded in tag seeding experiments were no more likely to be detected than recoveries of PTTP tag releases, i.e. that tag seeding experiments are not compromised by fishing vessel crew or potential tag finders. There are now sufficient tag seeding data to test that assumption. Here we develop statistical models to test whether tag seeding experiments are likely to have been compromised, based on the proportion of reported tag recoveries that were detected on fishing vessels. No evidence was detected for tag seeding experiments having been systematically compromised. However, reporting rates on fishing vessels were significantly higher for individual tag seeding experiments where observers thought it likely that crew had seen tag seeding take place. Additionally, due to sample sizes, the statistical power of the models may have been insufficient to detect small but significant differences given the large variability in observations.

We present a modified approach to estimating reporting rate priors that substantially improves fits to observations, resulting in reporting rate prior distributions that more accurately reflect underlying variability in flag-specific reporting rate estimates. Reporting rate prior distributions are presented for skipjack, yellowfin and bigeye tuna, assuming the regional structure from the 2014 assessments. Note that data from tag seeding experiments were removed from the modelled dataset if observers considered it likely that fishing vessel crew had seen seeding take place.

All tag recoveries reported to SPC are cross-validated using available additional datasets, e.g. VMS data and vessel logbook data, to determine the accuracy of recovery information reported to SPC. Here we develop statistical models that use tag seeding data to determine how accurately the cross-validation process estimates the reliability of tag recovery information, and the variables that influence the accuracy of these estimates. The results indicated that the cross-validation process provides estimates of accuracy that are both appropriate on a relative and absolute scale. Furthermore, the analyses suggest that the efficacy of the cross-validation process has increased with time. This is encouraging given the time and resources that are spent on cross-validation of tag recoveries.

Specific recommendations are summarised below:

- Tag seeding should be continued, targeted to fleets and regions where recoveries are most likely;
- The tag recovery cross-validation process should be continued, with due consideration of the resources required;
- A maximum of one tag seeding experiment per vessel per year should be implemented where possible, with a focus on sampling multiple vessels within fleets;
- Reporting rate prior distributions for MULTIFAN-CL assessments should be generated from flag-specific reporting rates based on beta-binomial models. This will ensure that fits to tagging data are not excessively penalised;
- The estimated reporting rate priors for regions 1 (all species), region 4 (skipjack) and region 7 (yellowfin and bigeye) should not be used. The uninformative prior distribution used for other tagging programmes (mean reporting rate = 0.5, penalty = 1) would be more appropriate for these regions.
- Data from recent tag seeding programmes in the Central Pacific should be included, once available; and,
- Future analysis of errors in recovery position and/or date should include exploration of the potential for experimental duplicates to result in correlated residuals, and therefore underestimation of uncertainty in the effects of explanatory variables.

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 including Project 35 and more recently Project 35b. In total, more than 700,000 tuna have been tagged and released of which 100,000 have been recovered and reported to SPC. Tag seeding experiments have been undertaken as a component of both the RTTP and PTTP, in which observers on purse seiners mark caught 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.

Tag seeding experiments rely on observers tagging fish such that seeded tags are treated equivalently to recoveries of tagged fish released during ordinary tagging cruises. Most importantly, this requires observers to seed fish without the knowledge of crew on the fishing vessel, or other potential tag finders. PTTP tag seeding experiments have been undertaken since 2007. The awareness amongst crew of tag seeding experiments may have increased with the number of tag seeding experiments. This would likely result higher detection rates of seeded tags on fishing vessels compared to the equivalent detection rates for tags released on PTTP tagging cruises. Comparison of fishing vessel detection rates between tag seeding experiments and PTTP tagging cruises therefore provides an opportunity to test whether tag seeding in general has taken place without the knowledge of fishing vessel crew.

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, expressing the proportion of recovered tags that are 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 (Berger *et al.*, 2014).

Tag recovery information reported to SPC undergoes rigorous data quality control, including cross-validation with available information from other datasets including VMS and vessel logbook data. This cross-validation process informs estimates of perceived uncertainty in recovery date and position for each tag recovery. The data quality control procedure is consistent across all tag recoveries. Consequently tag seeding experiments allow exploration of the accuracy of recovery information reported to SPC (e.g. Leroy *et al.*, 2015), and comparison with the perceived reliability of the tag recovery information from the cross-validation process.

As of April 2016, approximately 6 700 tags had been seeded during the PTTP as part of 322 tag seeding experiments, nearly doubling the number of seeded tags at the time of the most recent major analysis of the dataset by Berger *et al.* (2014). A reanalysis of the tag seeding dataset is timely in the context of both the additional data from tag seeding experiments and the 2016 skipjack assessment.

Three analyses are presented here:

1. Comparison of fishing vessel detection rates between tag seeding experiments and PTTP tagging cruises, to determine whether there is evidence that tag seeding experiments have been compromised.
2. Estimates of reporting rate priors based on tag seeding experiments.
3. Comparison of observed errors and perceived uncertainty in tag recovery date and location, to determine how accurately SPC's cross-validation process estimates reliability of tag recovery information and the factors that influence these errors.

2 Methods

The data required for analyses were extracted from Tagdager, the PTTP tagging database. All analyses were undertaken in R (R Core Team, 2015). Quantile residuals were used when exploring model diagnostics. Independence of residuals for models was confirmed using plots of the autocorrelation and partial autocorrelation functions. Multicollinearity between potential explanatory variables was tested using generalised variance inflation factors (VIFs) calculated using the R package car (Fox & Weisberg, 2011). The Akaike Information Criterion (AIC; Akaike, 1973) was used to evaluate support for explanatory variables. Effect plots of specific explanatory variables on the response scale were generated holding other explanatory variables at reference levels. Reference levels are provided in figure captions when used.

2.1 Reporting rates

Reporting rate models were constructed based on the approach of Berger *et al.* (2014). To summarise, flag-specific reporting rates were estimated based on tag seeding experiments. 100,000 samples were drawn from each flag-specific reporting rate distribution and these were combined to estimate reporting rate prior distributions for the skipjack, yellowfin and bigeye assessment regions based on catch-weighted averages of the flag-specific. Tag release and recovery information were extracted from Tagdager for all tag seeding experiments that commenced from 2007 to 2014 (inclusive), representing 6,237 of the 6,683 total seeded tags. Tag seeding experiments from 2015 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. This also maintains some consistency between the time period of seeding data used to estimate reporting rates, and the time period of PTTP release events used in the assessment model (McKechnie *et al.*, 2016). Assessment regions for skipjack, yellowfin and bigeye were taken from the 2014 assessment reports (Rice *et al.*, 2014; Davies *et al.*, 2014; Harley *et al.*, 2014).

The increase in seeded tags since 2014 and the collection of more detailed information with time provided opportunity for refinement of the approach used to estimate reporting rate priors. The methodology used here differed from that presented in Berger *et al.* (2014) in three aspects. Firstly, since 2009, observers have recorded whether they thought 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. Data from compromised tag seeding experiments were excluded to minimise upwards bias in reporting rate estimates, leaving 3,274 seeded tags from 154 tag seeding experiments. This filtering of data was not implemented in previous analyses, likely due to sample size limitations. Secondly, previous analyses used a quasi-binomial error distribution to account for extra-binomial variation, i.e. overdispersion. Beta-binomial models were also explored here, having the advantage of allowing parameterisation of overdispersion. A likely cause of overdispersion would be the presence of other unmodelled variables that have a significant effect on tag reporting rates but cannot be included in the models. Thirdly, previous analyses included tag type as an explanatory variable to account for potentially varying levels of tag shedding between steel and plastic anchored tags. Initial model runs suggested that the tag type effect was not significant, and more importantly introduced correlation in tag seeding experiment specific residuals. Correlation in residuals causes underestimation in standard errors of model parameters, which would result in underestimation of the variance in the reporting rate prior distributions. Consequently, tag type was not included as an explanatory variable in reporting rate models. It should be noted that tag type was also insignificant when reporting rate priors were estimated by Berger (2014).

Quasi-binomial and beta-binomial models were fitted in R packages mgcv (Wood, 2011) and gamlss (Rigby & Stasinopoulos, 2005) respectively. A binomial response variable was used to take account of correlation in residuals within tag seeding experiments. Observed reporting rates for vessels flagged to Korea, the Marshall Islands and USA were highly variable. A separate shared dispersion parameter was estimated for these flags in beta-binomial models (see model specifications below).

Base reporting rate models used a logit link and included vessel flag as the only explanatory variable in the formulation of mean reporting rates. Full specifications of the quasi-binomial and beta-binomial models, including the formulation of variance, are provided below.

Quasi Binomial

$$rec_t \sim \text{Quasi Binomial}(rel_t, \mu_t, \phi)$$

$$\text{Var}(rec_t) = \phi rel_t \mu_t (1 - \mu_t)$$

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_0 + \beta_1\, flag_t$$

where rec_t , rel_t and $flag_t$ are the total number of seeded tags reported as recovered, the total number of tags seeded and the fishing vessel flag, μ_t is the probability of detecting seeded tags (i.e. the reporting rate), ϕ is the variance inflation parameter for the quasi-binomial distribution and the subscript t references a specific tag-seeding experiment.

Beta Binomial

$$rec_t \sim \text{Beta Binomial}(rel_t, \mu_t, \sigma_t)$$

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

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_0 + \beta_1\, flag_t$$

$$\log(\sigma_t) = \begin{cases} \alpha_0 + \alpha_1 \text{ where } flag_t \in \{\text{Korea, USA, Marshall Islands}\} \\ \alpha_0 + \alpha_2 \text{ where } flag_t \notin \{\text{Korea, USA, Marshall Islands}\} \end{cases}$$

where all terms are as described for the quasi binomial model and σ_t is the dispersion term for the beta-binomial variance formulation for tag seeding experiment *t*.

Berger *et al.* (2014) commented that there are a range of other factors related to fish processing that likely influence tag reporting rates, e.g. offloading port or country. It is not straightforward to include these factors as explanatory variables in models of reporting rate as this information is generally only available for recovered tags, and tuna products from the same fishing trip do not always follow the same supply chains and so it is difficult to reconstruct explanatory variables for undetected tags. However information on well-specific destination country is in most cases collected by observers during tag seeding experiments, along with the wells in which tags were seeded. Destination country was extracted for seeded tags where one-to-one links could be made between wells with seeded tags and well-specific destination country. Alternative reporting rate models were constructed to attempt to take account of destination country of tags, a simplified proxy for product flow through the supply chain. Results for the base reporting models and alternative reporting rate models are provided separately in Sections 3.1.1 and 3.1.2 respectively. It was assumed that fishing vessel flag would be more influential on detection of seeded tags on fishing vessels (referred to hereafter as fishing vessel reporting rate), and that destination country would be more influential on detection of seeded tags in subsequent steps in the supply chain (referred to hereafter as post-fishing vessel reporting rate). Therefore fishing vessel and post-fishing vessel reporting rates were modelled separately. There was strong multicollinearity between fishing vessel flag and destination country (VIFs >> 100).

Both fishing vessel and post-fishing vessel reporting rate models were fitted in the R package gamlss (Rigby & Stasinopoulos, 2005) with beta-binomial variance and a logit link for mean reporting rate. A constant (shared) dispersion parameter was estimated for each model, with a log link.

Fishing vessel reporting rates were modelled as

$$FVrec_t \sim \text{Beta Binomial}(rel_t, \mu_t, \sigma)$$

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_0 + \beta_1 flag_t + \beta_2 year_t + \beta_3 obs\ info_t$$

where $FVrec_t$ is the number of fishing vessel recoveries of seeded tags, $year_t$ is the year, $obs\ info_t$ is a binary variable indicating whether the observer thought they had been seen seeding tags during the tag seeding experiment, μ_t is the mean fishing vessel reporting rate and the *t* subscript references a specific tag seeding experiment. The shared dispersion term for the beta-binomial variance formulation is denoted σ.

Post-fishing vessel reporting rates were modelled as

$$postFVrec_t \sim \text{Beta Binomial}(rel_t - FVrec_t, \mu_t, \sigma)$$

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \begin{cases} \beta_0 + \beta_1 destination_t + \beta_2 year_t, & \text{or} \\ \beta_0 + \beta_1 flag_t + \beta_2 year_t \end{cases}$$

where $postFVrec_t$ and $destination_t$ are the number of post-fishing vessel recoveries and the destination country for seeded tags from tag seeding experiment *t* respectively. Note that, as such, the numbers of seeded tags available for recovery post-fishing vessel were adjusted to take account of fishing vessel recoveries.

2.2 Exploratory analysis of tag recovery validation process

The accuracy and precision of the perceived reliability in reported information was explored by modelling errors in the recovery date and position of seeded tags as a function of reliability indices assigned during tag recovery cross-validation. Models of recovery position assumed normally distributed errors. Models of recovery date assumed Tweedie distributed errors (Tweedie, 1984) to account for the high proportion of tags where the recovery date was reported with error (47 % of records in the modelled dataset).

Predictive models of errors in recovery information were constructed to explore how different factors influence the reliability of recovery information. Explanatory variables included in the models were: the stage in the supply chain where the tag was detected; the year of the tag seeding experiment; whether the recovery vessel had been correctly identified; the time delay between tag seeding and tag detection; the country where the tag recovery was reported (not for recovery location due to multi-collinearity with the stage in the supply chain that the tag was detected); the length of the fishing vessel trip where the tag seeding experiment took place, measured in days and nautical miles for recovery date and position respectively; and, an interaction between supply chain and the time delay between seeding and detection. There was strong multi-collinearity between reporting country, fishing vessel flag and destination country. Reporting country was included as it was considered *a priori* to be most influential on recovery information accuracy. It should be noted that models fitted with either fishing vessel flag or destination country instead of reporting country had less support from observations based on AIC.

Tag recovery data were extracted from Tagdager for all seeded tags that had undergone the complete cross-validation process. Tag recoveries detected on fishing vessels were removed if the observer thought that crew had seen seeding taking place, on the assumption that recovery information would therefore be more reliable. This left 1,406 recoveries for models of recapture date and 1,047 recoveries for recapture location, out of the 3,699 total recoveries. Initial model runs identified strong correlation in residuals within tag seeding experiments, so random intercepts for tag seeding experiment were included in an attempt to address this. The inclusion of the random intercept term did not completely remove correlation in residuals within tag seeding experiments. The presence of correlation in residuals reduces standard errors in model parameters, which increases the chance of type I errors. Alternative approaches to removing the correlation in residuals were unsuccessful, e.g. compound symmetry in residuals at a tag seeding experiment or tag seeding event level.

Models were fitted using the mgcv library (Wood, 2011), with Tweedie variance function power parameters estimated as part of the model fitting procedure (using the function mgcv::tw).

2.3 Exploration of whether seeding trials have been compromised

Tag release and recovery data were extracted from Tagdager for all PTTP tagging cruises and tag seeding experiments. The proportion of total reported tag recoveries detected on fishing vessels was used as the basis of comparison, referred to as the fishing vessel detection rate. Fishing vessel detection rates were compared between recoveries of PTTP releases and tag seeding experiments where 'obs info' was available. As such, data from compromised seeding cruises are included in the modelled dataset, to avoid unnecessary reductions in statistical power, but any upwards bias can be accounted for by the 'obs info' effect. Fishing vessel detection rates were disaggregated by recapture vessel flag and release cruise, with release cruise defined as either a tag release cruise or tag seeding experiment. Flag-specific fishing vessel detection rates were modelled as a function of year, cruise type (i.e. PTTP v seeding) and 'obs info'

$$\operatorname{logit}\left(\frac{FVrec_c}{rec_c}\right) = \beta_0 + \beta_1 cruise\ type_c + \beta_2 year_c + \beta_3 obs\ info_c$$

where *c* denotes an individual PTTP tagging cruise or tag seeding experiment. 'Obs info' for PTTP recoveries was set to "not.seen" as fishing vessel crew would have been unaware that tagged fish had been recovered. Beta-binomial variance was assumed to account for extra-binomial variation, with models implemented in the R package gamlss (Rigby & Stasinopoulos, 2005). There were sufficient tag seeding experiments to construct flag-specific models for the US, Korea, the Philippines, Papua New Guinea, Chinese Taipei and Kiribati. The analysis, whilst using quantitative tools, should be considered more as a qualitative exploration of the data due to the relatively low sample sizes and the unbalanced nature of the datasets. Year and 'obs info' were included in all models regardless of statistical significance as both have a significant effect on fishing vessel reporting rates (see Section 3.1.2), and therefore detection rates.

3 Results

3.1 Reporting rates

3.1.1 Base model

Flag-specific reporting rate estimates for the base model are presented in Figure 1. The beta-binomial model better captured the variability in observed reporting rates compared to the quasi-binomial model. In particular, the variability in observed reporting rates for vessels flagged to Korea (KR), Marshall Islands (MH), Philippines (PH), PNG (PG) and USA (US) were underestimated by the quasi-binomial model. The beta-binomial dispersion parameter was estimated at 0.70 for Korea, Marshall Islands and USA, and 0.36 for all other flags. There were comparatively few seeding trips in the modelled dataset for vessels flagged to China (CN), Ecuador (EC), Fiji (FJ), New Zealand (NZ) and the Solomon Islands (SB), resulting in low precision in reporting-rates for these flags.

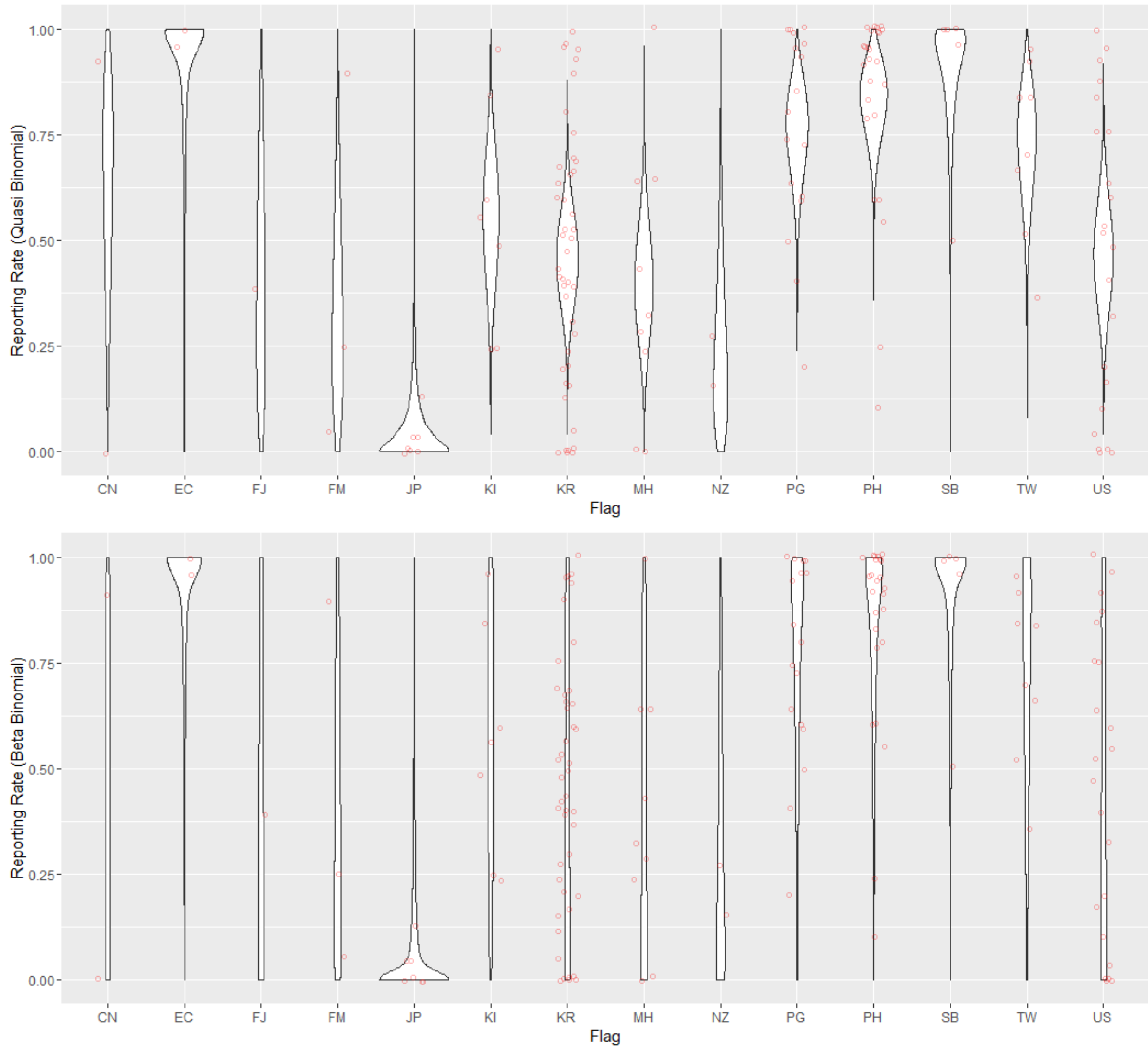

Figure 1 Estimated reporting rates by flag assuming quasi-binomial (top) and beta-binomial (bottom). The width of the 'violin' gives the probability density of a given reporting rate. Observed reporting rates from tag seeding experiments are provided (red circles, jittered).

The reporting rate estimates for the Japanese fleet were considered unlikely given the reported recoveries from the fleet and so the Taiwanese mean reporting report was applied to the Japanese fleet, as assumed by Berger *et al.* (2014). In the absence of empirical data, mean reporting rates of Spain and Vanuatu were set to those of Ecuador and PNG respectively and the mean reporting rate of the Philippines fleet was applied to Indonesia and Vietnam. Standard errors of all assumed means were set to the maximum across all flags, with beta-binomial dispersion parameters set to 0.36. A mean reporting rate of 0.5 was applied to El Salvador, i.e. approximately uniformly distributed across the unit interval.

Table 1 PTTP reporting rate prior distribution parameters for purse seine fisheries (all fleets) by species and region. The penalty term is inversely related to the variance of the distribution. MULTIFAN-CL currently implements normally distributed priors for reporting rates. Regions were taken from the 2014 assessment models.

		PTTP – Quasi-Binomial		PTTP – Beta-Binomial	
Species	**Region**	**Mean**	**Penalty**	**Mean**	**Penalty**
Skipjack	1	0.6179	4	0.6642	5
	2	0.5980	73	0.5892	41
	3	0.5440	186	0.5360	28
	4	0.7466	21	0.8081	23
	5	0.6911	113	0.6867	47
Bigeye	1	0.6179	4	0.6642	5
	3	0.5838	147	0.5647	44
	4	0.6315	96	0.6333	51
	7	0.7396	17	0.8079	22
	8	0.7149	182	0.7048	45
Yellowfin	1	0.6179	4	0.6642	5
	3	0.5954	105	0.5819	43
	4	0.5542	191	0.5479	31
	7	0.7473	21	0.8085	23
	8	0.7214	170	0.7168	45

The mean of the reporting rate priors was insensitive to the assumed variance formulation (Table 1). Reporting rate prior penalties were higher for the quasi-binomial estimates compared to the beta-binomial estimates, reflecting the lower variance in estimated reporting rates for the quasi-binomial model. Note penalties were calculated as $penalty = (2 * variance)^{-1}$, with the variance calculated on the nominal scale.

3.1.2 Alternative model accounting for supply chains

Significant effects on fishing vessel reporting rates were detected for vessel flag, year and whether or not the observer thought the crew had seen tag seeding taking place (Figure 2). As expected, greater proportions of seeded tags were detected on fishing vessels when observers thought that the crew had seen tag seeding taking place. Fishing vessel reporting rates were highest for vessels flagged to PNG and the Philippines, though mean reporting rate estimates for some flags had substantial uncertainty. An increasing trend in fishing vessel reporting rates with time was detected.

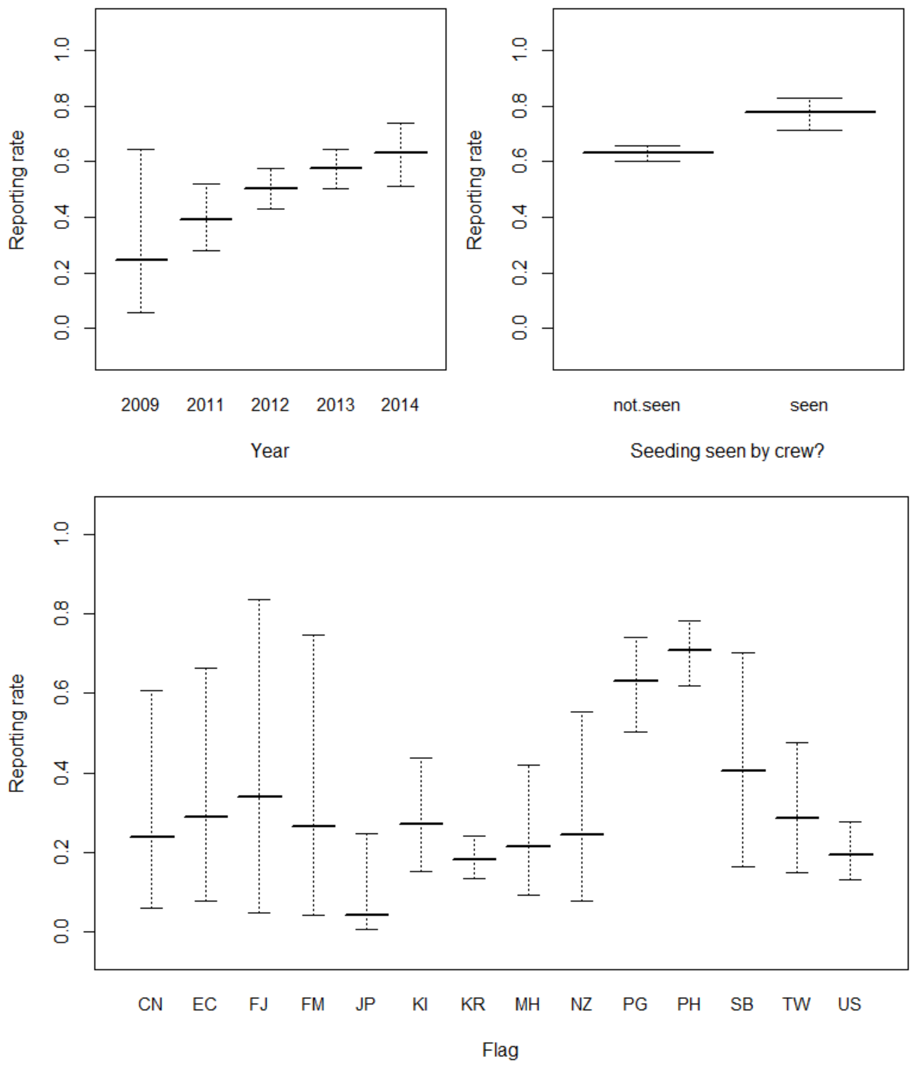

Figure 2 Mean fishing vessel reporting rate against year (top left), whether the observer thought the crew had seen tag seeding (top right) and flag (bottom). Reference levels for explanatory variables were: year = 2014; seeding seen by crew = not.seen; flag = PG.

Destination country was more influential on post-fishing vessel reporting rates than vessel flag (AIC of 802.6 against 838.6). It was not possible to include both destination country and vessel flag as explanatory variables due to extreme multicollinearity. No significant relationship was detected between post-fishing vessel reporting rates and whether observers thought fishing vessel crew had seen tag seeding taking place. The final post-fishing vessel reporting rate model included vessel flag and year as explanatory variables, both of which had significant effects on the response variable (Figure 3). Post-fishing vessel recovery rates decreased with time, though with some variability between 2007 and 2010. There was substantial variation in post-fishing vessel recovery rates between flags.

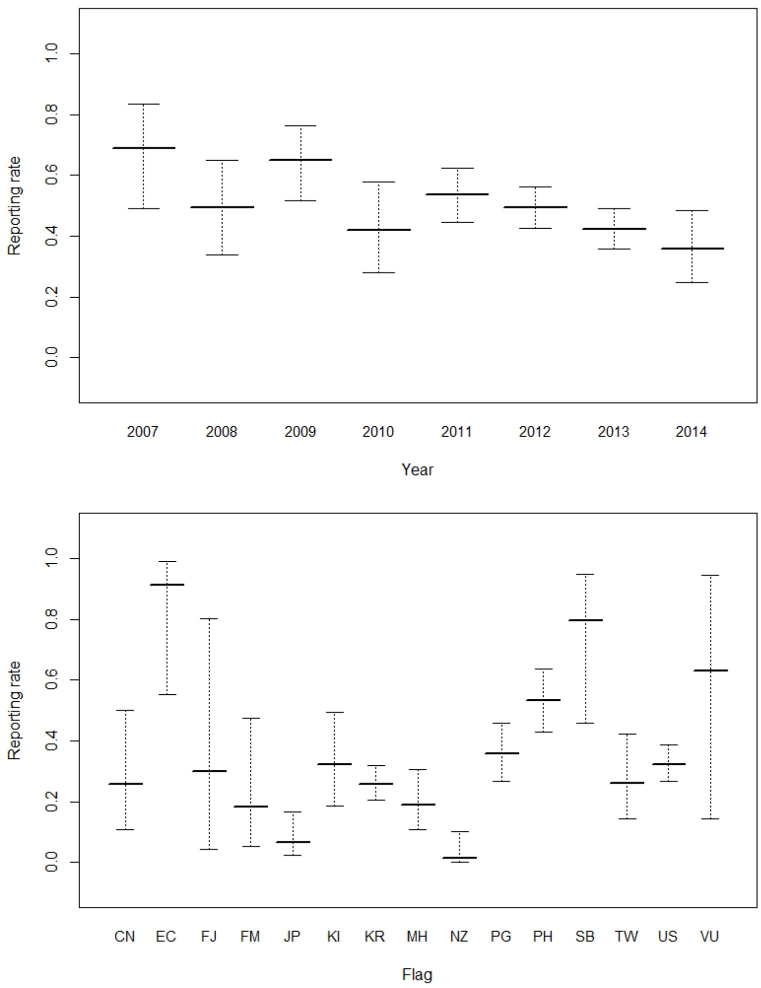

Figure 3 Mean post-fishing vessel recovery rates against year (top) and vessel flag (bottom). Reference levels for explanatory variables were: year = 2014, flag = PNG.

3.2 Exploratory analysis of tag recovery validation process

Errors in recapture date and position for seeded tags increased in line with their perceived uncertainty (Figure 4). Uncertainty in recovery information was overestimated for seeded tags with perceived reliability of +/- 2 degrees or worse.

The effect of explanatory variables on recovery date and position are included in Appendix A, Figure 5 and Figure 6 respectively. Errors in recovery date were strongly influenced by the country where the tag recovery was reported. As would be expected, errors in recovery date and position were significantly higher in cases where the recovery fishing vessel had not been correctly identified. Errors in recovery date and position were consistently low for tags detected in canneries. Errors in recovery date for tags detected on fishing vessels were relatively high, i.e. not significantly different to tags detected in cold storage, in fish markets or on carrier vessels. In contrast, errors in recovery position for tags detected on fishing vessels were relatively low, i.e. not significantly different to tags detected in canneries. Errors in recovery position also displayed decreasing trends with time. Errors in recovery date increased with the delay between tag seeding (i.e. ‘recovery’) and tag detection regardless of what stage in the supply chain tags were detected. Errors in recovery position increased with the time delay between tag seeding and tag detection for tags detected on fishing vessels, carrier vessels and in canneries. Errors in recovery position were not influenced by the time delay in detection for tags detected in cold storage and at fish markets. Tags detected on fishing vessels had comparatively low errors in recovery date and location, having controlled for the average time delay between seeding and detection (Table 2). Conversely, tags detected on carrier vessels had comparatively high errors in recovery date and location. Tags detected in canneries had high errors in recovery date but low errors in recovery location, with the opposite true of tags recovered in cold storage. The length of the observer trip had no significant effect on errors in recovery position and date and was removed from models, as this information was not available for all observer trips and thus would have reduced the sample size and thus statistical power of the analyses.

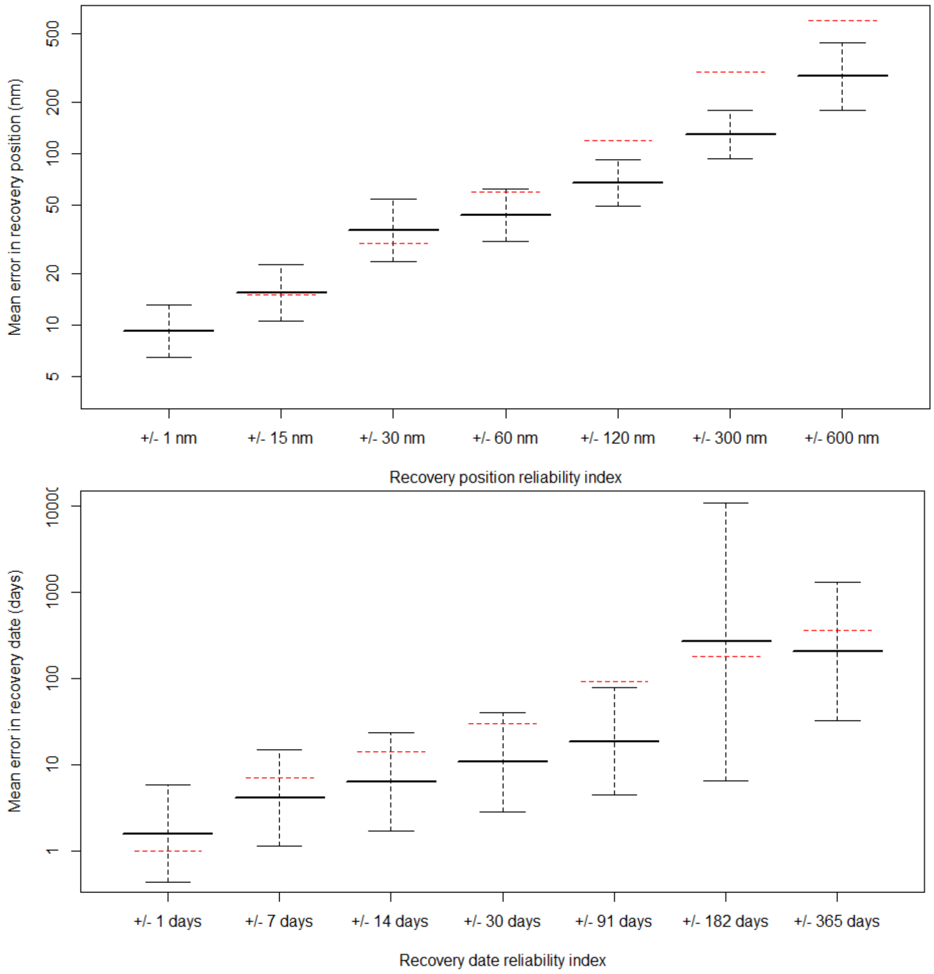

Figure 4 Mean errors in recovery position (top) and date (bottom) against reliability indices assigned during tag recovery cross-validation. Mean errors (thick black line) and the 95 % confidence intervals (thin black lines) are displayed along with the maximum error implied by the reliability index (dotted red line).

3.3 Exploration of whether seeding trials have been compromised

No significant differences were detected between fishing vessel detection rates of seeded tags and PTTP tagging cruises (Figure 7). The refitting of models without ‘obs info’ as an explanatory variable indicated that the lack of significant differences was not due to insufficient statistical power resulting from filtering of the dataset to exclude records missing ‘obs info’.

4 Discussion

Comparison of fishing vessel detection rates provided no evidence of significant differences between tag seeding experiments and PTTP cruises. However it is important to note that the low sample sizes will have impacted the statistical power of comparisons in the context of the high variance of fishing vessel detection rates. Furthermore, there were no clear temporal trends in the proportions of tag seeding experiments that were considered to be compromised by observers, either for individual flags or for the seeding programme as a whole. Therefore, whilst fishing vessel crew have been aware of specific tag seeding experiments, there was no evidence that the tag seeding programme as a whole was compromised. Tag seeding experiments thought to have been compromised were excluded, or treated appropriately, in analyses presented here so as to minimise bias in results.

For some fleets there was a tendency for multiple seeding experiments to be undertaken on the same vessel (mean 2 seeding experiments per vessel, maximum 9), sometimes in the same year. No relationship was observed between repeated tag seeding experiments on a given vessel and fishing vessel detection rates, or whether the cruise was considered to be compromised by observers. However, repeated sampling of vessels in a short period of time may increase the chance of fishing vessel crew becoming aware of the tag seeding process and so should be avoided where possible given the constraints of the observer programme. Furthermore, sampling of multiple vessels within fleets should ensure that sampled vessels are representative of the fleet as a whole and mitigate bias in flag specific reporting rates due to un-modelled vessel effects.

The proportion of total tags recovered (or seeded) that were detected on fishing vessels would provide a more appropriate metric to explore whether tag seeding experiments have been compromised, as it also takes account of differences in fishing vessel reporting rates between tag seeding experiments and PTTP tagging cruises. However trip-specific reporting rates appear to be highly variable, so the total number of recoveries of PTTP tags on a trip-by-trip basis cannot be estimated with a high degree of certainty.

Separate modelling of fishing vessel and post-fishing vessel reporting rates identified an increasing trend in fishing vessel reporting rates with time and a decreasing trend in post-fishing vessel reporting rates with time. The temporal trend in fishing vessel reporting rates could be explained by increased coverage of key ports by tag recovery officers as the PTTP has progressed, along with a general increase in awareness amongst fishing vessel crew of the PTTP. There are no ready explanations for the decreasing temporal trend in post-fishing vessel reporting rates. Regardless these temporal trends demonstrate the value of continued tag seeding, providing means of monitoring changes in tag recovery processes. Continued tag seeding experiments would be most informative if targeted at fleets and regions most likely to regularly recover tags from PTTP tagging cruises, and so will depend on where and when PTTP tagging cruises occur. Base models of reporting rate, combining fishing vessel and post-fishing vessel tag reporting, are preferred to separate modelling of the two processes to maintain consistency with the treatment of tagging data in MULTIFAN-CL. It is important to note that additional base model runs, including the year of tag seeding experiments in the parameterisation of mean reporting rate, did not detect a significant year effect with either quasi-binomial or beta-binomial errors which supports the use of time invariant reporting rates in MULTIFAN-CL.

Beta-binomial models of reporting rates performed better than the quasi-binomial models used in previous analyses, substantially improving fits to observations. Reporting rate prior distributions for MULTIFAN-CL assessments should be generated from the beta-binomial derived flag effects, to prevent excessive penalisation of deviation from the means of reporting rate prior distributions. This would increase flexibility within the assessment models to improve fits to other observations (e.g. length frequencies, CPUE indices) without compromising fits to the tagging data.

Reporting rate prior distributions for all species in region 1 were based exclusively on the estimated reporting rates for the Japanese fleet, as the sole purse seine fleet operating in the region, with these estimates derived from assumptions on the mean reporting rate and its variance. Reporting rate priors for the equatorial region west of 140 E, namely region 4 for skipjack and region 7 for yellowfin and bigeye, were predominantly informed by reporting rate estimates for the Philippines and Indonesia. Reporting rates of the Indonesian fleet were assumed to be similar to the Philippines fleet. However tag seeding experiments on Philippines vessels were exclusively conducted on trips east of 140 E. Thus the estimated reporting rates for sampled Philippines vessels may not reflect those of Philippines and Indonesian vessels operating west of 140 E, due to differences in vessel characteristics and supply chains. As such, the reporting rates priors for regions 1 (all species), region 4 (skipjack) and region 7 (yellowfin and bigeye) may not be appropriate for fleets operating in these regions. The uninformative prior distribution used for other tagging programmes (mean reporting rate = 0.5, penalty = 1) would likely be more appropriate for these regions.

The inclusion of destination country as an explanatory variable did not improve models for post-fishing vessel reporting rate. This may reflect inaccurate information on destination country, for example due to sorting and mixing of tuna catch between wells post-seeding. However there is strong correlation between destination country and vessel flag, and so vessel flag does incorporate some information on destination country. It should also be noted that carrier vessel VMS data are now provided to SPC, which in the future should provide the means to explore the accuracy of destination country information recorded by observers. Inclusion of other potential explanatory variables related to fish processing and transport would likely necessitate a Bayesian modelling framework coupled with detailed information on supply chains to overcome the absence of information for unreported tags.

Base models of reporting rate were only fitted to data from tag seeding experiments in which observers considered seeding to have been undetected by fishing vessel crew. This reduced the number of tags in the modelled dataset by approximately 50 %. Including or excluding compromised seeding experiments amounts to a compromise between mitigating upwards bias in mean reporting rate due to increased fishing vessel reporting rates at the expense of decreased reporting rate prior penalties. Additional model runs demonstrated that reporting rate prior parameters were insensitive to whether the dataset was filtered.

In the wider context of the PTTP, seeding data suggested that the cross validation process provides accurate estimates of tag recovery information in both relative and absolute terms. The perceived reliability of recovery position with reliability indices +/- 2 degrees (or worse) appear to be pessimistic, though these account for a relatively low proportion (30 %) of total PTTP recoveries. However, the perceived reliability of recovery date was broadly appropriate for all reliability indices. The tag recovery cross-validation process is resource-intensive and thus the confirmation of its efficacy is particularly encouraging. It is worth noting that the estimates of errors in recovery information provide a means to incorporate uncertainty in recovery information in analyses of tagging data, e.g. through likelihood weighting.

Previous studies have commented on apparent relationships between individual factors and uncertainty in PTTP recovery information (*e.g.* Leroy *et al.*, 2015). The models of errors in recovery date and position presented represent the first attempts to quantify the effects of factors on uncertainty in recovery information using statistical models. It is possible that some of the correlation in residuals is a result of experimental duplicates, i.e. tags that were seeded during the same release event, and detected by the same tag finder. This should be further explored in future analyses. Incorrectly identifying the recovery fishing vessel increased errors in recovery information. This is well known, and results from the use of VMS and logbook information from the reported recovery vessel in the cross-validation process. There was a decreasing trend in errors in recovery date and position with time, suggesting that there has been an improvement in the tag validation process with time. Errors in recovery information generally increased with the delay between tag recovery (i.e. seeding) and tag detection. For tags detected on fishing vessels, this could reflect a general increase in uncertainty of the origin of fish in wells due to well sorting and mixing, or simply reflect errors in well numbers reported to tag recovery officers. Tags detected in canneries had comparatively low estimates of mean errors in recovery position, accounting for the effect of other variables, with lower estimates than for recovers detected on carrier vessels. However, this is partly compensated by the higher rates of recovery fishing vessel misidentification for recoveries detected in canneries (21 %, compared to 7 % for tags recovered at other stages in the supply chain). The length of the observer trip had no significant effect on errors in recovery information. It would be interesting to see whether this also applies to tag seeding trips in the Central Pacific where fishing trips can be longer, once data from these trips are available.

Specific recommendations are summarised 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;
- The tag recovery cross-validation process should be continued, with due consideration of the resources required;
- A maximum of one tag seeding experiment per vessel per year should be implemented where possible, with a focus on sampling multiple vessels within fleets;
- Reporting rate prior distributions for MULTIFAN-CL assessments should be generated from flag-specific reporting rates based on beta-binomial models. This will ensure that fits to tagging data are not excessively penalised;
- The estimated reporting rate priors for regions 1 (all species), region 4 (skipjack) and region 7 (yellowfin and bigeye) should not be used. The uninformative prior distribution used for other tagging programmes (mean reporting rate = 0.5, penalty = 1) would be more appropriate for these regions.
- Data from recent tag seeding programmes in the Central Pacific should be included, once available; and,
- Future analysis of errors in recovery position and/or date should include exploration of the potential for experimental duplicates to result in correlated residuals, and therefore underestimation of uncertainty in the effects of explanatory variables.

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).

References

Akaike H. (1973). Information theory and an extension of the maximum likelihood principle. In: B.N. Petrov,F. Csaki (Eds.), Second International Symposium on Information Theory, Akademiai Kiado, Budapest, 1973, 267–281.

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.

Davies N., S. Harley, J. Hampton & S. McKechnie (2014). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA- WP-04 Rev1 25 July.

Fox J. & S. Weisberg S. (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Harley S., N. Davies, J. Hampton & S. McKechnie (2014). Stock assessment of bigeye tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA-WP-01 Rev1 25 July.

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.

McKechnie S., D. Ochi, H. Kiyofuji, T. Peatman, and S. Caillot (2016). Construction of tagging data input files for the 2016 skipjack tuna stock assessment in the western and central Pacific Ocean. WCPFC-SC12-2016/SA-IP-05, Bali, Indonesia, 3-11 August 2016.

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/.

Rice J., S. Harley, N. Davies & J. Hampton (2014). Stock assessment of skipjack tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA-WP-05 Rev1 25 July.

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

Tweedie, M.C.K. (1984). An index which distinguishes between some important exponential families. Statistics: Applications and New Directions. In: J. K. Ghosh and J. Roy (Eds.), Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, 579-604. Calcutta: Indian Statistical Institute.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)*, **73**, 3-36.

Appendix A: Additional Information

Table 2 Predicted mean error in recovery date (days) and location (nm) for tags detected at different stages in the supply chain. Time difference between seeding and detection was set at the respective average in the modelled dataset for errors in date of recovery. Vessel match, year and country of reporting were held constant at reference levels.

Where found	Time difference	Vessel match	Year	Country of reporting	Mean error	
					Date	Location
Fishing vessel	67.5	TRUE	2010	PG	5.9	40.6
Carrier vessel	17.5	TRUE	2010	PG	7.3	186.2
Fish market	32.4	TRUE	2010	PG	6.5	65.8
Cold storage	19.5	TRUE	2010	PG	4.1	124.7
Cannery	31.3	TRUE	2010	PG	20.4	58.6

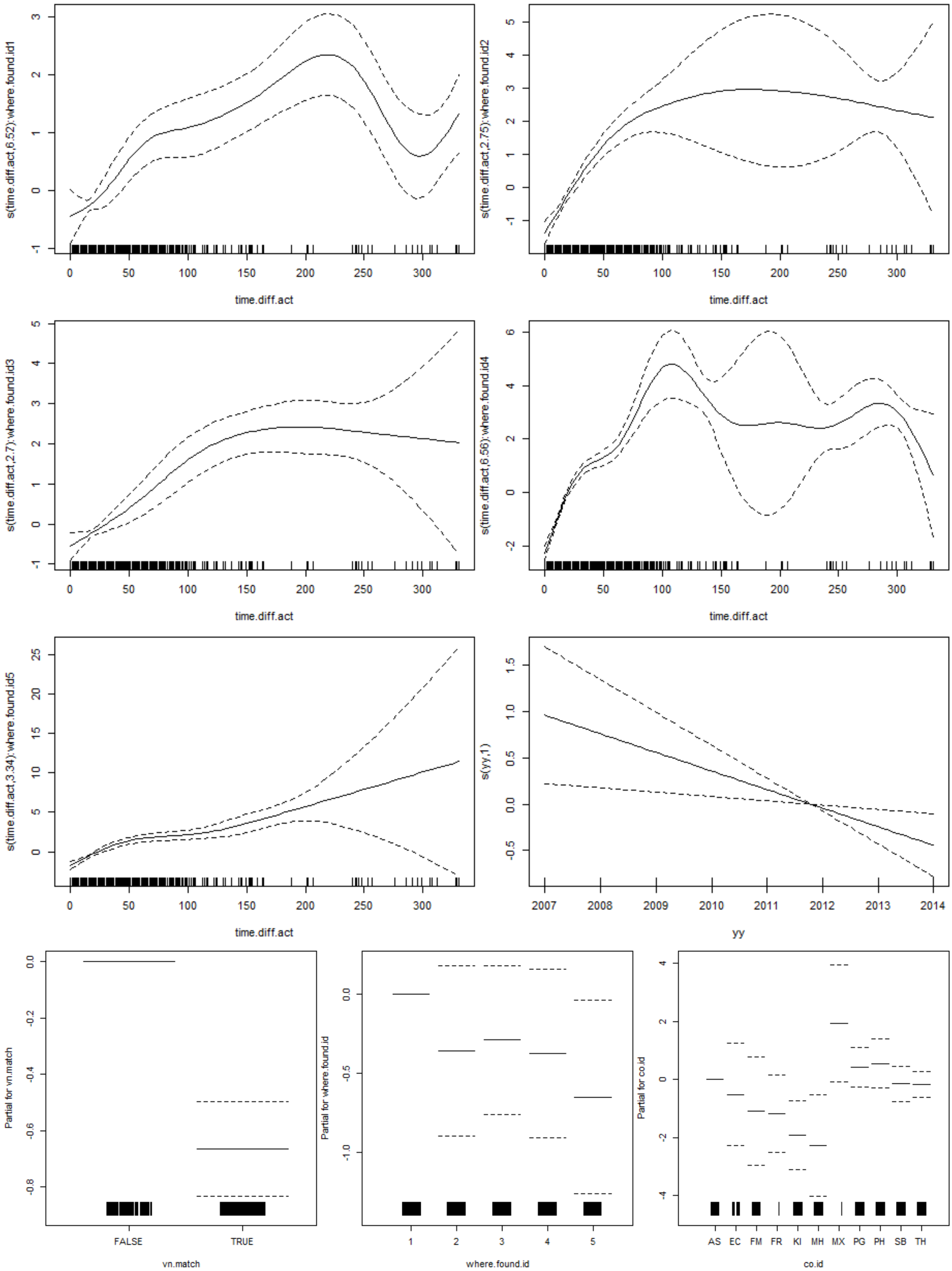

Figure 5 Effect plots for recovery date (log transformed) against: time difference between recovery and detection and where.found.id (top row, second row, third row left); year (third row right); correct identification of recovery fishing vessel (bottom left); where found id (bottom centre) and country where the tag recovery was reported (bottom right). Where found ids: 1 = cannery, 2 = cold storage, 3 = fish market, 4 = fishing vessel, 5 = carrier vessel, 6 = unknown.

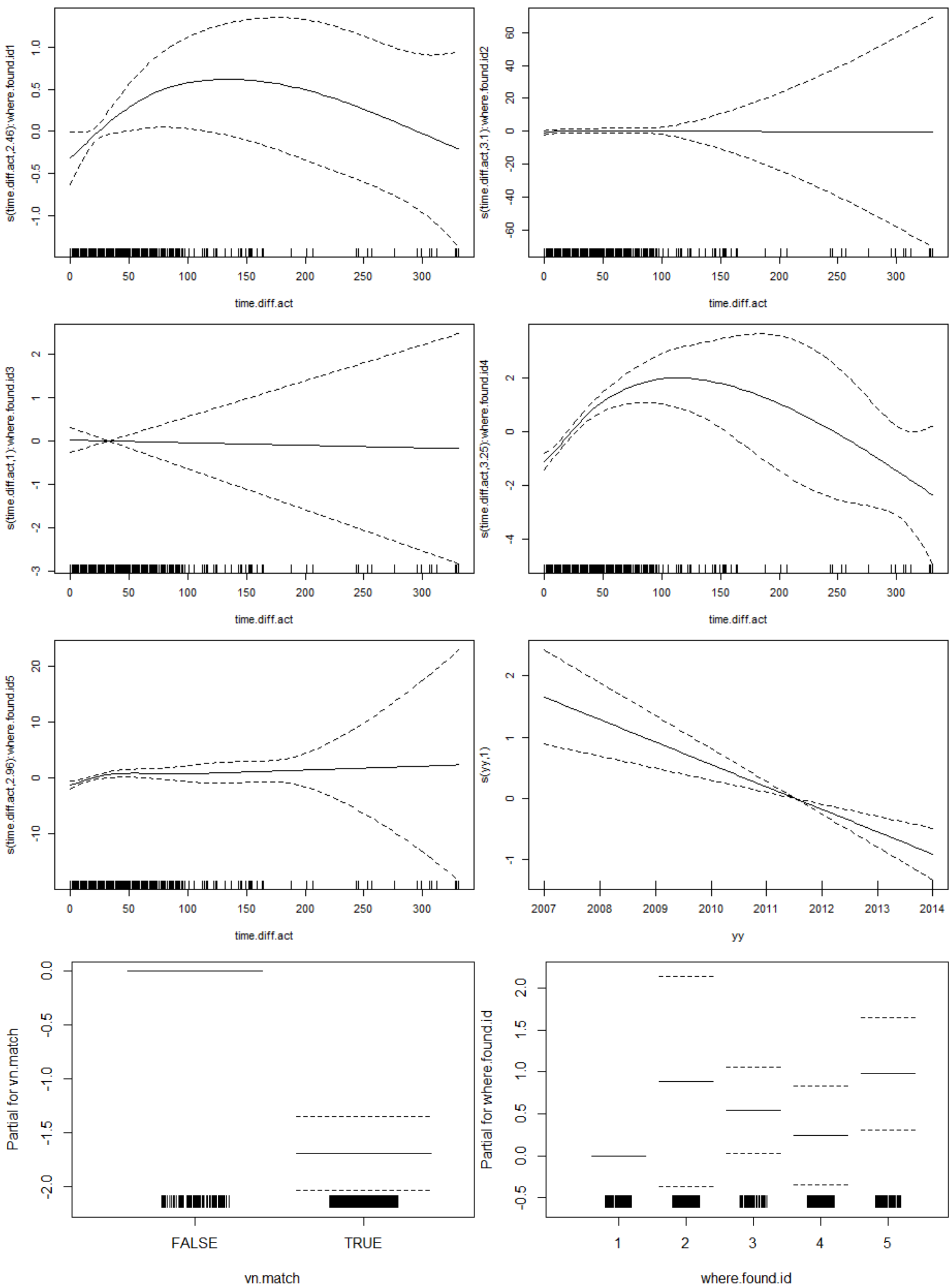

Figure 6 Effect plots for recovery position (log transformed) against: time difference between recovery and detection (top left); correct identification of recovery fishing vessel (top right); year (middle left); where found (middle right); and, country where the tag recovery was reported (bottom). Where found ids: 1 = cannery, 2 = cold storage, 3 = fish market, 4 = fishing vessel, 5 = carrier vessel, 6 = unknown.

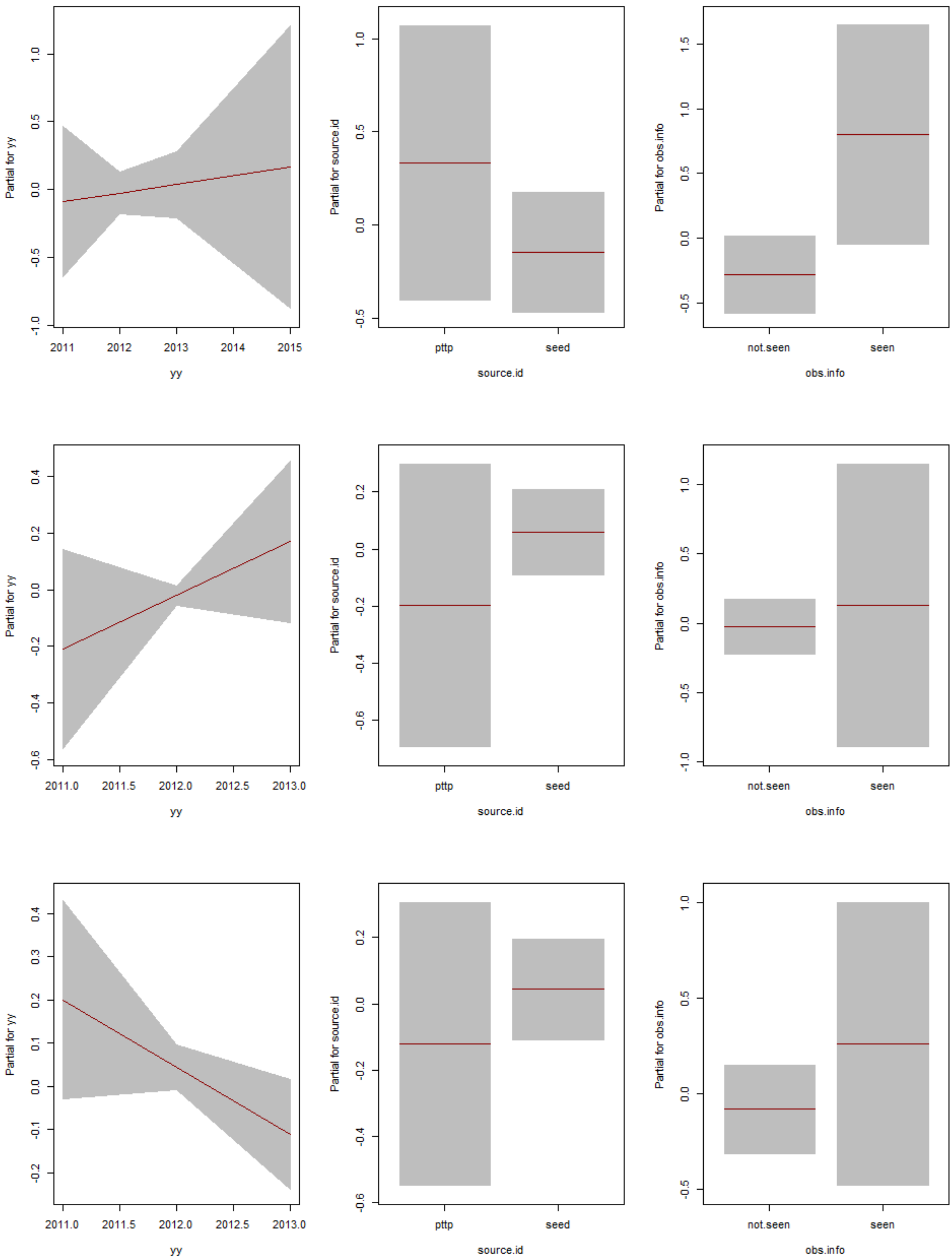

Figure 7 Effect plots for mean fishing vessel detection rates against: year (left column); cruise type – PTTP or seeding (middle); and obs info – whether the observer thought crew had seen seeding take place (right). Top row – US fleet, middle row – Korea fleet; bottom row – Philippines fleet. Figure is continued on next page for PNG, Chinese Taipei and Kiribati. Note that the y-axe ranges are not consistent.

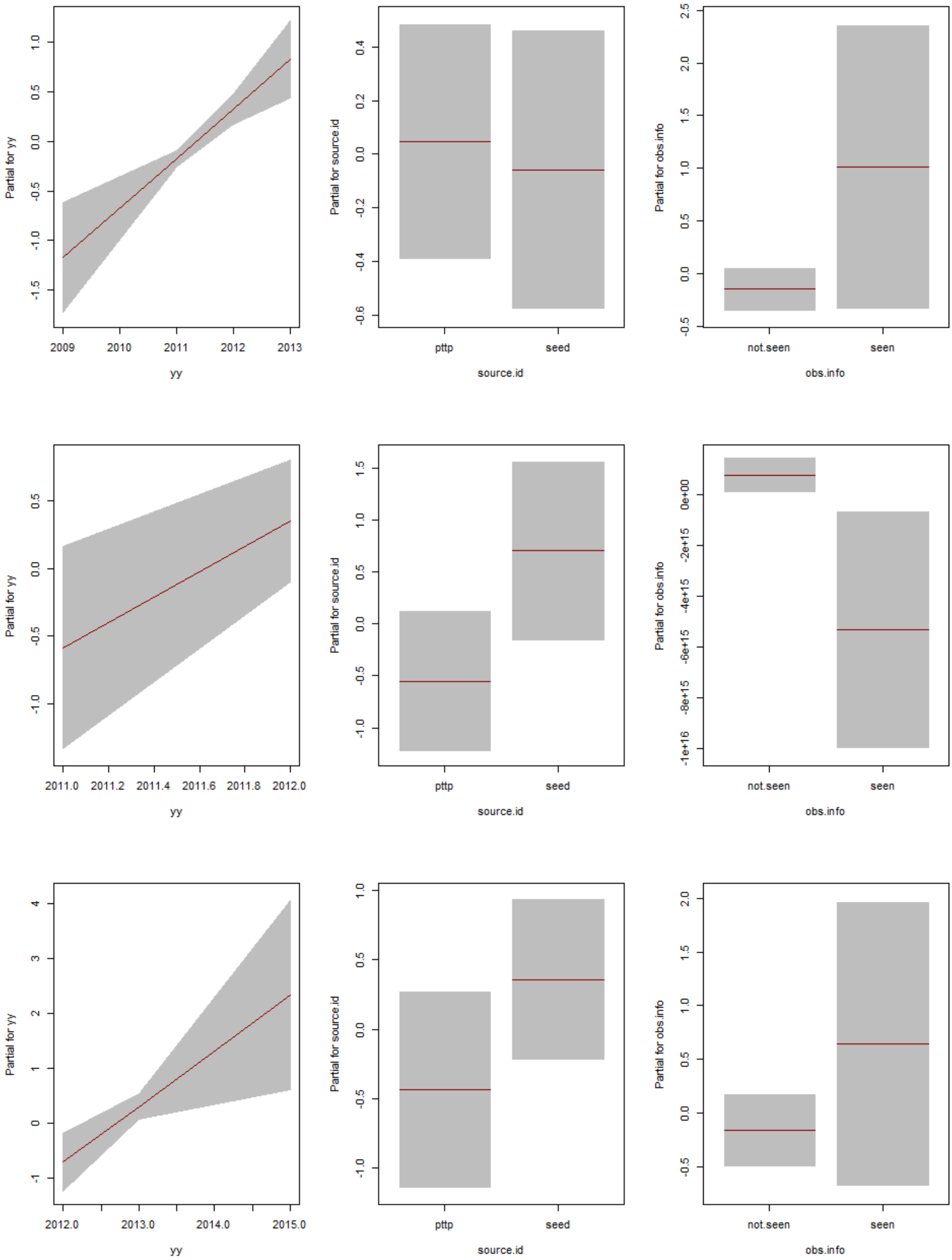

Figure 7 continued. Top row – Papua New Guinea fleet, middle row – Chinese Taipei; bottom row – Kiribati fleet.