CCSBT-CC/1610/BGD 02
(Previously CCSBT-ESC/1609/BGD 02 (Rev.1)

Ministry for Primary Industries
Manatū Ahu Matua

Australian Government
Department of Agriculture
ABARES

Updated estimates of Southern Bluefin Tuna Catch by CCSBT Non-Member states

August 2016

ESTIMATES OF SOUTHERN BLUEFIN TUNA CATCHES BY CCSBT NON-MEMBER STATES

Charles Edwards[1], Ashley Williams[2] and Simon Hoyle[3]

EXECUTIVE SUMMARY

This paper responds to a request from the Extended Commission (EC) to the Extended Scientific Committee (ESC) to further improve estimates of catch of southern Bluefin tuna (SBT) by non-Member fleets not reporting catch to the Commission for the Conservation of Southern Bluefin Tuna (CCSBT). We update and reconcile the results from previous work that applied different modelling approaches to provide preliminary estimates of potential non-Member catch of SBT. The general approach for both modelling approaches involved estimating the predicted catch rate from CCSBT data and applying that catch rate to non-Member effort in order to predict potential unreported catch.

Information on non-Member longline fishing effort in the Indian Ocean, the Western Pacific, and the Atlantic was obtained from the Indian Ocean Tuna Commission (IOTC), Western and Central Pacific Fisheries Commission (WCPFC), and the International Commission for the Conservation of Atlantic Tunas (ICCAT), respectively. Catch rates were estimated using two modelling approaches: generalised linear modelling and random forest regression, parameterised with the same data. In order to obtain a sufficiently large dataset of CCSBT catch and effort data with which to fit the models, we converted Japanese catches in number of fish to catches in weight, by modelling fish size patterns in space and time. We then modelled catch rates (in kilograms per hook) by year, month, vessel fleet (flag) and 5° grid, using each method. These predicted catch rates were applied to non-Member fishing effort by year, month, and 5° square in order to predict the unreported SBT catches.

When predicting the catches, it was necessary to make assumptions concerning the catchability of the non-Member fleets. Two alternative catchabilities were assumed, namely those of the Japanese and Taiwanese fleets, taken to represent alternative fishing behaviours (targeted and non-targeted respectively). These provided upper and lower bounds for the predicted catch. There are some differences between the results of the two modelling approaches, which are discussed.

INTRODUCTION

In 2013, the Extended Commission of CCSBT requested that the Extended Scientific Committee (ESC) provide advice on the potential impact of unaccounted mortality on the stock assessment and rebuild strategy for southern bluefin tuna (SBT).

The sources of unaccounted mortality included:

- Unreported catch or uncertainty in retained catch by Members from, for example:
  - surface fisheries,
  - artisanal catch,
  - non-compliance with existing measures (e.g. catch over-run);
- Mortality from releases and/or discards;
- Recreational fisheries;

[1] National Institute of Water and Atmospheric Research, 301 Evans Bay Parade, Wellington, New Zealand.
[2] Australian Bureau of Agricultural and Resource Economics and Sciences, Department of Agriculture and Water Resources, Canberra ACT 2601, Australia.
[3] National Institute of Water and Atmospheric Research, 217 Akersten St, Port Nelson, New Zealand.

- Catches by non-Members; and
- Any other sources of mortality that the Extended Scientific Committee is able to provide advice on (including depredation).

In 2014, the ESC19 noted that the impacts of unaccounted mortality on the stock assessment and rebuilding plan could be substantial, based on a range of projections from different unaccounted mortality scenarios. The ESC19 also noted that such scenarios were plausible given the available data, information, and anecdotal market reports. Based on this advice, the EC directed the ESC to undertake specific analyses to provide estimates of non-Member catch.

In 2015, two separate papers were presented to the ESC20 that provided estimates of non-Member catches of SBT (Chambers and Hoyle, 2015, Hoyle and Chambers, 2015). Each paper applied a different modelling approach and used a different subset of data. Predicted non-Member catches differed substantially between the two papers, but the reasons for the different estimates could not be resolved at the time. Furthermore, these papers did not consider non-Member catches of SBT in the Atlantic Ocean, so likely underestimated the total non-Member catch of SBT. The EC stressed the importance of obtaining the best possible estimates of non-Member catch, and requested the ESC to further improve estimates of non-Member catch and to report this information to the EC in a transparent manner. In particular, the EC requested that the method used for estimating non-Member catches needs be clearly described together with information on the fleets that are likely to catch SBT.

The objective of this paper is to respond to the request from the EC to the ESC by updating and reconciling the results from Hoyle and Chambers (2015) and Chambers and Hoyle (2015) and providing updated estimates of SBT catch by CCSBT Non-Member states. We apply the same two modelling approaches used by Hoyle and Chambers (2015) and Chambers and Hoyle (2015), but use the same variables in each model and apply them to the exact same dataset to enable valid comparisons of results between the two modelling approaches. Also, we extend the area in the analyses beyond the Indian and Pacific Oceans to include predicted catches from the Atlantic Ocean.

Cooperating non-Members report catch to CCSBT, and, in this paper, we have grouped them with Members in all analyses. The term “non-Member”, therefore, refers to non-cooperating non-Members unless specified otherwise. There is no reliable information available on SBT catch by non-cooperating non-Members. Information from a number of sources has indicated that a market for SBT exists in China. Although a small amount of catch in this market is supplied by catch from Members and cooperating non-Members, it may also be supplied with SBT that is not reported to CCSBT, since imports into China registered by the CCSBT have been found to be lower than total exports (CCSBT Secretariat, 2014).

Analysis of the effort data reported to the IOTC and WCPFC shows a large degree of overlap in SBT fishing grounds for these tuna fisheries (Larcombe, 2014). However, SBT catch by non-Members of CCSBT is not reported to WCPFC, even though these tuna fleets likely take quantities of SBT bycatch in the albacore, bigeye, and yellowfin target fisheries. Observer reports presented at the 2014 Scientific Committee of the WCPFC showed SBT catch on some trips in other tuna target fisheries, but only a very small proportion is reported. There may also be bycatch of SBT in pelagic longline fisheries in the south Atlantic. In general, the extent to which non-Member SBT catches are due to targeted or bycatch fishing is unknown.

OVERVIEW OF METHODOLOGY

The data preparation and analyses can be summarized in the following steps:

a) Obtain catch, effort, and size data from Member and cooperating non-Member states reporting to CCSBT by 5° square, year, and month, for the Pacific, Indian, and Atlantic Oceans.

b) Model length data in order to estimate catch weight in tonnes for CCSBT member fleets that report catches in numbers only (i.e. Japanese fleet).

c) Create adjusted CCSBT effort data for Japan, Korea, and Taiwan that includes unreported, zero-catch effort recorded in the WCPFC, IOTC, and ICCAT data bases.

d) Fit statistical models to catch and effort data for all CCSBT fleets using the adjusted or unadjusted effort, and estimate spatial and temporal covariates contributing to the catch per unit effort.

e) For each ocean use the model results to predict the non-Member SBT catch per unit effort for spatial (5° square) and temporal (year and month) strata, and based on two alternate assumptions: all non-Member effort has the same catchability as estimated for Japan, and; all non-Member effort has the same catchability as estimated for Taiwan. These fleets represent fisheries in which SBT is largely a target (Japan) or a bycatch species (Taiwan). This was repeated using model results obtained from fitting to either adjusted or unadjusted CCSBT effort data.

f) Obtain longline fishing effort reported to the WCPFC, IOTC, and ICCAT by non-Member, by 5° grid, year, and month.

g) Estimate catch for non-Member states by multiplying inferred catch rates by the effort per strata and summing across strata.

DATA PREPARATION

Spatial definitions

The CCSBT spatial strata are shown in Figure 1. The spatial definitions for each ocean, used throughout this analysis, are as follows:

	Latitude	Longitude	CCSBT area
Pacific Ocean	20°S to 55°S	150°E to 290°E	4, 5, 6, 7, 12
Indian Ocean	20°S to 55°S	20°E to 150°E	2, 3, 4, 7, 8, 9, 14
Atlantic Ocean	20°S to 55°S	290°E to 20°E	9, 10, 11, 15

These areas are non-overlapping, meaning that each unit of fishing effort can be assigned to only one of the three oceans. The areas of competence for the WCPFC, IOTC, and ICCAT largely fit within the Pacific, Indian, and Atlantic Ocean definitions given above. However there is a small area of overlap between the south-eastern edge of the IOTC area, which ends at 150°E, and the south-western edge of the WCPFC area, which ends at 141°E. This area of overlap forms part of CCSBT areas 4 and 7. In the current analysis, WCPFC data from this area were discarded in favour of data from the IOTC. This is because fewer data are omitted from the IOTC database for reasons of vessel identification (see below).

The latitudinal boundaries were set following inspection of the CCSBT catch data. SBT catches are bounded by latitude from approximately 20°S to 55°S (Figure 2), which was used as a consistent latitudinal boundary across oceans for the analyses. This excludes CCSBT areas 1, 13, and most of area 11. The spatial distributions of CCSBT effort and catch are shown in Figure 3 and Figure 4. SBT catches in the Atlantic are largely confined to area 9, which is contiguous with the Indian Ocean. For this reason, analyses were conducted for the Atlantic and Indian Oceans combined, with the Pacific analysed separately.

In CCSBT data, the latitude and longitude numbers indicate the north-western corner of a 5° grid square, while in the WCPFC data they indicate the south-western corner. For IOTC and ICCAT data, they indicate the corner closest to 0° latitude and 0° longitude. In this paper, all spatial data are managed at the 5° square level, and all latitudes and longitudes have been converted to indicate the centre of the grid square. Some ICCAT and IOTC data were recorded at higher levels of aggregation, such as a 10° grid square. In these few instances, all data were assigned to the mid-point of that square.

Effort data acquisition and preparation for the Pacific Ocean

Non-Member catch and effort data for the Pacific Ocean were obtained via public domain (PD) data requests to the WCPFC, and were received in two formats: 1) PD-flag stratified by year, month, 5° square, and flag, and 2) PD-total, stratified by year, quarter, 5° latitude, and Member/non-Member. Both public domain datasets omit strata that include fewer than three vessels in order to avoid potential identification (WCPFC, 2007), which meant that more data were omitted from the less aggregated dataset PD-flag. However, PD-flag dataset contains higher spatial resolution. Therefore, we used the PD-total dataset to calculate a multiplier, so that for the non-Member effort in PD-flag, for each year, quarter, and latitude band, the effort could be scaled to match the total effort in PD-total, whilst retaining the distribution across 5° grid and month. The total Pacific Ocean effort by year and area, calculated via this procedure, is shown in Figure 5.

Effort data acquisition and preparation for the Indian Ocean

Indian Ocean non-Member effort data were obtained from the IOTC website (http://www.iotc.org/documents/ce-longline). For IOTC data, in cases when an individual vessel can be identified, the data are aggregated prior to release by time, area, or flag to preclude such identification. Thus, no catch and effort data were omitted from the IOTC dataset. A small amount of IOTC effort was reported in days rather than hooks, and these were omitted.

Effort data acquisition and preparation for the Atlantic Ocean

Atlantic Ocean non-Member effort data were obtained from the Task II catch and effort database on the ICCAT website (https://www.iccat.int/en/accesingdb.htm). For longline, these data are aggregated by flag, year, month, and 5° latitude and 5° longitude grids. As for the WCPFC data, in order to avoid identification of vessel, data aggregations are only reported for a particular stratum if they contain observations from a minimum of three vessels. Unfortunately, for this dataset, no information on the total effort was available for scaling, which meant that total reported effort was likely underestimated. The total Atlantic Ocean effort by year and area is shown in Figure 6.

CCSBT catch, effort and size data

Catch and effort data for parties reporting to the CCSBT were obtained from the file 'CatchEffort_2016_January.txt' in the 2016 CCSBT data compilation CD. Data were prepared by extracting all records with gear code longline ('LL'), and removing records with missing values for year or effort.

The overall objective was to predict non-Member catches in weight by multiplying non-Member effort (in hooks) by predicted catch rates (in weight per hook). These predicted catch rates were estimated using statistical models parameterised with the CCSBT catch and effort data. Catches reported to CCSBT prior to 2007 are known to be unreliable (Polacheck, 2012). Therefore, we used data from 2007 onwards to parameterise the models and to predict non-Member catches.

Conversion from catch numbers to weight

Following the argument given by Hoyle and Chambers (2015), we considered Japanese catch and effort data to be essential for estimating predicted catch rates, because of the spatial and temporal coverage of the Japanese fleet, and their relatively consistent fishing methods. However, most of the Japanese data are reported in catch numbers, not weight, which made it necessary to convert the catches in number to catches in weight.

To estimate catch weights for the Japanese longline fishery, we needed to estimate the average weight per fish per strata. This was achieved by fitting a statistical model to the Japanese length frequency sampling data held by CCSBT.

The longline length sampling data held by CCSBT were obtained from the CCSBT website (http://www.ccsbt.org/site/sbt_data.php). Data were prepared by extracting records for the Japanese fleet with longline gear code 'LL', and by removing records with class precision > 2 cm. All years of length data were used (1965 - 2014). The average length in each length class was assumed to be the middle value, i.e. 107.5 cm for a fish in the length class 108 cm with class precision of 1 cm, since the label indicates the upper end of the length class. When preparing the data for analysis, length records were replicated according to the 'adjusted frequency'. This involved randomly sampling the length records with replacement, with a probability proportional to the frequency and with the number of samples thinned due to computational memory constraints. These length data were then converted to weight using the length to processed weight conversion factors agreed at the 1994 SBT Trilateral Workshop on Age and Growth, 17 Jan - 4 Feb, 1994 (Table 1). Processed weights were converted to whole weight by adding 15%, as agreed at the 1994 workshop.

To allow prediction, we applied a general linear model (GLM) implemented in R 3.2.4 (R Core Team, 2016) to model the square root transform of the weight per fish as a function of year, month, and statistical area:

$$E\left[\sqrt{weight_{per\ fish}}\right] \sim year + month{:}\,area \qquad (1)$$

where E[.] refers to the expectation. The month and the statistical area were combined into a categorical variable 'month:area' to avoid problems with interaction terms, since there were different amounts of data across months in different statistical areas. The model assumed that inter-annual variation was consistent for month:area combinations. There were statistically significant interactions between year and month:area effects, but these were ignored to enable sizes to be predicted for more strata. Applying a square root transformation to the weights helped to normalize the residuals.

The above model was fitted to two subsets of the CCSBT size sampling data, corresponding to the Pacific Ocean and the Indian and Atlantic Oceans combined. Residuals from the generalized linear models were relatively normally distributed after transformation, and there was only limited variation in variance among statistical areas (see Figure 19 and Figure 20 in the Appendix). Estimates were unavailable for several month-by-area combinations that lacked size sampling data, and were copied from adjacent months for the same statistical areas. Estimates of mean weight were obtained for the majority of statistical area-month combinations, and inferred for the remainder (see Figure 21 and Figure 22 in the Appendix). This allowed mean weights to be predicted for year by month by statistical area strata, and shown to be a reasonable fit to the empirical data (see Figure 23 and Figure 24 in the Appendix).

The model could then be used to predict the mean weight per fish for each year and month:area combination, in each of these two regions. However, the model was only used to predict the mean weight for strata containing relatively few empirically measured fish. Specifically, strata with total adjusted size sampling frequencies of at least 100 fish were assigned the empirical mean weight, whilst strata with adjusted frequencies of fewer than 100 fish were assigned a predicted mean weight based on the fitted model.

When using the model for prediction, due to the distribution of the data and the square root transformation, back-transformed nominal mean weights tended to be lower than the true mean. We removed this bias by, for each stratum, sampling 2000 residuals with replacement and adding them to the predicted mean to generate 2000 parametric bootstrap samples, back-transforming by squaring the samples, and taking the mean of the back-transformed samples as the predicted weight for the stratum. This bias correction approach is analogous to smearing (Duan, 1983).

For all CCSBT effort that reported SBT catch in retained number but not weight (predominantly Japanese but also some from South Africa), predicted retained weights were calculated by multiplying retained numbers by estimated mean weight per fish for the appropriate stratum. In order to examine the accuracy of the estimation process, we used the same approach to predict retained weights for CCSBT effort that reported catch in both numbers and weight. We plotted these results by flag and ocean, with observed weight plotted against predicted weight. These are shown in Figure 7 and Figure 8.

The relationship between predicted and observed retained weight varied between flags, with predicted weights very close to observed weights for Japan in both the Pacific Ocean (Figure 7) and Indian and Atlantic Oceans combined (Figure 8). In the Indian and Atlantic Oceans, the predicted weights were, on average, similar to reported weights for Korea, higher than reported for Taiwan, slightly higher for Australia, and lower for South Africa. For the Pacific Ocean, the predicted weights were similar to reported weights for Korea, slightly higher than reported for Australia, and lower than reported for New Zealand and Taiwan. These differences between predicted and reported weight may reflect differences in average fishing locations and fishing behaviour between fleets. The predictive model takes into account year, month, and statistical area, but there are also consistent differences in mean weight within statistical areas that the model does not take into account.

Most importantly, the predictions of catch weight for Japanese effort appear to be sufficiently reliable for reconstruction of the catch rate data for use in downstream model-based prediction of the non-Member catch rate, with the proviso that observed weights were only available in the Japanese data for comparison for a single year.

Adjustment for unreported effort

The reporting of effort by Japan, Taiwan, and Korea depends on the spatio-temporal strata in which fishing takes place. If fishing occurs in areas 4 to 9, during months 4 to 9, then all effort and catch is reported to CCSBT. These “core” strata typically account for the majority of the catch. However, if fishing occurs outside of these core strata then effort is only reported when there is a positive catch of SBT. This means that in the more lightly fished areas and outside of the normal SBT season, catch rates for Japan, Taiwan, and Korea may be overestimated because of unreported zero-catch effort.

To adjust the CCSBT effort to account for unreported effort we used the effort reported to the WCPFC, IOTC, and ICCAT. For each year, month, and grid cell combination, outside of the core strata, and for Japan, Korea, and Taiwan only, we compared the effort reported to the CCSBT with the effort reported to either the WCPFC, IOTC, or ICCAT as appropriate. For each comparison, the maximum effort value was selected. Catches were not adjusted since CCSBT data were assumed to include all SBT catch. For each adjustment, we were able to record a reporting rate as the ratio of reported to unreported effort. Average reporting rates for the Indo-Atlantic and Pacific Oceans are shown in Figure 9. Raw and unadjusted effort for the Japan, Korea, and Taiwan fleets is shown in Figure 10.

The application of this approach to adjust the effort created an additional data set with which to estimate the catch rate. Although we consider the adjusted effort values to give a more accurate reflection of the catch rate, both the adjusted and unadjusted data were nevertheless carried through and analysed in parallel.

ANALYSES AND RESULTS

Non-Member catches were estimated by predicting catch rates per stratum from the CCSBT member data, and multiplying by non-Member effort in the same strata. Two modelling approaches were applied to predict non-Member catches: generalised linear models (GLM) and random forest (RF) regression. These models were applied to the same data sets and used the same response and predictor variables to enable a valid comparison of the results from the two modelling approaches. These analyses were conducted separately for Pacific Ocean data and the Indian and Atlantic Oceans combined.

Spatial overlap

The spatial distribution of the non-Member effort obtained from the WCPFC, IOTC, and ICCAT has been mapped in Figure 11. This distribution overlaps substantially with the reported effort by CCSBT parties shown in Figure 3.

Generalised linear model (GLM)

Catch rate estimation

In the GLM approach, catch rates were estimated using a two-part model fitted to the data to estimate the effects of year, quarter, month, flag, and 5° square. As in previous analyses (Hoyle and Chambers, 2015) the probability of zero catch is modelled using a binomial model, but we replace the lognormal transformation for positive catches with a power transformation. The previous analyses added a constant to the CPUE to normalize the residuals, but this led to negative predicted catches in some strata. A power transformation of $CPUE^{1/5}$ effectively normalized residuals.

The year and quarter were combined into a single categorical variable 'year:qtr', and the 5° square was defined as a categorical variable by its combined coordinates 'lat:lon'. A cubic spline *ns()* with 4 degrees of freedom was used to describe the influence of month, treated as a continuous variable. We also included a 'core' covariate to identify whether the effort took place within or outside of the core spatio-temporal strata.

The approach involves first modelling the probability of a non-zero catch with a binomial GLM, and then modelling the distribution of CPUE for non-zero catches with a power-transformed normal model. It can therefore be written as:

$$P[\mathrm{weight}_{\mathrm{hooked}} > 0] \sim year{:}qtr + core + ns(month) + flag + lat{:}lon + ns(hooks) \quad \text{Model A}$$

$$E\left[\left(\frac{\mathrm{weight}_{\mathrm{hooked}}}{hooks}\right)^{1/5} \mid \mathrm{weight}_{\mathrm{hooked}} > 0\right] \sim year{:}qtr + core + ns(month) + flag + lat{:}lon \quad \text{Model B}$$

where P[.] is the probability, and E[.] is the expectation. In this approach the number of hooks was also included as a predictor for the binomial model described by a cubic spline, with 20 degrees of freedom for the Pacific Ocean, and 3 degrees of freedom for the combined Indian and Atlantic Oceans. This was because records with more effort were expected to be more likely to include non-zero catch. We conducted two analyses, using either the adjusted or unadjusted CCSBT effort records when modelling the catch rate, and in both instances all records with zero effort were ignored.

To estimate non-Member catches, we predicted total hooked weight per unit effort using the estimated coefficients and associated covariates in the non-Member effort data. If we write:

$$\pi = P[\text{weight}_{\text{hooked}} > 0]$$

$$\mu = E\left[\left(\frac{\text{weight}_{\text{hooked}}}{hooks}\right)^{1/5} \mid \text{weight}_{\text{hooked}} > 0\right]$$

then, using a non-parametric smearing approach to back-transform the model prediction (Duan, 1983), the predicted catch rate per unit of effort is:

$$E\left[\frac{\text{weight}_{\text{hooked}}}{hooks}\right] = \frac{\sum_{i=1}^{N} \pi.(\mu + e_i)^5}{N}$$

where e_i is an error residual sampled from the distribution of Model B, and N is a large number, 2000 in this case. Values of μ and π for each data record where obtained using the 'predict.glm' function in R, applied to each model in turn. Because the number of data records used to parameterise Model B is smaller than the number of data records used in Model A, in some cases, a particular strata level returned an estimated coefficient for the latter model but not the former. In these cases it was assumed that $E\left[\frac{\text{weight}_{\text{hooked}}}{hooks}\right]$ = 0.

Standardizations for both Pacific and Indian Ocean CCSBT data fitted the data relatively well, but with positively biased residuals for small numbers of hooks, particularly in the Indian Ocean (see Figure 25 and Figure 26 in the Appendix). The spatial effects showed the expected patterns of higher catch rates further south (Figure 12 and Figure 13).

Catch prediction

Catches were predicted by aggregating the non-Member effort data by stratum, predicting the catch rate using the model-based procedure above, and multiplying this expectation by the effort. Predictions were made for the most recent eight years of data (2007 to 2014). Areas without reported effort by Members at any time could not be allocated a catch rate. They were assigned an SBT catch rate of zero, on the assumption that CCSBT Members as a whole have fished in all areas in which SBT can be taken in significant numbers.

We checked the estimates by predicting catches for member fleets using the CCSBT input data, and comparing them with reported catches. These are shown in Figure 14 and Figure 15 assuming the adjusted CCSBT effort data. Catch predictions with CCSBT data gave total catch estimates that were close to the observed estimates and without significant bias. This result suggests that the model is acceptable for predicting non-Member catch.

Finally, non-Member catch was estimated for the Pacific, Indian, and Atlantic Oceans, by year and by statistical area, and for each catchability assumption (Japanese or Taiwanese), by multiplying non-Member effort by the predicted catch rate per stratum, and summing across strata to produce estimates per year and statistical area (Table 2 and Table 3). This procedure was repeated using the catch rate model estimates obtained from both the adjusted and unadjusted CCSBT effort data.

Random forest analysis

Catch rate estimation

Catch rates were predicted using the random forests machine learning algorithm, similar to Chambers and Hoyle (2015). Random forests involve fitting an ensemble of regression trees, then averaging the predictions across all trees. The random forest algorithm first selects many (e.g. >1000) bootstrap samples from the data, each of which contains ~63% of the original observations (Strobl et al., 2009).

Observations that are not selected in each bootstrap sample are referred to as out-of-bag observations. A regression tree is then fitted to each bootstrap sample, but only a subset of randomly selected predictor variables are used at each node. The trees are fully grown with no pruning, then each tree is used to predict the out-of-bag observations. The predicted value of an observation is calculated by averaging the out-of-bag predictions for that observation across all trees. The out-of-bag estimates are considered a cross-validation of the accuracy of estimates because they are not used in the fitting of trees. The relative importance of each predictor variable is then determined from the misclassification rate for the out-of-bag observations.

Random forests is a non-parametric modelling approach that has considerable flexibility for handling correlated variables and complex non-linear interactions (Strobl et al., 2009). Therefore, it was not necessary to create single categorical variables for year and quarter (year:qtr), and 5° squares (lat:lon) as was done for the GLM approach. We used the *randomForest* R-package (Liaw and Wiener, 2002) to predict catch rates of SBT. In contrast to the random forest model used by Chambers and Hoyle (2015), we fitted the random forest model to the same predictor variables that were used in the GLM: year, quarter, month, flag, latitude, and longitude, to enable a valid comparison of results between the two modelling approaches. The random forest model can be written as:

$$\left(\frac{weight_{hooked}}{hooks}\right) \mid hooks > 0 \sim year + qtr + month + flag + lat + lon$$

All of the predictor variables were treated as continuous variables except flag, which was a categorical variable. A sufficiently large number of trees (1000) were used, and different random seeds were applied for the analysis of each ocean to ensure stability in variable importance and predictions. We found only trivial differences in results when using different numbers of predictor variables at each tree node. Therefore, we used the default value applied in the *randomForest* package (number of predictor variables divided by 3).

The effect of adjusting the CCSBT effort data had a minimal effect on the relative importance of variables in predicting catch rates of SBT (Figure 27). The relative importance of variables varied between oceans, but latitude was most important for all oceans. Flag and month were important for the Indian and Atlantic Oceans, while flag and year were important for the Pacific Ocean (Figure 27).

Fitted random forests models are difficult to interpret comprehensively (Prasad et al., 2006). The partial effects plots (Figure 28) provide some indication of the effects of individual predictor variables of the model on SBT catch rates in the Pacific and Indian and Atlantic Oceans. The effect of adjusting the CCSBT effort data had a minimal effect on the partial effects in all oceans. The partial effects of year, latitude, month, and flag were similar among oceans. The partial effect of year increased markedly since 2007, which is consistent with various monitoring series for SBT CPUE (Chambers, 2014, Itoh and Takahashi, 2014). Similarly, the partial effect of latitude increased with latitude which is consistent with higher catch rates south of 35°S. The partial effect of month was generally greater during the austral winter and spring, consistent with the period of greater SBT catches. The partial effect of flag indicated that catch rates of SBT by the Taiwanese fleet were different to other fleets.

The spatial patterns in predicted catch rates for SBT from the random forest models were similar to those from the GLMs, and showed the expected patterns of higher catch rates in more southern latitudes (Figure 12 and Figure 13).

Catch prediction

Catch predictions derived from the random forest model followed a similar procedure as the GLM approach described above. That is, catches were predicted by multiplying the aggregated non-Member effort data by stratum by the predicted catch rates from the random forest model generated by using the *randomForest* 'predict' function. Predictions were made for the years 2007 to 2014 inclusive, and areas without reported effort by Members were assigned an SBT catch rate of zero.

We checked the estimates by predicting catches for member fleets using the adjusted CCSBT effort data, and comparing these predictions with reported catches (Figure 14 and Figure 15). Catch predictions with CCSBT data produced total catch estimates that were close to the reported estimates and without significant bias, indicating that the model is acceptable for predicting non-Member catch.

Finally, non-Member catch was predicted for the Pacific, Indian, and Atlantic Oceans, by year and by statistical area, for each catchability assumption (Japanese or Taiwanese), by multiplying non-Member effort by the predicted catch rate per stratum, and summing across strata to produce estimates per year and statistical area (Table 4 and Table 5). This procedure was repeated using the catch rate model estimates obtained from both the adjusted and unadjusted CCSBT effort data.

DISCUSSION

The estimation approaches applied here generate catch predictions based on the assumption that non-Member catch rates match those of Members, but targeted effort will catch more SBT than non-targeted effort. The higher Japanese catch rates represent targeted effort, while the lower Taiwanese catch rates represent non-targeted effort. This partitioning of effort types is an approximation, because a number of different fishing and targeting practices occur in the areas of interest. However, the main aim of this study was to identify the approximate plausible range of catches. A more comprehensive exploration of the possible fishing methods and catch rates would require access to operational data, and was beyond the scope of this study.

An additional uncertainty included in the results concerns unreported effort by some of the CCSBT Member vessels, which may have biased our estimate of the catch rate. This is represented in the analysis by using either the adjusted or unadjusted CCSBT effort data to estimate the catch rate, and this distinction is retained for the prediction of catches by the non-Members. This adjustment may be incomplete, because it does not account for zero-catch effort that may be missing from the effort data reported to CCSBT by Australia and New-Zealand. Furthermore, we are unable to verify whether the unreported effort by Japan, Korea, and Taiwan is comparable to the reported effort in terms of the fishing practices employed. In a manner similar to our alternate assumptions concerning the catchability of non-Member fleets, we therefore chose to present both sets of results as representative of approximate bounds on the true catch rate.

A comparative summary of the results from the GLM and random forest approach is given in Figure 18. The estimates are similar, but some differences remain, despite a consistent use of the same data to parameterise the model in each case. The magnitude of these differences depends on the catchability assumption, and is different for each ocean. Identifying the precise cause of these discrepancies is beyond the scope of this study, and the results are presented here as equally valid alternatives.

There are a number of differences between the random forest and GLM methodologies. The random forest approach is able to model interactions between effects such as time, area, and flag, while the GLM models assume no interaction terms. These interactions are traded off against less flexible main effects, since the random forest model uses continuous time (year-qtr) and spatial (latitude and longitude) terms, while the GLM is fitted with categorical year-qtr and 5° square effects. The random forest estimates for the member data appear slightly better than the GLM estimates (Figures 14 and 15), but this does not necessarily imply better estimates for the non-member data. A useful comparison would be to apply a General Additive Model (e.g. mgcv) with smoothers on the temporal and spatial effects, to bring the GLM approach closer to the random forest approach.

The GLM modelling approach assumes that, apart from differences between core and non-core areas, relative catch rates among locations do not change by quarter, whereas SBT move seasonally, which should result in seasonal changes in the relative spatial catch rates. However, although season-location interactions were statistically significant in the catch rate model the effect was quite small, and adding these interaction terms resulted in the loss of catch rate estimates for relatively large areas which may have affected the ability to predict catches. We therefore chose to omit the interactions. If effort by CCSBT Members is higher in areas and seasons with higher SBT catch rates, ignoring interactions is likely to bias estimates towards periods with higher catch rates, which (given the spatial distribution of effort) may positively bias estimates of non-member catch**.**

Acknowledgments

The analyses described in this report are based on publicly available data from the IOTC, WCPFC, and ICCAT. The analyses also use the CCSBT data provided by Members and cooperating non-Members. Thanks to Colin Millar of the CCSBT Secretariat for helping to obtain the WCPFC catch and effort data, and to Peter Williams and Manu Schneiter of the Secretariat of the Pacific Community for preparing the WCPFC data and advising us on how to use it. This work was funded by the Ministry for Primary Industries (New Zealand), and the Australian Bureau of Agricultural and Resource Economics and Sciences' Fisheries Resources Research Fund. We are grateful for a helpful review of this document by Owen Anderson (NIWA, Wellington, New Zealand).

References

CCSBT SECRETARIAT 2014. Southern Bluefin Tuna Trade data: Annual analyses. *Ninth Meeting of the CCSBT Compliance Committee.* Auckland, New Zealand.

CHAMBERS, M. & HOYLE, S. 2015. Estimates of non-Member catch of SBT in the Indian and Pacific Oceans. *20th Meeting of the CCSBT Extended Scientific Committee, Incheon, South Korea, August 2015.* Australian Bureau of Agricultural and Resource Economics and Sciences, Australia.

DUAN, N. 1983. Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association,* 78**,** 605-610.

HOYLE, S. & CHAMBERS, M. 2015. Estimating Southern Bluefin Tuna Catches by Non-Members of CCSBT. *20th Meeting of the CCSBT Extended Scientific Committee, Incheon, South Korea, August 2015.* Ministry for Primary Industries, New Zealand.

ITOH, T. & TAKAHASHI N. 2014. Update of the core vessel data and CPUE for southern bluefin tuna in 2014. 5th Meeting of the CCSBT Operating Model and Management Procedure Technical Committee. Seatlle, USA.

LARCOMBE, J. 2014. Fleet overlap in the IOTC area. *19th Extended Scientific Committee of the CCSBT.* Auckland, New Zealand.

LIAW, A., & WIENER, M. 2002. Classification and regression by randomForest. R News, 2, 18–22.

POLACHECK, T. 2012. Assessment of IUU fishing for Southern Bluefin Tuna. *Marine Policy,* 36**,** 1150-1165.

PRASAD, A.M., IVERSON, L.R. & LIAW, A. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181-199.

R CORE TEAM 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

STROBL, C., MALLEY, J., & TUTZ, G. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, andrandomforests. Psychological Methods, 14, 323–348.

WCPFC 2007. Rules and Procedures for the Protection, Access to, and Dissemination of Data Compiled by the Commission (as revised by WCPFC4 2007).

Figures

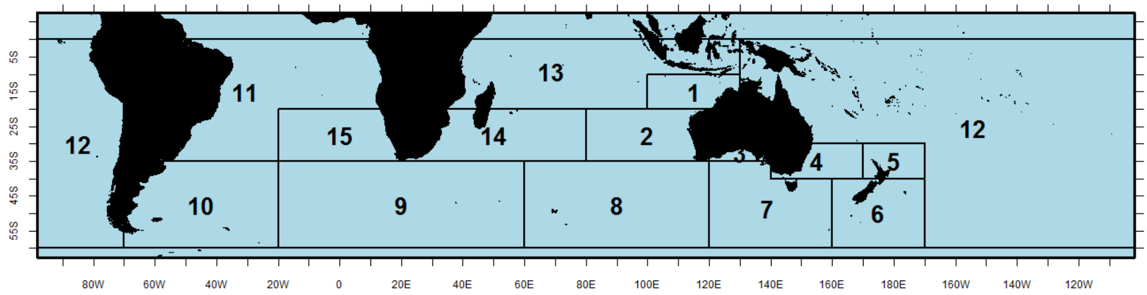


Figure 1. Map showing the CCSBT statistical areas

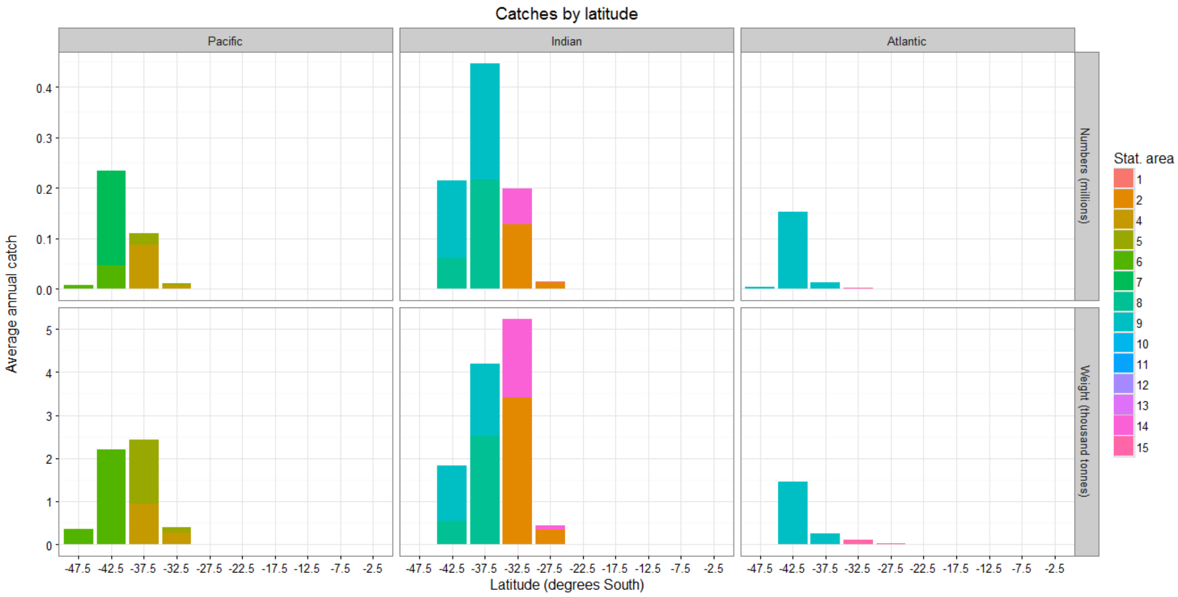


Figure 2. Average annual SBT catches by latitude for each ocean and statistical area, for 2007 to 2014, reported as numbers and weight.

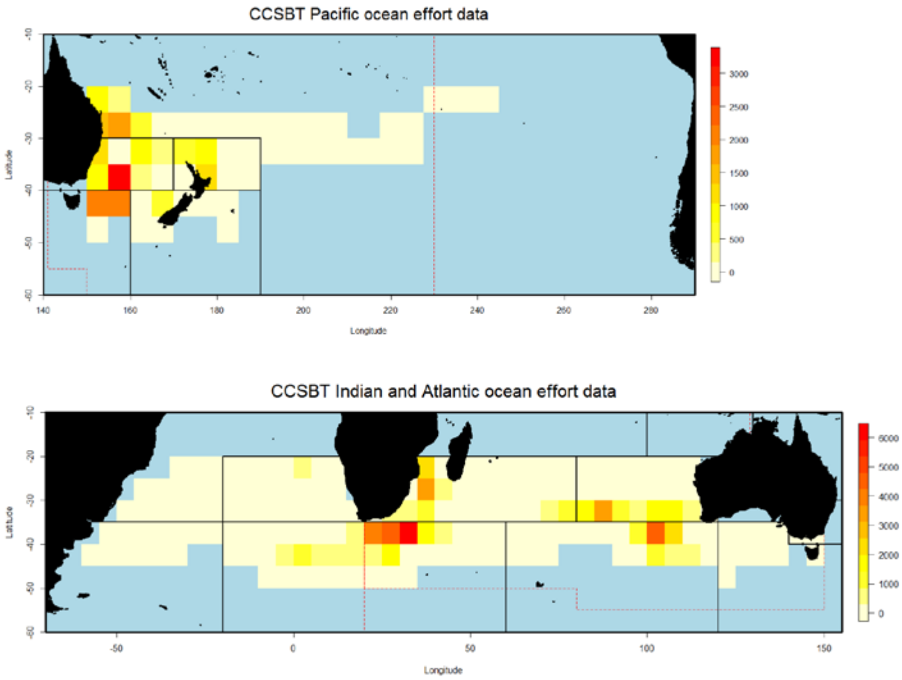


Figure 3. Average annual effort (in thousands of hooks) reported to CCSBT since 2007 inclusive, mapped by 5° square. The data have been trimmed spatially according to the limits defined on page 4. CCSBT statistical areas are shown as solid lines (see Figure 1); the approximate boundaries for each Regional Fishery Body are shown as dashed lines.

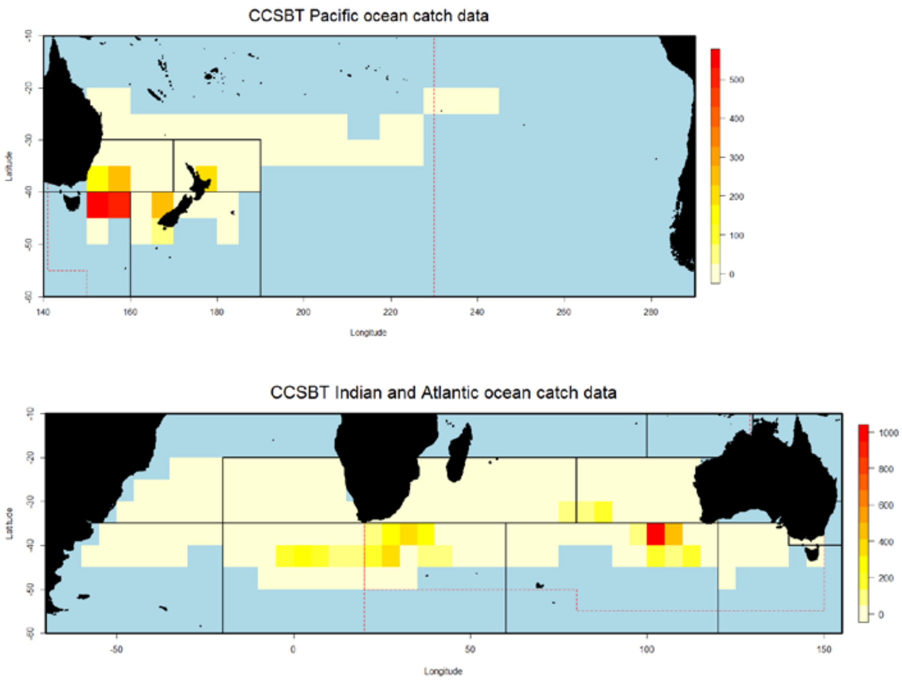


Figure 4. Average annual catch (in thousands of tonnes) reported to CCSBT since 2007 inclusive, mapped by 5° square. All catch data reported as numbers have been converted to weight. The data have been trimmed spatially according to the limits defined on page 4. CCSBT statistical areas are shown as solid lines (see Figure 1); the approximate boundaries for each Regional Fishery Body are shown as dashed lines.

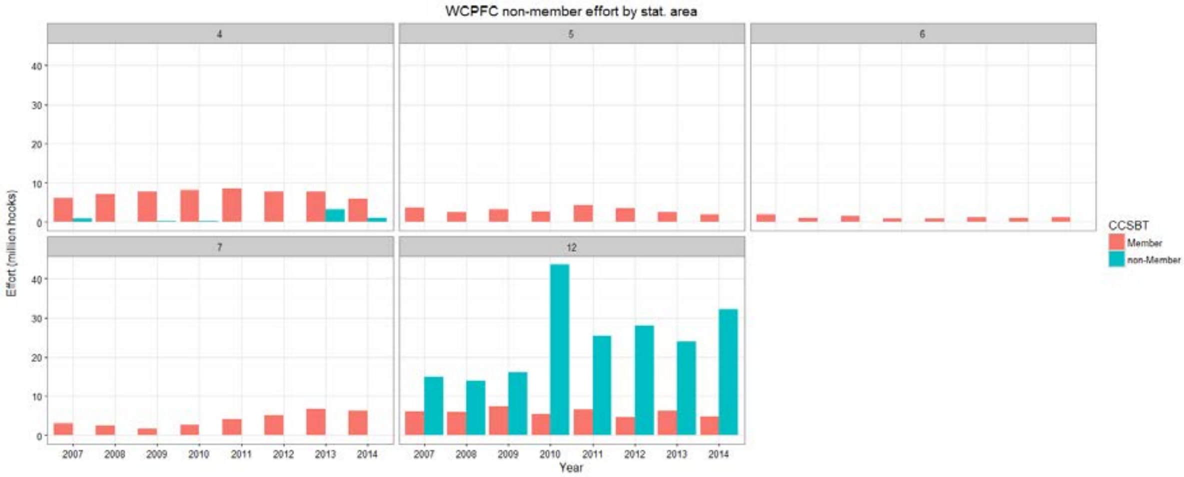


Figure 5. Public domain non-Member effort reported to the WCPFC since 2007. Unadjusted CCSBT effort is also shown for comparison. No non-Member effort has been reported to the WCPFC in areas 5, 6 and 7 (bounded at 150°E). In areas 4 and 12 average non-Member effort is approximately 0.7 and 24.7 million hooks per year, respectively.

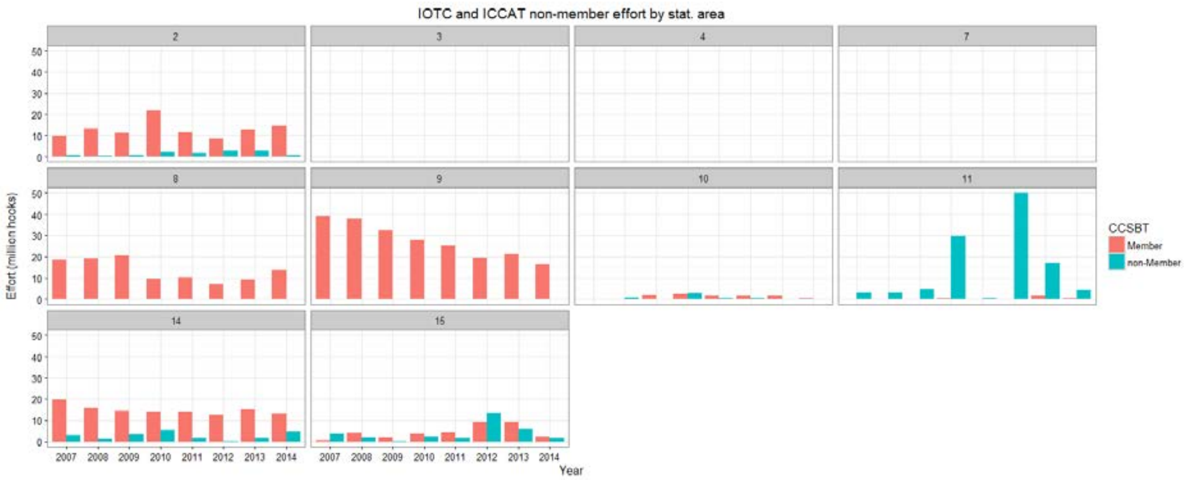


Figure 6. Non-Member effort reported to the ICCAT and IOTC since 2007. Unadjusted CCSBT effort is also shown for comparison. There is no CCSBT effort in area 3 and for areas 4 and 7 (bounded at 150°E) it is relatively small, being 4 and 23 thousand average hooks per year. There is no ICCAT or IOTC non-Member effort in these areas. In the remaining areas, average non-Member effort is: area 2, 1.5 million hooks; area 8, 2 thousand hooks; area 9, 59 thousand hooks; area 10, 0.6 million hooks; area 11, 14 million hooks; area 14, 2.6 million hooks; area 15, 3.8 million hooks.

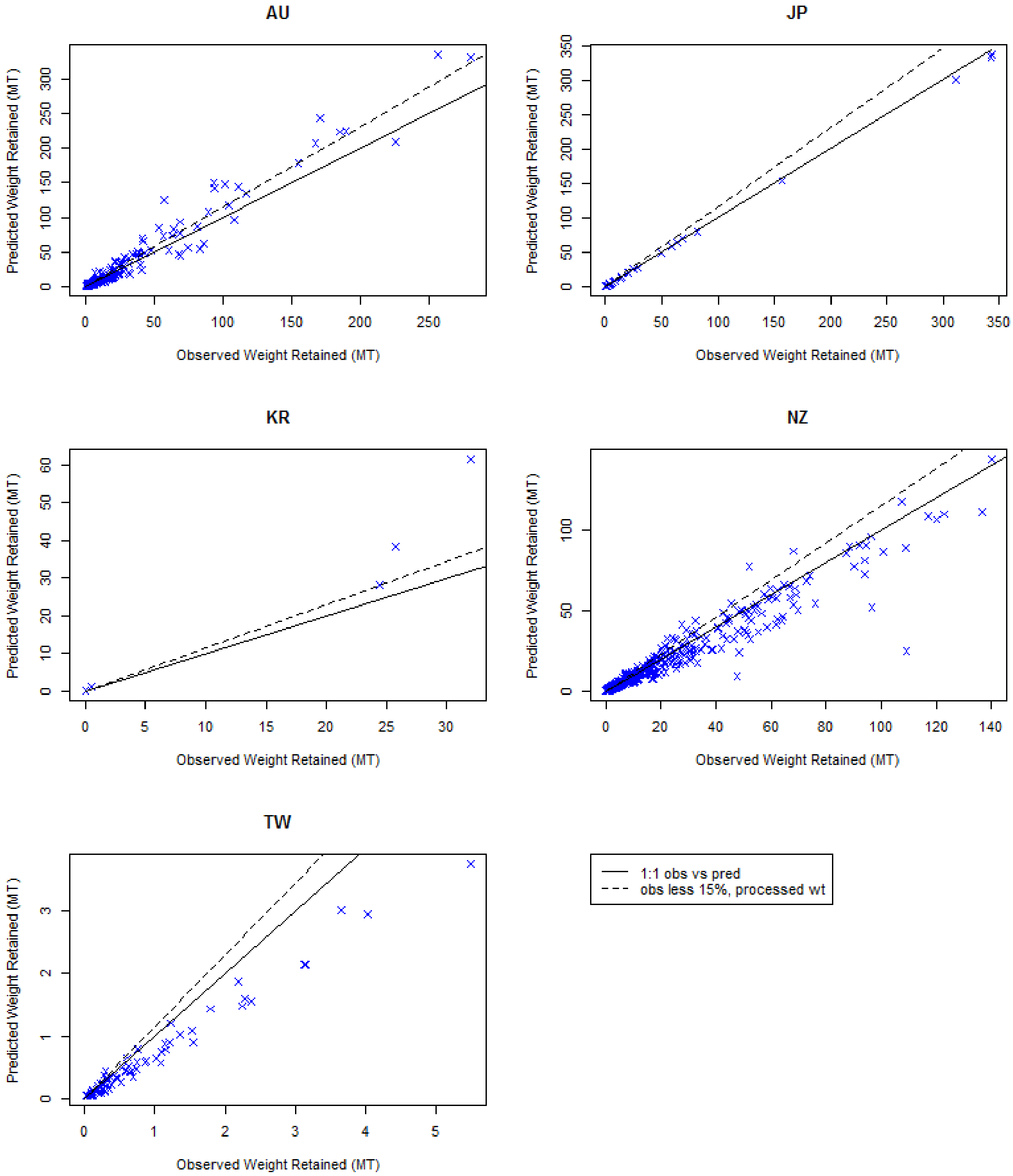


Figure 7. Comparisons of observed and predicted catches in metric tonnes (MT) by flag in the Pacific Ocean, with predictions based on multiplying numbers caught by mean observed or predicted weights in the Japanese catch for each stratum (year-month-statistical area). Flags are Australia (AU), Japan (JP), Republic of Korea (KR), Taiwan (TW), and South Africa (ZA).

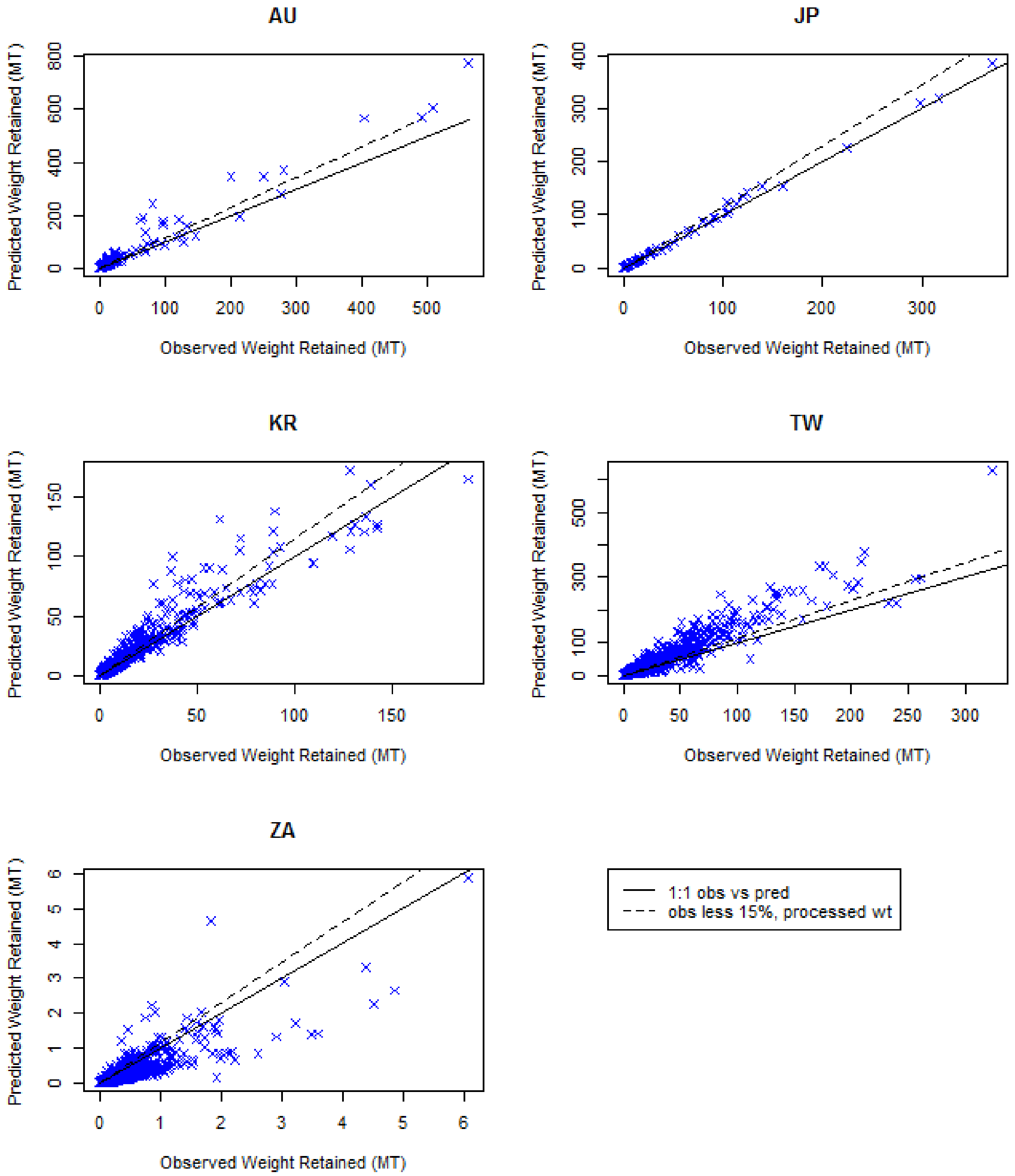


Figure 8. Comparisons of observed and predicted catches in metric tonnes (MT) by flag in the Indian and Atlantic Oceans, with predictions based on multiplying numbers caught by mean observed or predicted weights in the Japanese catch for each stratum (year-month-statistical area). Flags are Australia (AU), Japan (JP), Republic of Korea (KR), New Zealand (NZ), and Taiwan (TW).

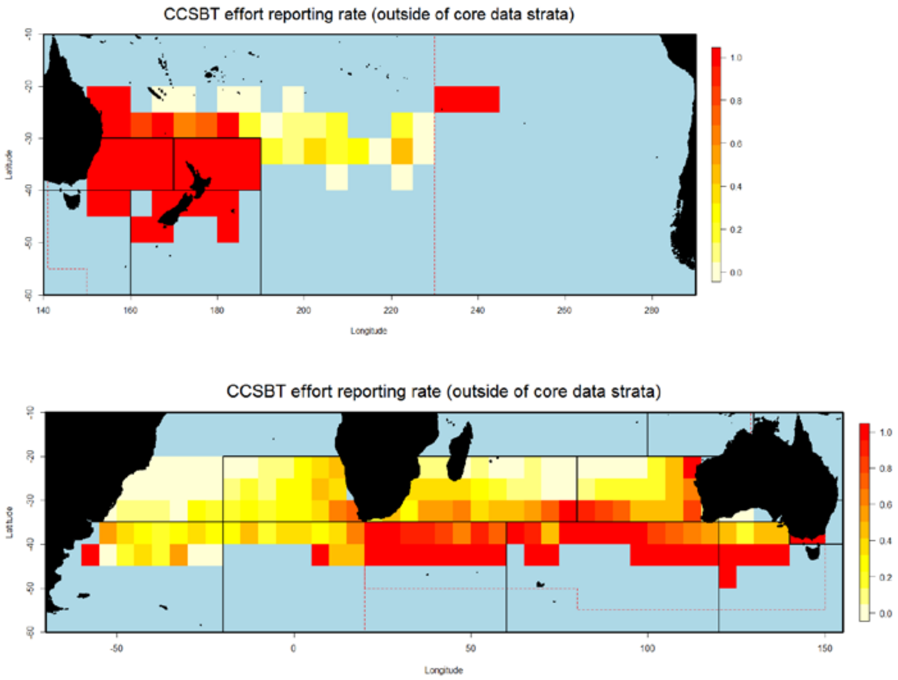


Figure 9. Effort reporting rates of Members to CCSBT outside of the core spatio-temporal strata, shown as an average across all fleets and years (2007 to 2014). Low reporting rates indicate a high proportion of zero catches, which are not reported to the CCSBT by JP, KR and TW.

Figure 10. Adjusted effort for JP, KR and TW, shown by statistical area for years 2007 to 2014.

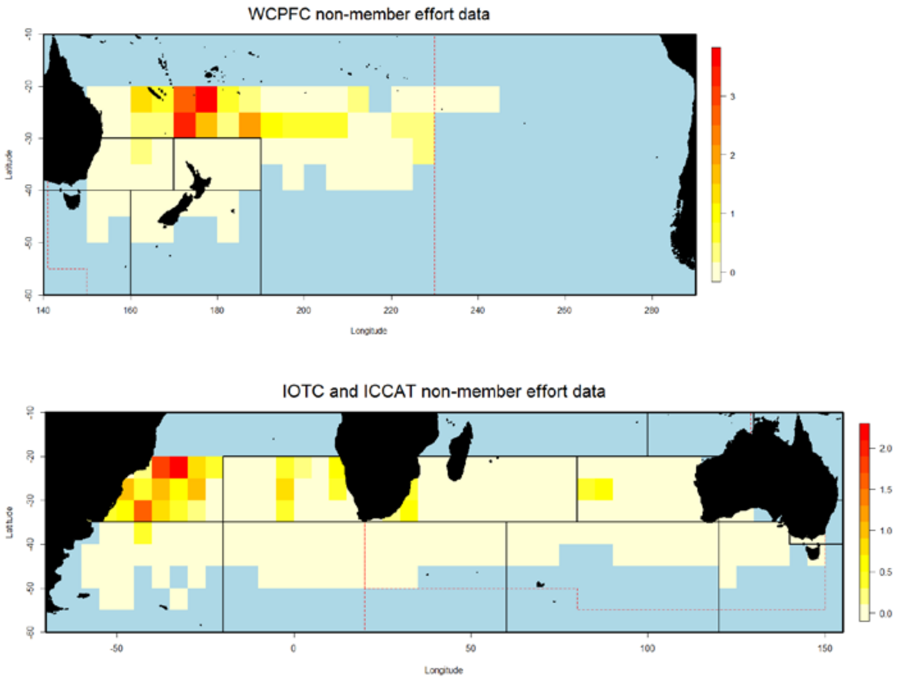


Figure 11. Average annual non-Member effort (in thousands of hooks) for 2007 to 2014 reported to Regional Fisheries Bodies in the Pacific and Indo-Atlantic Oceans.

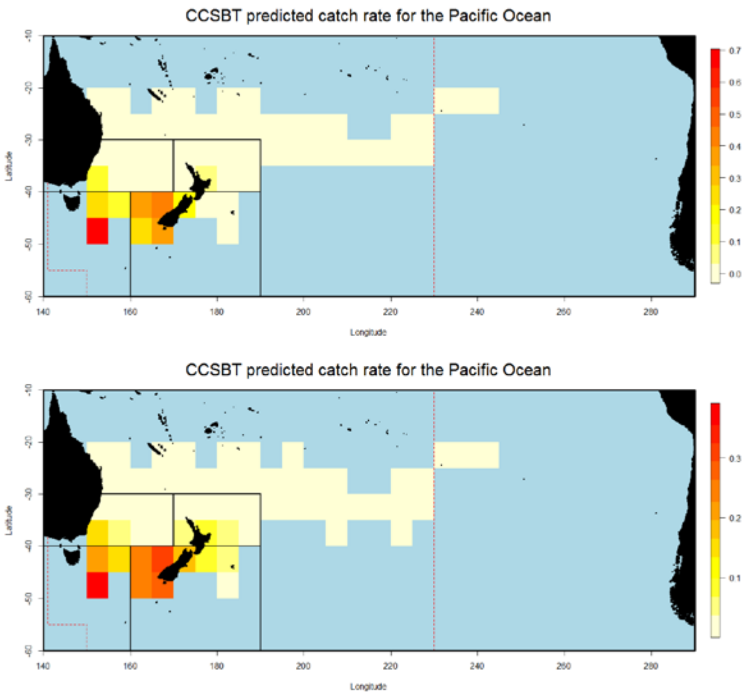


Figure 12. Average annual predicted CCSBT catch rates of southern Bluefin tuna by 5° square in the Pacific Ocean, in kilograms per hook, from the GLM (top) and random forest (bottom) analyses, assuming an adjusted CCSBT effort rate.

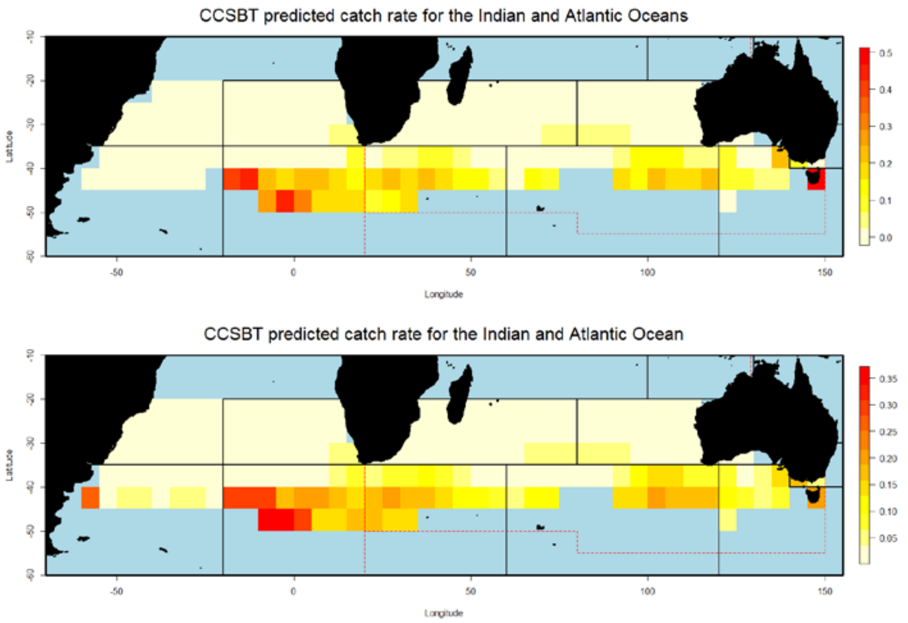


Figure 13. Average annual predicted CCSBT catch rates of southern Bluefin by 5° square in the Indian and Atlantic Oceans, in kilograms per hook, from the GLM (top) and random forest (bottom) analyses, assuming an adjusted CCSBT effort rate.

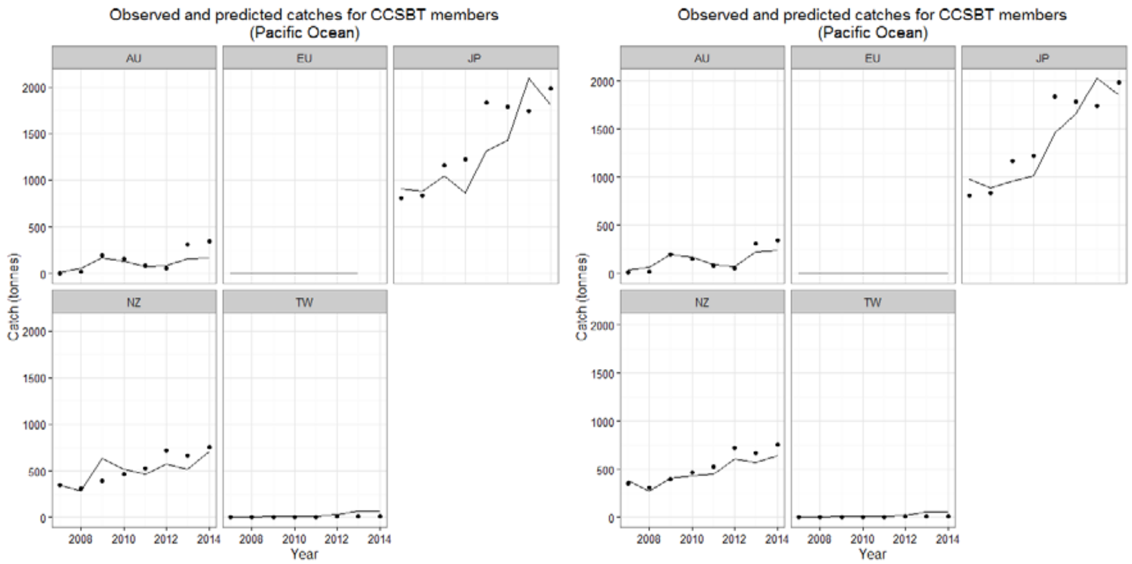


Figure 14. Comparison of observed (circles) and predicted (lines) catches in the Pacific Ocean for CCSBT Members, from the GLM (left) and random forest (right) models, based on multiplying the predicted catch rate in kilograms per hook by adjusted effort in the CCSBT data. No observed catch data were available for the EU fleet, and the predictions shown assume TW catchability. Using the GLM, the average annual predicted EU catch was less than one tonne (and less than two tonnes assuming JP catchability). Similarly, using the random forest model, the average predicted EU catch was less than one tonne (and less than two tonnes assuming JP catchability).

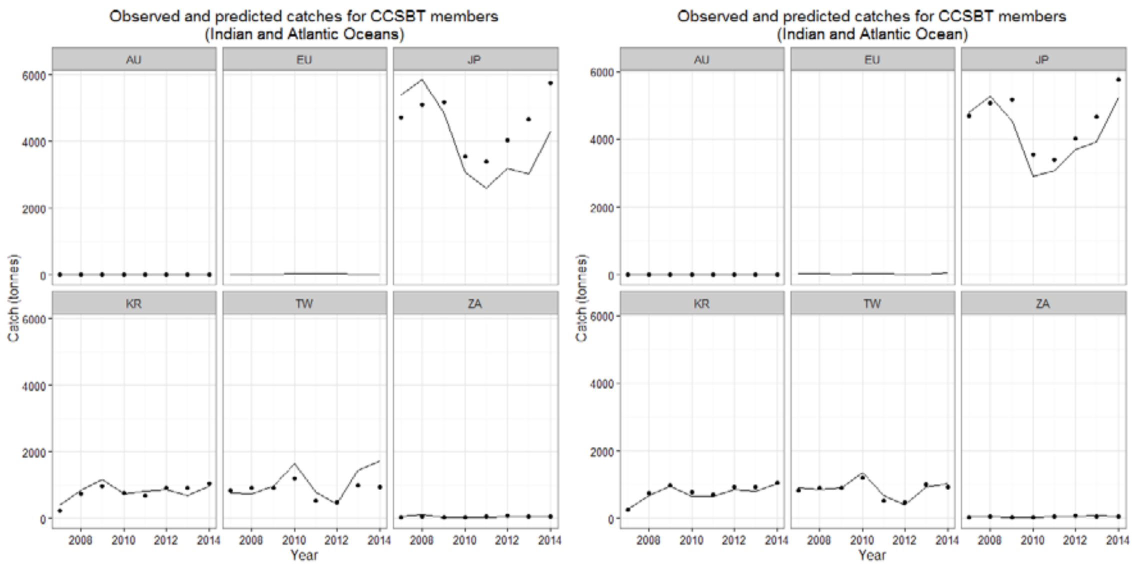


Figure 15. Comparison of observed (circles) and predicted (lines) catches in the Indian and Atlantic Oceans for CCSBT Members, from the GLM (left) and random forest (right) models, based on multiplying the predicted catch rate in kilograms per hook by adjusted effort in the CCSBT data. No observed catch data were available for the EU fleet, and the predictions shown assume TW catchability. Using the GLM, the average annual predicted EU catch was 18 tonnes and 66 tonnes assuming JP catchability). Using the random forest model, the average annual predicted EU catches was 26 tonnes (and 99 tonnes assuming a JP catchability).

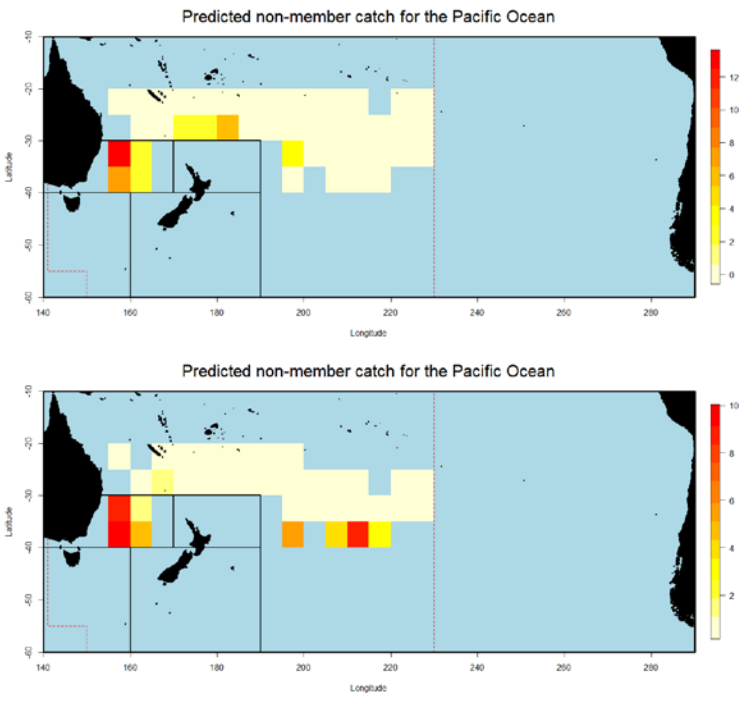


Figure 16. Map of average annual predicted non-Member catch in the Pacific Ocean, in tonnes, from the GLM (top) and random forest (bottom) analyses. Results average two alternate assumptions concerning catchability of the non-Member fleet (i.e. assumed Japanese or Taiwanese catchability) and use the adjusted CCSBT catch rate data used to parameterise the model.

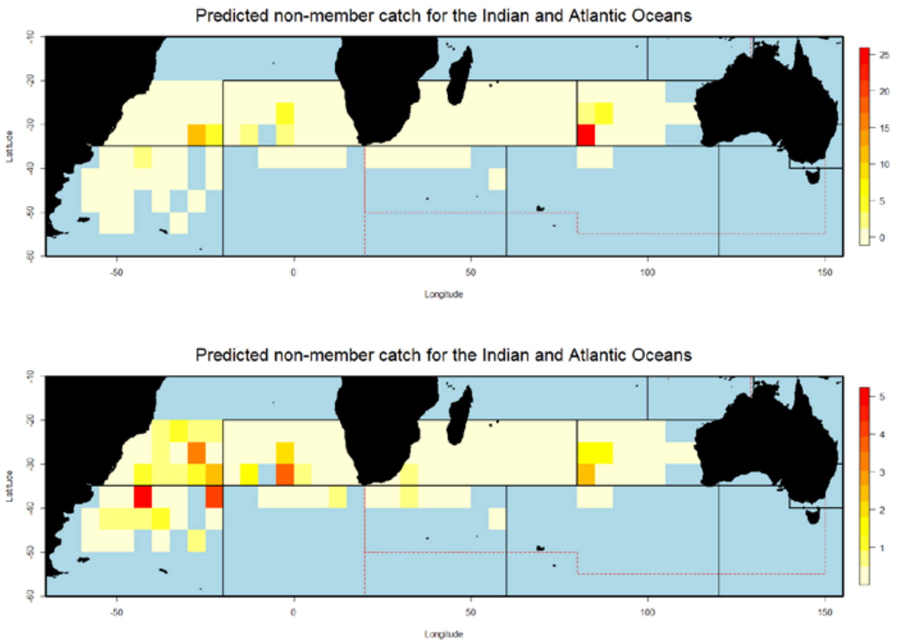
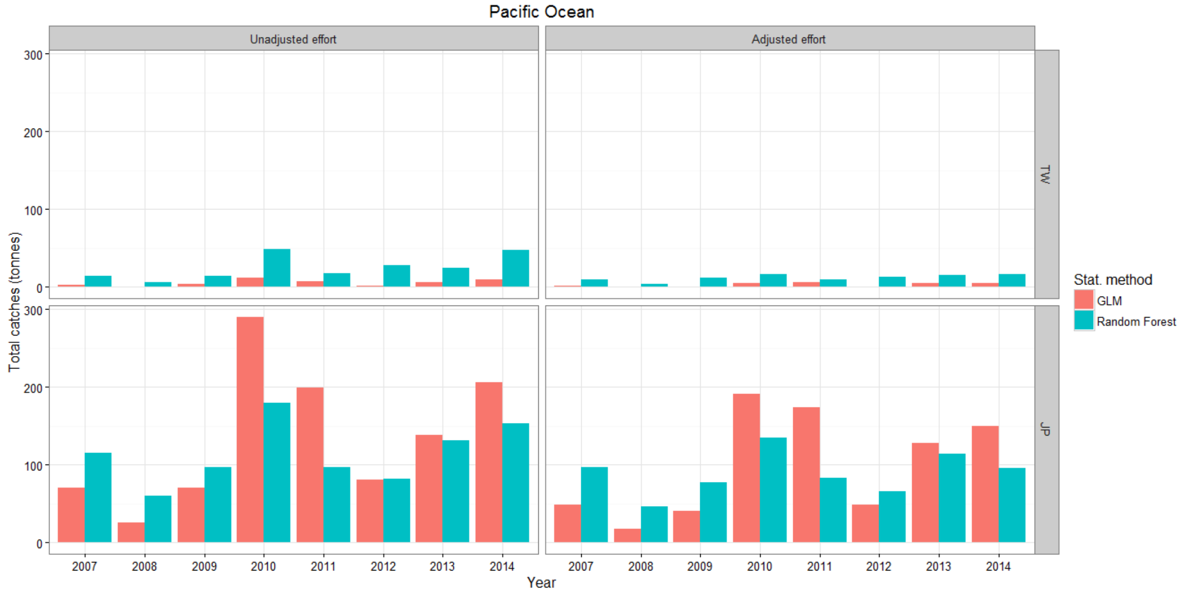


Figure 17. Map of average annual predicted non-Member catch in the Indian and Atlantic Oceans, in tonnes, from the GLM (top) and random forest (bottom) analyses. Results average two alternate assumptions concerning catchability of the non-Member fleet (i.e. assumed Japanese or Taiwanese catchability) and use the adjusted CCSBT catch rate data used to parameterise the model.



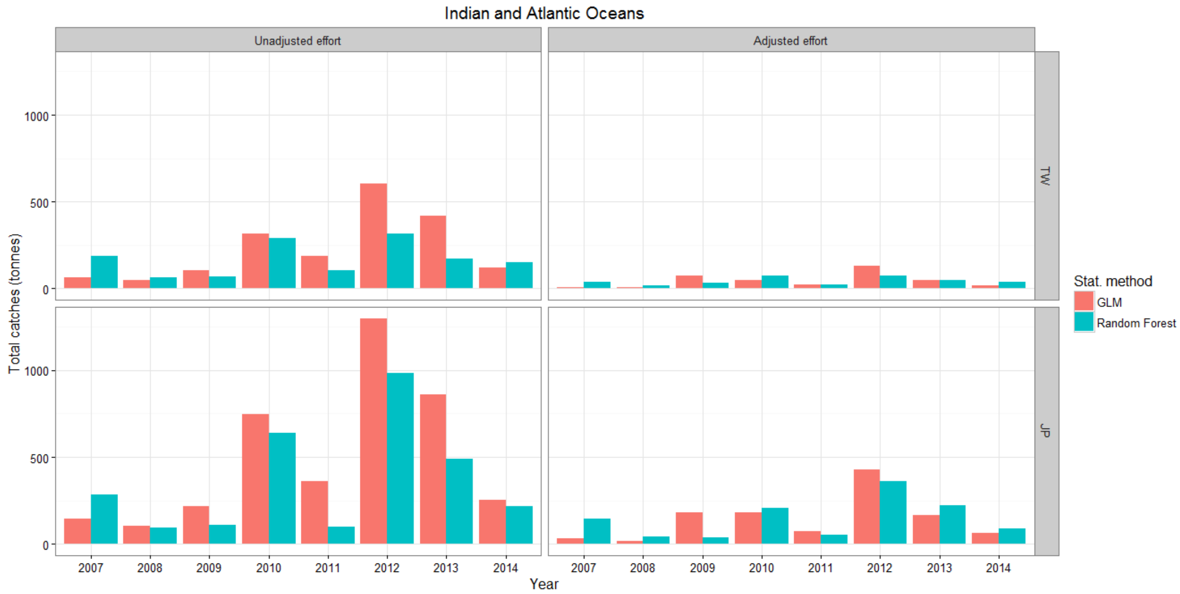


Figure 18. Total catches per year for each ocean, predicted by the GLM and random forest models. Outputs from the alternative flag assumptions are shown, in addition to the effect of using the adjusted or unadjusted CCSBT catch rate.

Tables

Table 1. Length to processed weight conversion factors agreed at the 1994 SBT Trilateral Workshop on Age and Growth, 17 Jan - 4 Feb, 1994. The parameters are used in the equation: $Weight = A.Length^B$; with *A* and *B* defined separately for adults and juveniles. Juveniles were defined as less than 130 cm, adults as greater than or equal to 130 cm.

Statistical area	Quarter	A_JUV	B_JUV	A_ADULT	B_ADULT
1	1	1.3545E-05	3.0214	7.3465E-06	3.157
2	1	1.3545E-05	3.0214	7.3465E-06	3.157
3	1	1.3545E-05	3.0214	5.5706E-06	3.2164
4	1	1.3545E-05	3.0214	5.5706E-06	3.2164
5	1	1.3545E-05	3.0214	8.3688E-06	3.1429
6	1	1.3545E-05	3.0214	8.3688E-06	3.1429
7	1	1.3545E-05	3.0214	5.5706E-06	3.2164
8	1	1.3545E-05	3.0214	3.9080E-07	3.7529
9	1	1.3545E-05	3.0214	5.1065E-06	3.2393
10	1	1.3545E-05	3.0214	5.1065E-06	3.2393
1	2	8.9030E-06	3.1225	1.8240E-07	3.9056
2	2	8.9030E-06	3.1225	1.8240E-07	3.9056
3	2	8.9030E-06	3.1225	5.5706E-06	3.2164
4	2	8.9030E-06	3.1225	5.5706E-06	3.2164
5	2	8.9030E-06	3.1225	2.9786E-06	3.3411
6	2	8.9030E-06	3.1225	7.3465E-06	3.157
7	2	8.9030E-06	3.1225	5.5706E-06	3.2164
8	2	8.9030E-06	3.1225	1.8240E-07	3.9056
9	2	8.9030E-06	3.1225	5.1065E-06	3.2393
10	2	8.9030E-06	3.1225	5.1065E-06	3.2393
1	3	1.5216E-05	3.0009	1.8240E-07	3.9056
2	3	1.5216E-05	3.0009	1.8240E-07	3.9056
3	3	1.5216E-05	3.0009	1.5380E-06	3.4754
4	3	1.5216E-05	3.0009	1.5380E-06	3.4754
5	3	1.5216E-05	3.0009	3.9490E-06	3.2886
6	3	1.5216E-05	3.0009	3.9490E-06	3.2886
7	3	1.5216E-05	3.0009	1.5380E-06	3.4754
8	3	1.5216E-05	3.0009	1.8240E-07	3.9056
9	3	1.5216E-05	3.0009	4.7780E-07	3.7032
10	3	1.5216E-05	3.0009	4.7780E-07	3.7032
1	4	1.3545E-05	3.0214	7.3465E-06	3.157
2	4	1.3545E-05	3.0214	7.3465E-06	3.157
3	4	1.3545E-05	3.0214	1.5380E-06	3.4754
4	4	1.3545E-05	3.0214	1.5380E-06	3.4754
5	4	1.3545E-05	3.0214	3.9490E-06	3.2886
6	4	1.3545E-05	3.0214	8.3688E-06	3.1429
7	4	1.3545E-05	3.0214	1.5380E-06	3.4754
8	4	1.3545E-05	3.0214	3.9080E-07	3.7529
9	4	1.3545E-05	3.0214	4.7780E-07	3.7032
10	4	1.3545E-05	3.0214	4.7780E-07	3.7032

Table 2. Predicted catches in tonnes by non-Member fleet for the GLM model for the Pacific, based on alternative assumptions that non-Member catchabilities match those of Taiwan (TW) or Japan (JP), and assuming either adjusted or unadjusted CCSBT effort. Blank cells indicate that no data were available. "Other" flags consist of summed values from CK, NC, PF and TO.

Note: Predicted catches by flag have been removed and are only available in the Members version of this report.

JP	Total
2007	70.11
2008	25.88
2009	70.74
2010	290.31
2011	198.42
2012	80.62
2013	138.29
2014	205.69

JP - adjusted	Total
2007	47.98
2008	17.05
2009	40.39
2010	190.74
2011	173.34
2012	48.45
2013	127.85
2014	149

TW	Total
2007	3.03
2008	0.73
2009	4
2010	12.25
2011	7.57
2012	2.01
2013	6.04
2014	9.98

TW - adjusted	Total
2007	1.47
2008	0.21
2009	0.86
2010	4.63
2011	6.24
2012	0.68
2013	4.79
2014	5.39

Table 3. Predicted catches in tonnes by non-Member fleet from GLM model for the combined Indian and Atlantic Oceans, based on alternative assumptions that non-Member catchabilities match those of Taiwan (TW) or Japan (JP), and assuming either adjusted or unadjusted CCSBT effort. Blank cells indicate that no data were available. "Other" flags consist of summed values from MU, MY, TH, BR, BZ, GH, TT, UY, VC and VU.

Note: Predicted catches by flag have been removed and are only available in the Members version of this report.

JP		Total
	2007	146.53
	2008	103.15
	2009	218.14
	2010	748.51
	2011	361.69
	2012	1300.67
	2013	859.70
	2014	254.47

JP - adjusted		Total
	2007	32.83
	2008	18.23
	2009	183.46
	2010	181.26
	2011	72.66
	2012	427.41
	2013	165.37
	2014	60.54

TW		Total
	2007	64.56
	2008	47.33
	2009	105.93
	2010	313.04
	2011	185.33
	2012	604.46
	2013	416.17
	2014	122.33

TW - adjusted		Total
	2007	8.25
	2008	5.01
	2009	74.30
	2010	48.54
	2011	21.31
	2012	130.57
	2013	48.92
	2014	17.03

Table 4. Predicted catches in tonnes by non-Member fleet from the random forest model for the Pacific, based on alternative assumptions that non-Member catchabilities match those of Taiwan (TW) or Japan (JP)), and assuming either adjusted or unadjusted CCSBT effort. "Other" flags consist of summed values from CK, NC, PF and TO.

Note: Predicted catches by flag have been removed and are only available in the Members version of this report.

JP	Total
2007	114.72
2008	60.28
2009	97.24
2010	179.79
2011	97.22
2012	81.66
2013	131.44
2014	153.24

JP - adjusted	Total
2007	96.18
2008	46.29
2009	76.97
2010	134.41
2011	83.31
2012	65.15
2013	113.41
2014	95.76

TW	Total
2007	14.29
2008	6.53
2009	14.23
2010	48.04
2011	17.82
2012	27.59
2013	24.75
2014	46.92

TW - adjusted	Total
2007	9.88
2008	3.84
2009	11.58
2010	16.39
2011	9.06
2012	13.13
2013	15.48
2014	16.04

Table 5. Predicted catches in tonnes by non-Member fleet from the random forest model for the Indian and Atlantic Oceans, based on alternative assumptions that non-Member catchabilities match those of Taiwan (TW) or Japan (JP), and assuming either adjusted or unadjusted CCSBT effort "Other" flags consist of summed values from MU, MY, TH, BR, BZ, GH, TT, UY, VC and VU.

Note: Predicted catches by flag have been removed and are only available in the Members version of this report.

JP		Total
	2007	285.32
	2008	92.04
	2009	108.27
	2010	638.82
	2011	100.7
	2012	982.56
	2013	489.36
	2014	215.19

JP - adjusted		Total
	2007	142.48
	2008	43.2
	2009	38.08
	2010	208.11
	2011	52.06
	2012	359.75
	2013	223.1
	2014	89.64

TW		Total
	2007	185.85
	2008	64.4
	2009	68.7
	2010	289.26
	2011	104.52
	2012	314.95
	2013	172.72
	2014	149.86

TW - adjusted		Total
	2007	35.8
	2008	16.88
	2009	31.76
	2010	74.54
	2011	24.1
	2012	75.19
	2013	45.85
	2014	35.75

Appendix: Additional diagnostic plots

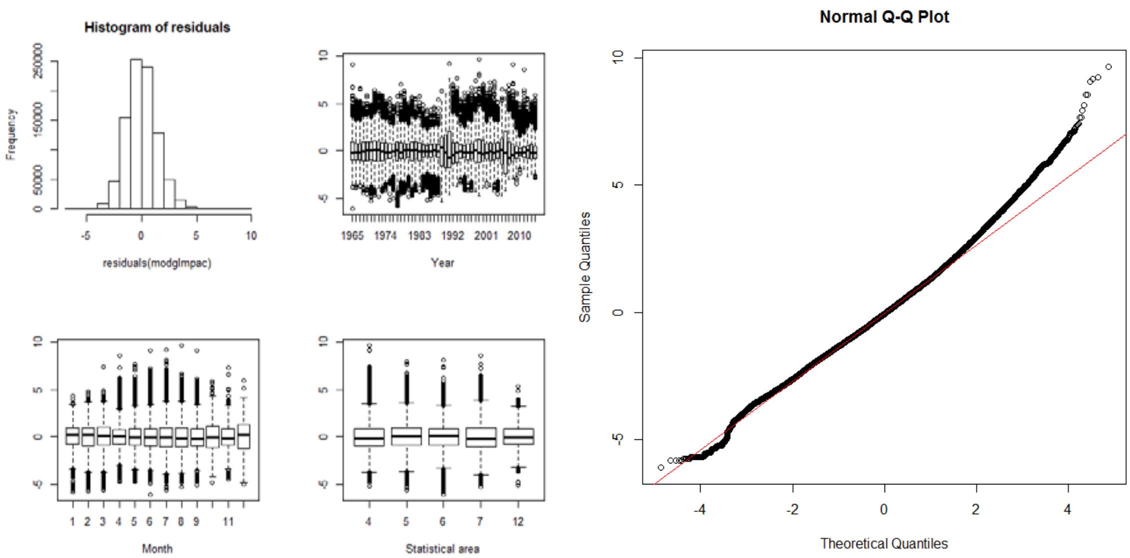


Figure 19. Histogram and boxplots of residuals from the analysis of size data for the Pacific Ocean.

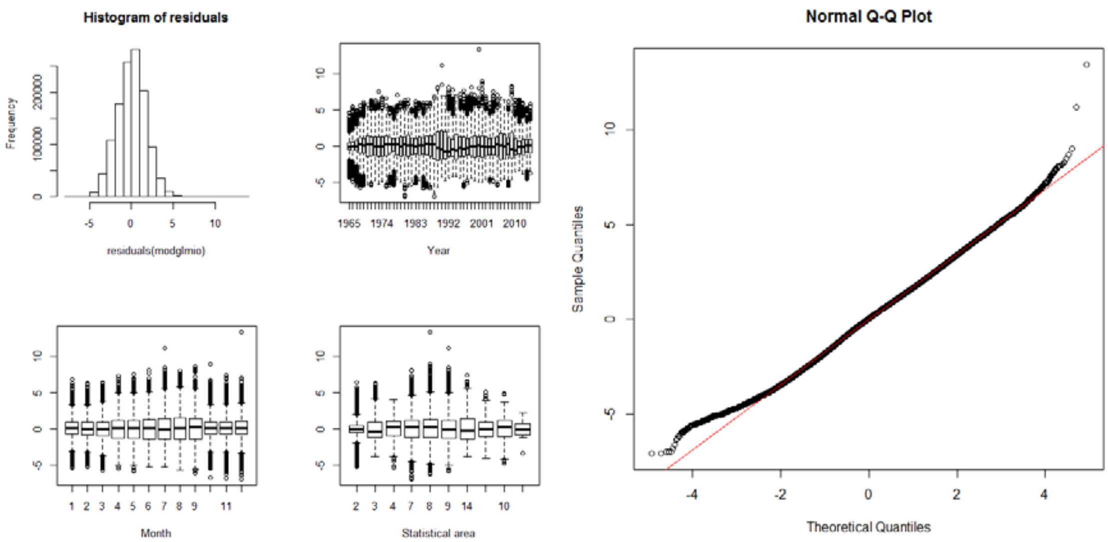


Figure 20. Histogram and boxplots of residuals from the analysis of size data for the Indian and Atlantic Oceans.

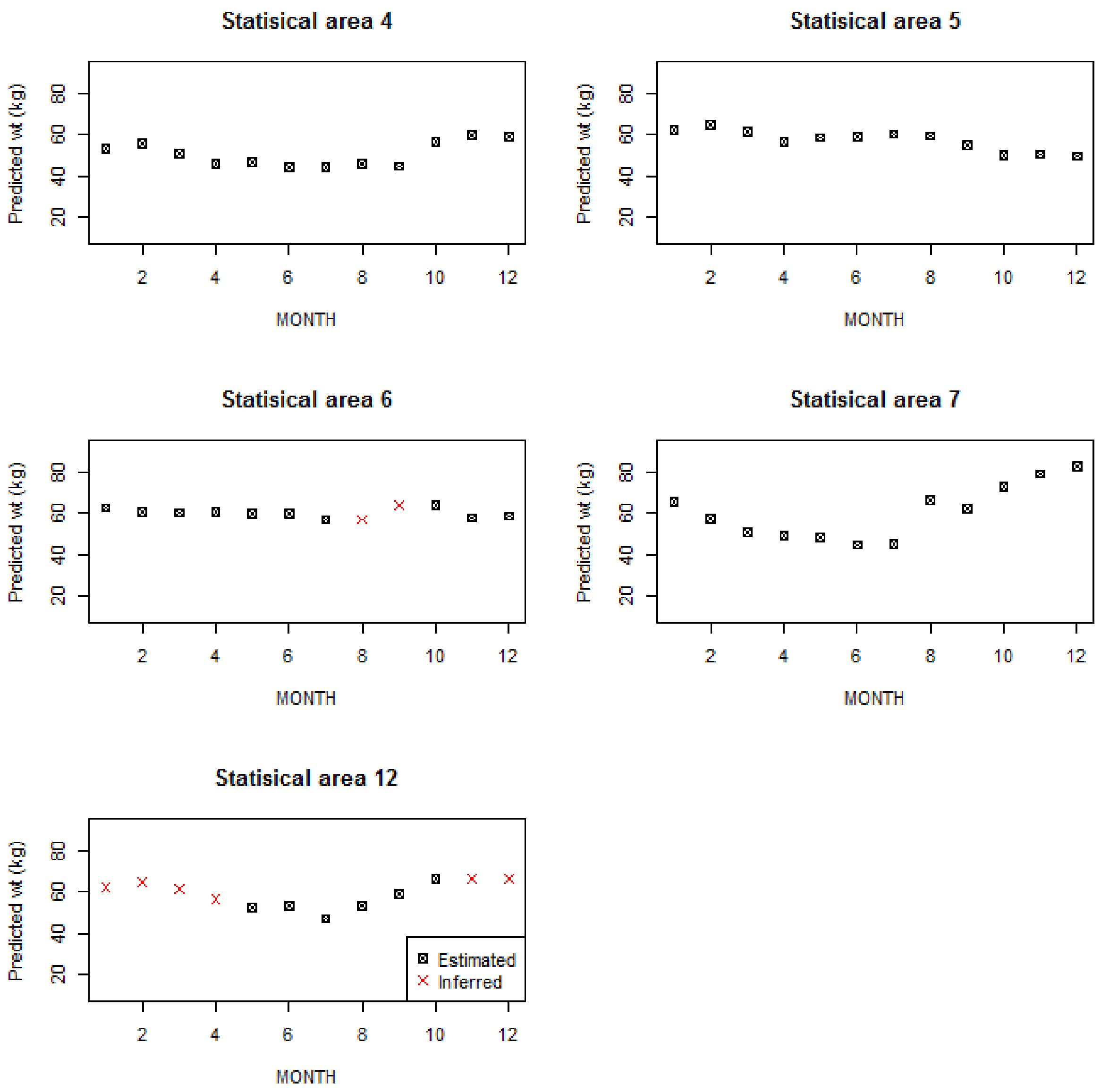


Figure 21. Predicted mean weights by month and statistical area for the Pacific Ocean. Estimated weights are plotted with black square, and inferred weights with red X's. In the month where both estimated and inferred values are plotted, the estimated value was considered unreliable and replaced with inferred values

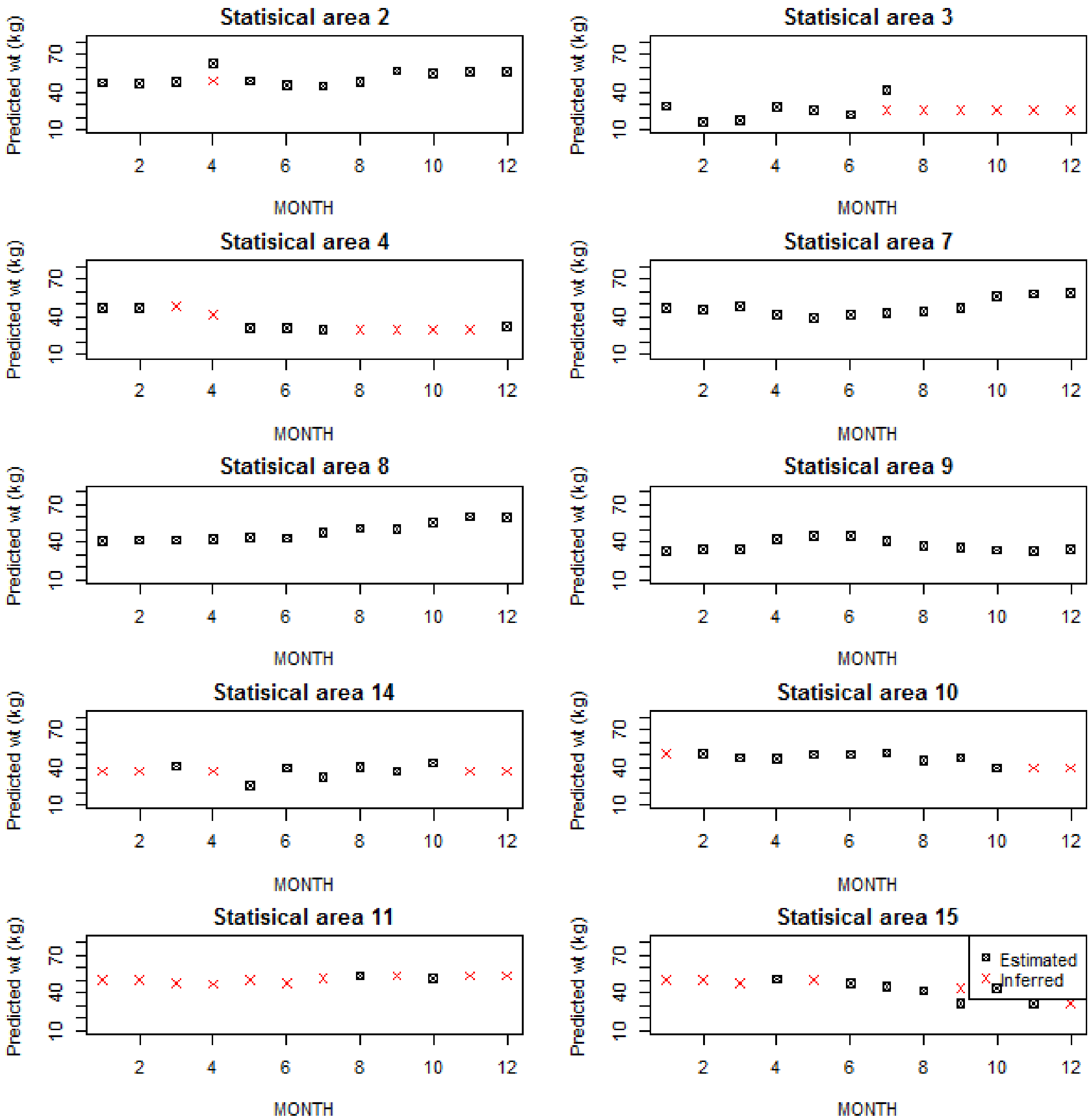


Figure 22. Predicted mean weights by month and statistical area for the Indian and Atlantic Oceans. Estimated weights are plotted with black square, and inferred weights with red X's. In the few months where both estimated and inferred values are plotted, the estimated values were considered unreliable and replaced with inferred values.

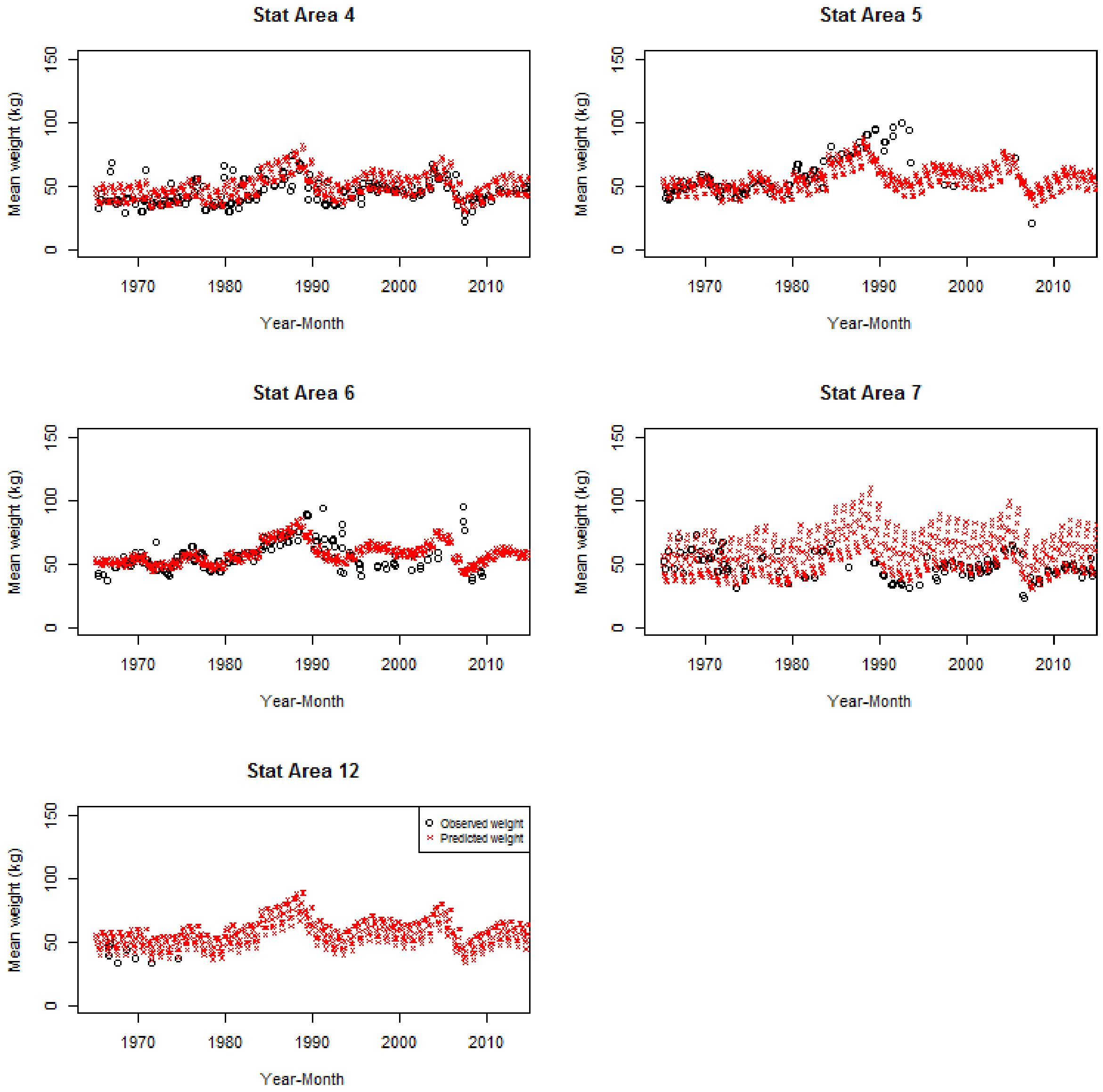


Figure 23. Observed (black circles) and predicted (red crosses) mean weights for Japanese sets in the Pacific Ocean.

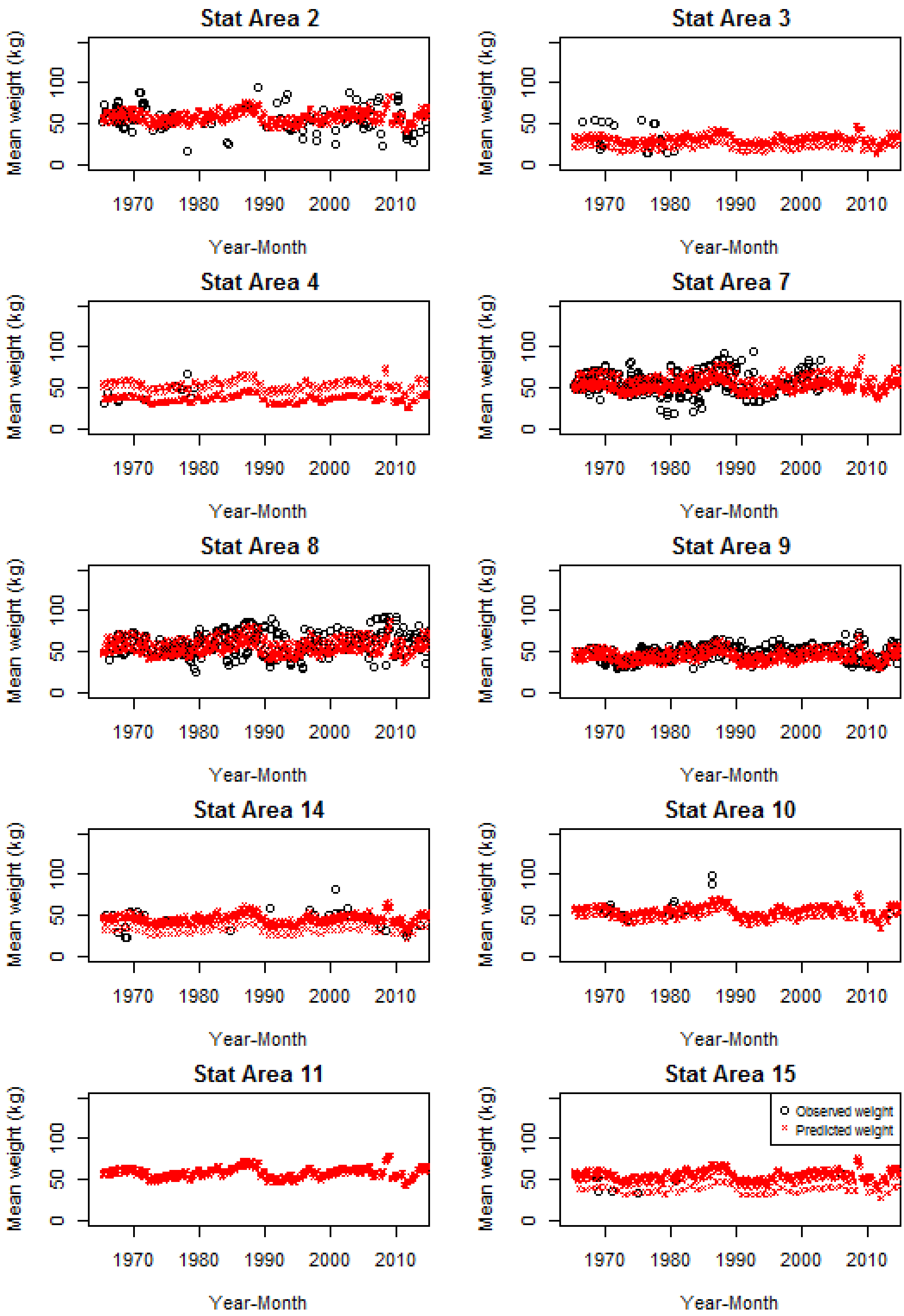


Figure 24. Observed (black circles) and predicted (red crosses) mean weights for Japanese sets in the Indian and Atlantic Oceans.

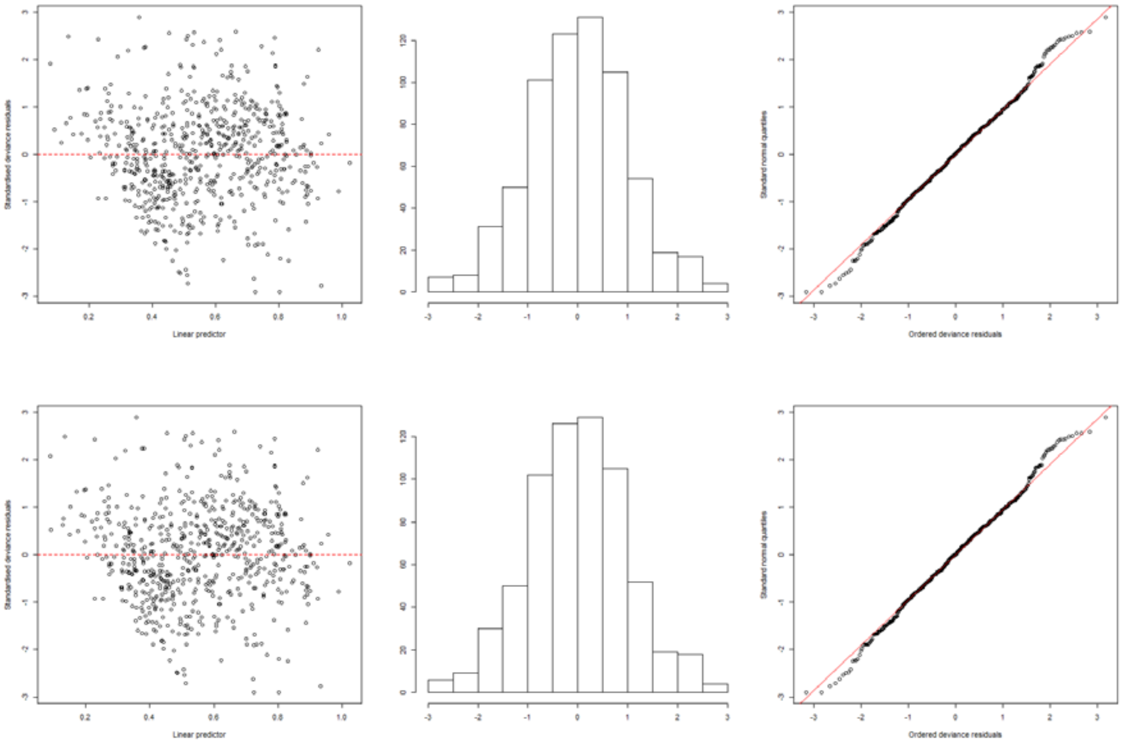


Figure 25. Residual diagnostics for the Pacific Ocean positive GLM, applied to the adjusted (above) and unadjusted (below) CCSBT effort data.

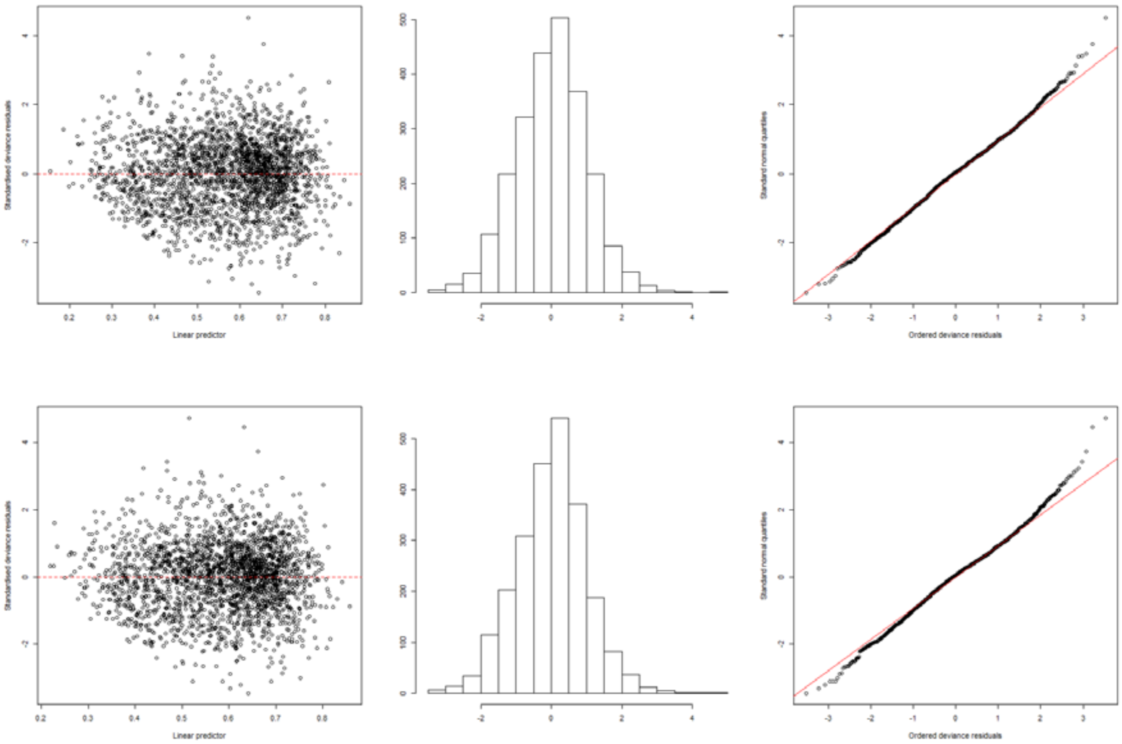


Figure 26. Residual diagnostics for the Indian and Atlantic Oceans positive GLM, applied to the adjusted (above) and unadjusted (below) CCSBT effort data.

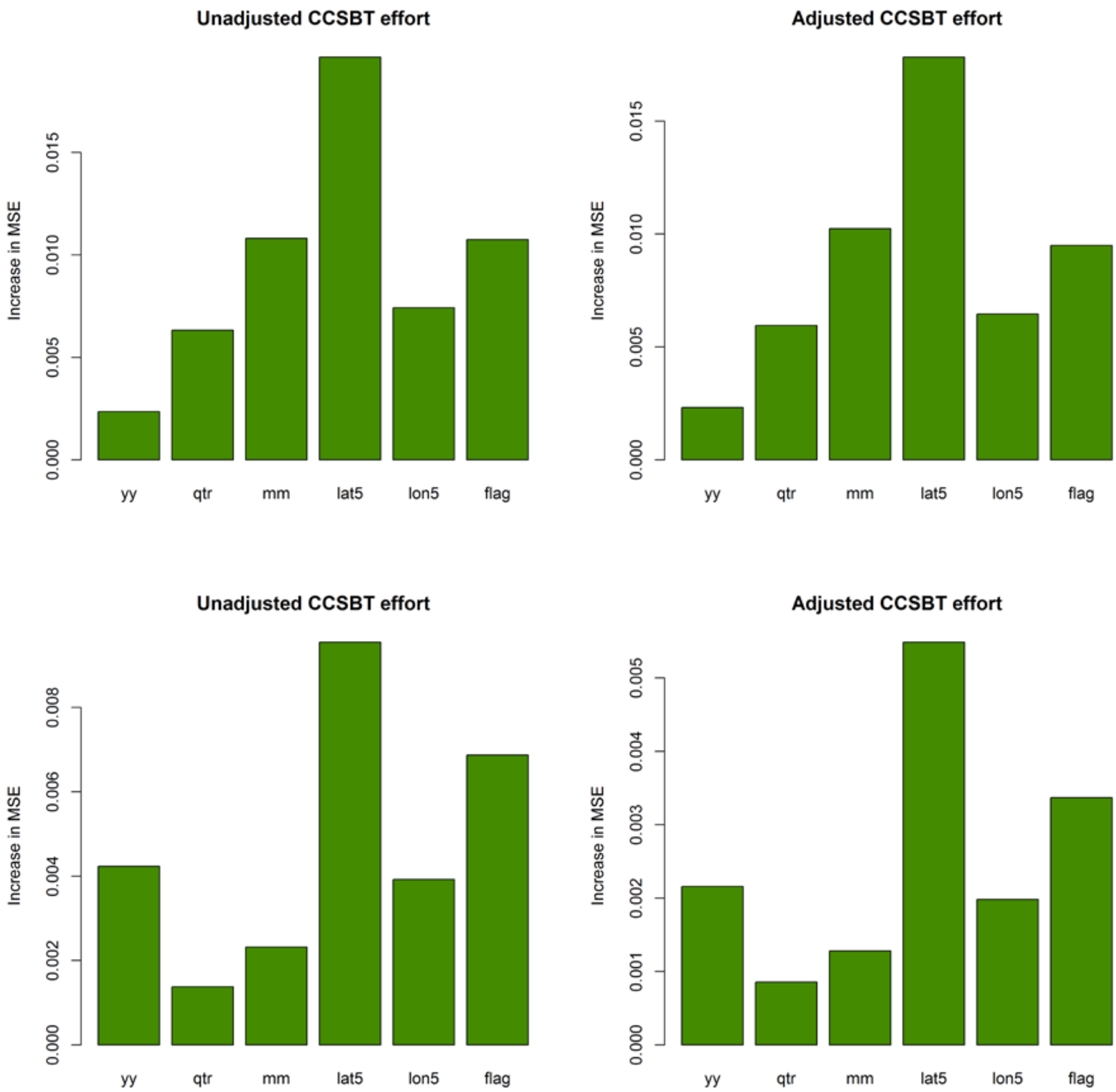


Figure 27. Variable importance for predicting catch rates of SBT using random forests for the Pacific (top) and Indian and Atlantic (bottom) Oceans and for adjusted and unadjusted CCSBT effort. MSE is the mean squared error.

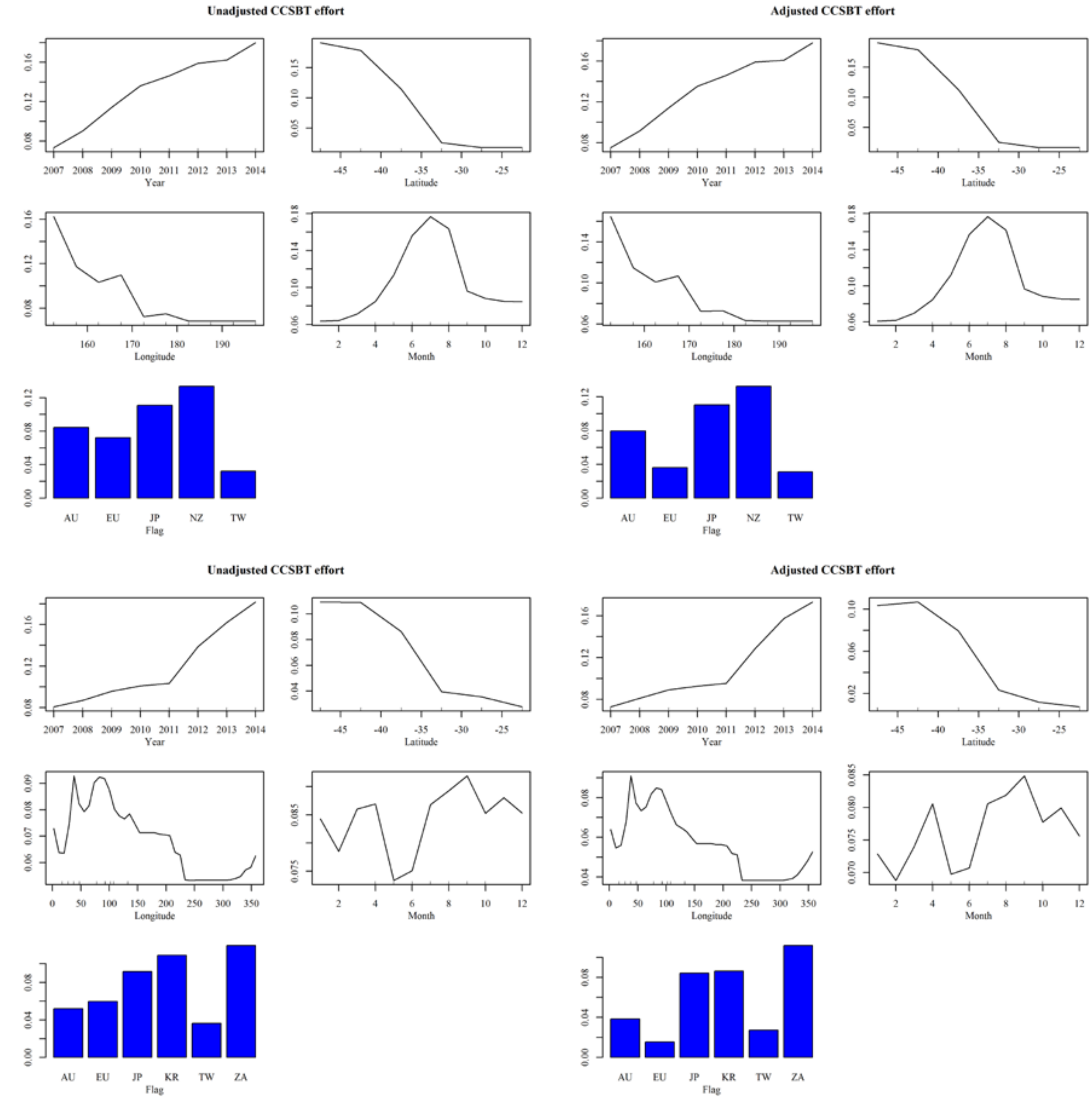


Figure 28. Partial effects of variables in the random forests model for predicting catch rates of SBT in the Pacific (top) and Indian and Atlantic (bottom) Oceans and for adjusted and unadjusted CCSBT effort.