SCIENTIFIC COMMITTEE
NINTH REGULAR SESSION

6-14 August 2013
Pohnpei, Federated States of Micronesia

Evaluation of Tag Mixing Assumptions for Skipjack, Yellowfin and Bigeye Tuna Stock Assessments in the Western Pacific and Indian Oceans

WCPFC-SC9-2013/ SA-IP-11

Dale Kolody[1] and Simon Hoyle[2]

[1] Wealth from Oceans National Research Flagship, CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania, 7001, Australia)
[2] (Secretariat of the Pacific Community Oceanic Fisheries Programme, BP D5, 98848, Noumea, New Caledonia)

Evaluation of Tag Mixing Assumptions for Skipjack, Yellowfin and Bigeye Tuna Stock Assessments in the Western Pacific and Indian Oceans

Dale Kolody (dale.kolody@csiro.au, Wealth from Oceans National Research Flagship, CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart, Tasmania, 7001, Australia)

Simon Hoyle (Secretariat of the Pacific Community Oceanic Fisheries Programme, BP D5, 98848, Noumea, New Caledonia)

Abstract

Tagging (or mark-recapture) studies generally require the assumption that tagged and untagged individuals (of a particular demographic group) are equally likely to be captured. For tuna, this usually translates into the assumption that tagged fish released from a relatively small region (selected for logistical reasons) mix rapidly over a much broader region of interest. In this paper, we apply the Comparison of Paired Recovery Distribution (CPRD) analysis to skipjack (*Katsuwonus pelamis*), yellowfin (*Thunnus albacares*) and bigeye tuna (*Thunnus obesus*) from large-scale tagging programmes in the Western Pacific and Indian Oceans. The CPRD analysis uses chi-square and nearest neighbour permutation statistical tests to calculate the probability that two tag recovery distributions from different release events are drawn from the same spatial distribution. The release events were separated in space and/or time and recovered in the same time window. If there is evidence that the two recovery distributions differ, it follows that tags from the two release events are not fully mixed with each other, such that at least one would also not be mixed with the untagged population either. Recovery events were defined with spatial boundaries corresponding to the most recent stock assessments, and analyses were restricted to release areas that were entirely within the recovery region. In the WCPO, there was strong evidence of incomplete mixing for 5 quarters following release for skipjack and yellowfin tunas and 1 quarter for bigeye tuna. For all 3 species, the observed periods of incomplete mixing is clearly a minimum, as there were insufficient observations to make inferences with respect to longer periods at liberty. In the Indian Ocean, there is strong evidence for incomplete mixing for skipjack for 3 quarters following release, 2 quarters for yellowfin and 1 quarter for bigeye. In contrast to the WCPO, useful numbers of CPRD events of longer duration were identified but did not show consistent compelling evidence of incomplete mixing. The difference between the Pacific and Indian Ocean results may reflect genuine characteristics of mixing rates (e.g. tuna in archipelagic waters seem to migrate shorter distances on average that tuna in oceanic waters). However, the failure to detect incomplete mixing in the Indian Ocean may also reflect methodological limitations arising from the opportunistic nature of tag observations (e.g. the spatial distribution of Indian Ocean tag releases and recoveries was more restricted than in the WCPO). We expect that the mixing problem is serious enough to potentially introduce large biases to at least some of these stock assessments, but the magnitude of the biases may not be easy to quantify.

Introduction

Tagging studies are potentially very valuable for making inferences about tuna populations, with tag-based estimators for abundance, natural mortality, fishing mortality and movement (e.g. Brownie et al. 1985) often integrated within stock assessment models (e.g. Maunder and Punt 2013). These estimators assume that tagged individuals are equally vulnerable to recapture as untagged individuals (for a particular demographic group which may be defined on the basis of size, age, sex, spatial region, etc.). This occurs if either tags or effort are distributed randomly with respect to the population of interest. However, tuna are usually tagged in logistically convenient but restricted areas, and the fishing fleet do not distribute their effort at random with respect to the fish population due to their operating constraints and efficiency motives. Analysts usually assume that released tags rapidly mix with the untagged population, such that tagged and untagged recapture probabilities should be nearly equal (at the relevant scale) after a predefined mixing period. However, the validity of this assumption is rarely explicitly examined in the context of the coarse spatial structure of most tuna stock assessments, and substantial biases may result from tag-based estimators as a consequence. For example, if (within a particular demographic group) tags remain resident near their point of release, and the fishery operates far from the release site, then tag-based estimators can be expected to under-estimate fishing mortality and over-estimate abundance of the population. Part of the impetus for this paper arises from recommendations of the independent review of Western and Central Pacific Ocean (WCPO) bigeye tuna stock assessment (Ianelli et al. 2012), in which tag mixing problems were suspected of having a strong influence on biomass estimates.

Kolody and Hoyle (in review) developed two methods for examining tag mixing assumptions, applied them to WCPO skipjack tuna, and used simulations to explore the magnitude of estimator biases that might be expected due to incomplete tag mixing. In this paper, we use one of these approaches, the Comparison of Paired Recovery Distributions (CPRD) analysis (similar to the method developed by Latour et al. 2001), to examine mixing assumptions for WCPO and Indian Ocean skipjack (*Katsuwonus pelamis*), yellowfin (*Thunnus albacares*) and bigeye tunas (*Thunnus obesus*). The CPRD analysis was selected here because it is easy to apply and has very simple data requirements (date, position and size of releases and recaptures). The alternative approach developed in Kolody and Hoyle (in review) has some additional advantages, but requires spatial catch data by size and fleet and must account for variability in tag reporting rates.

WCPO Tuna Tagging Programmes and Stock Assessment

Since large-scale tuna tagging began in the WCPO in the 1970s, approximately 500 000 skipjack, 160 000 yellowfin and 50 000 bigeye have been released (Kearney 1992, Kaltongga 1998, Caillot et al. 2012). The distributions of releases and recoveries are shown in Figure 1 (tagging data from Japanese programmes, mostly located near Japan, were not available and have not been included). Releases were primarily conventional plastic dart (spaghetti) tags inserted into juvenile tuna (fork length 30-60 cm) from dedicated pole and line vessels.

The most recent stock assessments for WCPO skipjack (Hoyle et al. 2011), yellowfin (Langley et al. 2011) and bigeye (Davies et al. 2011) were conducted using MULTIFAN-CL software (e.g. Hampton and Fournier 2001) with a 3 region spatial structure for skipjack and 6 region for yellowfin and bigeye tunas, as shown in Figure 2. Each spatial unit contains a sub-population that is assumed to be internally homogenous, and is linked to the other sub-populations via quarterly migration (implemented using instantaneous age-dependent bulk transfer coefficients). The skipjack assessment assumed a tag mixing period of 1 quarter (i.e. tags are assumed to be fully mixed within a spatial unit in the quarter immediately following release), while yellowfin and bigeye assumed 2 quarters. During the mixing period, fishing mortality estimates for tagged and untagged fish are calculated separately, and tag recoveries are excluded from the model objective function.

Indian Ocean Tuna Tagging Programmes and Stock Assessment

The majority of tuna tagging in the Indian Ocean was conducted from 2005-8 as part of the Regional Tuna Tagging Programme (e.g. Hallier 2008), with approximately 78 000 skipjack, 55 000 yellowfin and 35 000 bigeye tags released. Several thousand additional tags have been released as part of various small-scale tagging programmes, most notably in the Maldives (starting in the 1990s). The distributions of releases and recoveries are shown in Figure 6. Similar to the WCPO tagging programmes, releases were primarily conventional plastic dart tags inserted into juvenile tuna from pole and line vessels.

The most recent stock assessments that used the tagging data for Indian Ocean skipjack (Sharma et al. 2012), yellowfin (Langley et al. 2012) and bigeye (Kolody et al. 2010) were conducted using MULTIFAN-CL and Stock Synthesis (e.g. Methot 2000) software, while independent tagging analyses have been undertaken simultaneously (Dortel et al. 2012, Eveson et al. 2012). The yellowfin assessment uses a 5 region structure (Figure 7). Spatial structure for skipjack and bigeye tunas has been explored, but recent advice has favoured spatially-aggregated models, largely out of recognition of the absence of informative data with which to estimate movement. The skipjack assessment assumed a tag mixing period of 2 quarters, yellowfin 4 quarters (free-school sets only, as proposed by Langley and Million 2012) and bigeye 2 or 4 quarters (depending on the individual model).

Methods

The Comparison of Paired Recovery Distributions (CPRD) analysis developed in Kolody and Hoyle (in review) is very similar to the approach described previously in Latour et al. (2001). The CPRD analysis is a simple method of testing the null hypothesis that tagged fish are well-mixed. It requires only the dates and positions of tag releases and recoveries (by size class). CPRD involves the following steps:

1. Assign each recovered tag to a unique release and recovery event. For the purposes of this paper, we define a tag release event as all tag releases in a relatively small spatial area (e.g. small boxes in Figure 2, Figure 7) during a single quarter. Tag recovery events are defined as all tags from the same release event that are recaptured within the same quarter in a particular stock assessment region (e.g. large boxes in Figure 2, Figure 7).
2. Identify pairs of synchronous recovery events that correspond to different release events.
3. For each pair of recovery distributions from 2, calculate the probability that the observed recoveries represent samples from the same spatial distribution. If the distributions are

significantly different, this indicates that the release events are probably not fully mixed with one another, and it follows that at least one of the tag releases is not fully-mixed with the untagged population.

4. Repeat 3 for all CPRDs and summarize the results in relation to explanatory variables of interest.

The main difference between our approach and that developed by Latour et al. (2001) is that we include tag recovery comparisons from release events that are separated in space. Having multiple tag release sites is a useful luxury that might not be available from most tagging programmes. This makes the test of mixing more powerful since it can identify spatial differences that might not be evident for releases from a single location at different times (as shown in the logistic regression results below).

Conventional tag recovery distributions depend upon many factors, including the distribution of the tagged fish, and the distributions of fishing effort, tag reporting rates, and reporting errors in tag recovery dates and positions. Because of these complicating factors, in most tuna fisheries it is very unlikely that the tag recovery distribution would accurately represent the true distribution of all the tagged fish caught, or the tagged (or untagged) distribution still at liberty. However, in the CPRD analysis, both recovery distributions are from the same recovery time window, so all of the factors above that influence the observed distributions are equivalent. This means that it is meaningful to compare distributions despite these complicating factors. Nevertheless, each of these factors may increase the probability of Type II errors (i.e. failure to identify incomplete mixing when the two distributions are actually different).

There are a couple other potential complications to the analysis. Fishery selectivity can create a difference in tag recovery distributions because tagged fish of different sizes are not equally vulnerable. We attempted to minimize this effect by only comparing recovery distributions of similarly-sized fish (with recovery sizes estimated from the release length, time-at-liberty and mean growth curve estimates). Skipjack were separated into size-classes of <38, 38 to <58 and 58+ cm, yellowfin and bigeye <60, 60 to <100 and 100+ cm. Large and frequent tag recovery date errors would bias results if observed tag recovery events included recoveries from outside the desired time window. For example, if a substantial number of recoveries from the period before full-mixing had been achieved were misreported into a later period in which full-mixing had been achieved, this could cause a Type I error (i.e. inflated probability of finding significant differences between distributions even though the recovery events are actually samples from a single fully-mixed population). This is probably not a frequent problem, but could have serious implications for most tag analyses if it were. Additionally, the aggregation of tags (e.g. due to the effect of long-term fidelity of individuals to specific schools) is a form of incomplete mixing that the CPRD analysis will also identify. The aggregation process would not necessarily bias tag estimators (e.g. if the schools are well-mixed with each other), but would result in statistical over-dispersion relative to the ideal situation of truly random mixing.

Various test statistics could be used to compare the recovery distributions. We explored two simple options with simulation testing: i) chi-square contingency table on a two dimensional grid of binned data (similar to Latour et al. 2001), and ii) a Nearest-Neighbour Permutation (NNP) test (e.g. Bailey and Gatrell 1995). The R software package *spatstat* function *nncross* was used for the nearest neighbour calculations (available from R Core Team 2013). Kolody and Hoyle (in review) found that the results of the two tests often differed for individual CPRD events, but supported qualitatively very similar conclusions when applied across a large number of CPRD events. However, the chi-square test can be sensitive to the choice of bin definitions (5^o X 5^o bins were used here), and proved to be less powerful than the NNP test for smaller sample sizes in simple simulation tests (e.g. < ~30 tags per recovery distribution). The NNP test consisted of:

1. Calculate the mean NN for all tags from the first release group, where each individual NN represents the great circle distance between the first release group tag recovery position and the location of whichever tag recovery from the second release group is closest.

2. Generate a non-parametric null distribution for mean NN using Monte Carlo simulations:

    a. Pool the observed tag recovery positions from both release events, and then randomly reassign each tag to one of the two release events (maintaining the original sample sizes).

    b. Calculate mean NN from the randomized tag recovery distributions.

    c. Repeat steps 2a-b 1000 times, creating a mean NN frequency distribution.

3. If mean NN from 1 is near the tails of the null distribution from 2c, this is interpreted as evidence for incomplete mixing. We would generally expect the mean NN value for two distinct tag distributions to be at the right-hand tail of the null distribution (i.e. consistent with a large degree of spatial separation), and accordingly used a one-tailed test of significance ($P_{NNP} > 0.95$). However, for consistency of interpretation with the chi-square results, the mean NNP *P*-values are summarized as $P = 1 - P_{NNP}$, such that $P < \alpha = 0.05$ is identified as significant evidence of incomplete mixing.

Note that the result of the NNP test differs depending on which release event is defined as first and second. This is disconcerting if one is only interested in a small number of comparisons. We would expect very little difference to the overall result if a large number of comparisons are made. But there is potentially a systematic difference in how release events are assigned (e.g. shorter time-at-liberty and hence generally larger recovery numbers are always assigned to release group 1). As this is a new approach, we report the P-values for the chi-square test and the NNP test with the default event group labelling, and with the release event labelling reversed, to check for consistency.

Six example CPRD events are shown in Figure 3 (bigeye tuna releases in the WCPO) to illustrate the degree to which the results of the statistical tests tend to be consistent with each other and what one might conclude from visual inspection. Qualitatively, we would conclude that the tags in the top 2 panels are not mixed, and the 3 P-values from the chi-square and nearest neighbour permutation tests are in agreement. Similarly, we would tend to conclude that tags in the bottom 2 panels are probably mixed, and none of the P-values were significant. The middle two panels were selected as marginal cases. Visually, we would probably conclude that the tags in the right panel are mixed, while those in the left are not. But this view is the opposite of the conclusion of the first (default) NNP test value (which is the value emphasized in the results) and the chi-square test, but not the

NNP test with the region assignments reversed. In most cases, it seems like the statistical tests perform as we would intuitively expect (a particularly unusual example is highlighted in the discussion).

As a simple summary of (*n*) multiple CPRD results, we treated each individual *P*-value (P_i), from the first NNP test, as an independent sample from a Bernoulli process, and calculated the probability of obtaining the observed number (*k*) of statistically significant ($P_i \leq \alpha = 0.05$) results (or more) as the binomial probability:

$$P_B = \sum_{j=k}^{n} \binom{n}{k} \alpha^j (1-\alpha)^{n-j} .$$

The independence assumption might not always be completely valid in the sense that some recovery events appear in multiple CPRD observations. Similarly, recovery events from different size classes may arise from the same release event in a few cases.

We would expect that the probability of detecting incomplete mixing would vary in relation to the observed number of tag recoveries per release event, the distance separating release events and the duration since release. Toward this goal, we fit the following binomial logistic generalized linear model (R function *glm*, R Core Team 2013):

mixSig ~ *minMixTime* + *dist* + *minNTags* + *species*

Where:

> *mixSig* = binary response indicating whether the individual CPRD did or did not show significant (α = 0.05) evidence of incomplete tag mixing,
>
> *minMixTime* = (continuous variable) quarters between release and recovery time window (lesser of the two mixing periods if the two release events are asynchronous),
>
> *dist* = (continuous variable) great circle distance between the two release events (centres of the release areas shown in Figure 2),
>
> *minNTags* = (continuous variable) the lesser number of tag recoveries from the two recovery events, and
>
> *species* = categorical variable.

With additional development, this might provide a more objective basis for informing tagging model assumptions (i.e. spatial/temporal resolution and assumed mixing periods). However, in the current form, this is an oversimplification that ignores other important factors (e.g. the influence of the spatial pattern of the fishing fleet and reporting rates, the definition of recovery regions, and behavioural differences of the fish under different oceanographic conditions).

Results and Discussion

WCPO

While there are a very large number of tag releases and recoveries (Figure 1), only a few hundred pairs of events met the CPRD minimum sample size requirement. The frequencies with which particular release and recovery areas are represented are summarized by species in Table 1. Skipjack and yellowfin had a reasonable number of events, with the vast majority in the western equatorial region. Bigeye had only a small number of CPRDs, with more in the eastern than western equatorial region (due to the high proportion of bigeye released at the TAO buoys).

In the western WCPO, there was strong evidence of incomplete mixing for 5 quarters following release for skipjack and yellowfin tunas and 1 quarter for bigeye tuna (Table 2). In the eastern WCPO, incomplete mixing was evident for 1 quarter for skipjack and bigeye. For all 3 species, there were insufficient observations (3 or less) with which to make strong inferences about mixing over longer periods at liberty. The results of the chi-square and NNP tests were similar (when substantial numbers of CPRDs were observed). Note that the skipjack results are qualitatively the same, but slightly different from those described in Kolody and Hoyle (in review) because the release area boundaries were slightly changed (and there may have been additional minor changes to the tag database between extractions).

Figure 4 shows all of the CPRD events for WCPO bigeye tuna with at least 2 quarters of mixing for both release events (these are the only events from which all tags would have been included in the stock assessment). Only one of the five CPRDs showed significant evidence of incomplete mixing. All of the statistical tests appear to be consistent with what one would conclude from a visual inspection of the recovery locations.

The Analysis of Deviance table for the logistic regression model indicates that the probability of detecting incomplete mixing is strongly dependent on the mixing period and the distance between tag release points (Table 3). The number of tag recoveries and species were much less important, but still significant. The predicted probabilities are shown in relation to the main explanatory variables in Figure 5. The relative importance of the independent variables is consistent with expectations, and the approach seems like it might be useful for identifying the (minimum) spatial-temporal conditions that would satisfy tag mixing assumptions (e.g. answer the question - what combination of spatial resolution and mixing period would result in $P \sim 0.05$, the detection of random mixing purely by chance?). However, there are some caveats with the proposed model: e.g.

- important interactions are probably being ignored
- we might expect nonlinear responses to some of the independent variables
- the *dist* variable tells us something about tag mixing as a function of the separation distance of release events, but the more informative approach for deciding on an appropriate spatial structure would require repetition of the analysis with a range of alternative recovery area options. i.e. for a CPRD with release groups a fixed distance apart, a small recovery area should generally show less evidence of incomplete mixing than a large recovery area (and a smaller area will be more likely to conform to mixing assumptions).

- The model is still dependent on the one-sided nature of the CPRD analysis - absence of evidence for incomplete mixing is not the same as evidence that complete mixing has been achieved.

The CPRD analyses strongly suggest that tag mixing is not occurring at the rate assumed in the WCPO stock assessments for skipjack and yellowfin tunas. We are unable to recommend an appropriate mixing period on the basis of this analysis, except to note that it should be longer than 5 quarters in the western WCPO, if the current spatial structure is maintained. However, this is really suggestive of a fundamental mismatch between the assessment spatial structure and the tag dynamics. Reviewing the spatial structure would presumably be more fruitful than extending the mixing period. Adequate mixing might take considerably longer than 6 quarters (and conceivably never occurs, e.g. some fish might simply develop a sustained preference for familiar areas). Other analyses have suggested that bigeye and yellowfin tag displacements may be of a similar magnitude for releases from the same area, e.g. Hampton and Gunn (1998) compare tag recovery distributions from Coral Sea releases. If this is the case, then we could expect that bigeye mixing characteristics are similar to yellowfin in these areas. However, WCPO bigeye seem to have a larger proportion of the population in oceanic waters than the other two species, such that the average displacements and mixing rates for bigeye may exceed that of the other species.

Indian Ocean

The number of times that specific release and recovery areas are represented in the Indian Ocean CPRD analysis are summarized by species in Table 4. The majority of release events were from east Africa for all species. Large numbers of skipjack release events were also represented from Maldives and Seychelles.

In the Indian Ocean, there is strong evidence for incomplete mixing for skipjack for 3 quarters following release, 2 quarters for yellowfin, and 1 quarter for bigeye (Table 5). In contrast to the WCPO, moderate numbers of CPRD events of longer duration were identified with useful sample sizes, and did not show strong and consistent evidence of incomplete mixing. Thus the analysis does not provide direct evidence that tag mixing assumptions for recent Indian Ocean stock assessments are inappropriate.

However, we would caution that this represents an optimistic interpretation because of the limitations of the CPRD analysis. A different analysis, based on the qualitative comparisons of catch and tag distributions (Langley and Million 2012), concluded that 4 quarters of mixing were required for Indian Ocean yellowfin purse seine free-school sets, while adequate mixing was not observed for FAD sets. The most notable limitation for both these sorts of analyses arises from the opportunistic nature of fisheries tag recoveries. There is often qualitative evidence to suggest that tuna tend to be less mobile in archipelagic waters than in the open ocean. In the Pacific Ocean, the purse seine fleets, which have reasonable numbers of returned tags (and reporting rate estimates from tag seeding experiments), tend to operate under more diverse conditions than in the Indian Ocean. The Pacific fleet often operates near the tag release site and in a mix of archipelagic and open ocean waters. In contrast, the Indian Ocean purse seine fleet operates mostly in open ocean, far from the main tag release point near the coast of east Africa. We can make some further qualitative observations by examining individual CPRD events (Figure 8):

- A and B suggest that there is some short-term residency around the east Africa release site, but it is not necessarily persistent. Some fish clearly do disperse rapidly as well.
- C and D suggest that east African releases can and do mix from the African coast through to the Maldives archipelago.
- However, E and F indicate that releases from the Maldives at least sometimes are retained near the Maldives (even after 4 quarters), and do not necessarily mix substantially into the western PS region.

The increased importance of open ocean habitat in the Indian Ocean may increase the effective mixing rate of tags relative to the Pacific. However, the pattern of releases and operational characteristics of the Indian Ocean fleet makes it more likely to decrease the sensitivity of the analysis. In the Indian Ocean, the CPRD analysis cannot be expected to make inferences about tag mixing rates outside of the western equatorial purse seine region, except for a limited number of observations from the Maldives. Tag recoveries from many coastal regions and the eastern equatorial region are not well represented in this analysis.

Note that Figure 8E provides a good illustration of the counter-intuitive discrepancy among P-values that can occasionally arise from the statistical tests. The NNP P-values are very different depending on the release group designations. This happens because the recoveries from the two release groups are close together and overlap in the Maldives region, but only group 1 is represented in the distant western purse seine region, where tags are broadly separated. In the NNP test, the mean NN value for group 1 is very large relative to the null distribution, and the mean NN for group 2 is very small. Both results would be identified as significant if two-tailed tests were employed in this case (note that for the NNP tests, reported P-Value = 1-actual P-Value as discussed in the methods). Also curious is the lack of significance of the chi-square test. Presumably this simply relates to the small sample size of group 2, and the fact that the arbitrary 5 degree binning option results in a large number of bins with 0 or 1 observation. The chi-square test does not consider the spatial relationship among the bins, so the higher density of non-zero bins in the Maldives is irrelevant. Overall, the statistical tests seem to be generally consistent with each other and intuition, but individual results may be deceptive. Perhaps there are more appropriate tests, but we found no reason to doubt the broad conclusions of the analysis.

Conclusions

This study demonstrates that tag mixing assumptions are not being met in recent WCPO stock assessments for skipjack and yellowfin in the western Pacific Ocean. The CPRD results did not identify a similar problem for bigeye, though this may be due to the limited number of observations, rather than substantive behavioural differences among species. The analysis did not identify incompatibility between mixing rates and Indian Ocean assessment assumptions. However, the analysis can never prove that adequate mixing has occurred, and the character of the Indian Ocean observations is such that the CPRD analysis is less powerful than in the Pacific.

The CPRD analysis does not quantify the magnitude of the estimator biases that are likely to result from incomplete tag mixing. The magnitude and direction of biases depend on the distributions of the tagged and untagged populations, and the fleet distribution, in a manner that is not easy to quantify (e.g. it is particularly difficult to make inferences about fish density in regions which have no fishery, poor data, or unknown tag reporting rates). Simulations loosely guided by WCPO skipjack dynamics (Kolody and Hoyle, in review) illustrated that biases caused by incomplete tag mixing could be substantial, but are not necessarily inevitable. The tag mixing dynamics could be contributing to the large assessment sensitivities identified in the WCPO bigeye assessment review (Ianelli et al. 2012), however, there could be other substantial contributing factors as well.

The CPRD analysis, and the related approach discussed in Kolody and Hoyle (in review) could be used to define a spatial structure that is more compatible with mixing assumptions. This would simply involve comparing different (increasingly higher resolution) spatial structures until the evidence for incomplete mixing disappeared within a satisfactory mixing period. This would be a useful analysis, but there are limits to this approach (in addition to the fact that full mixing cannot be proved), including:

- Smaller regions result in fewer tag recoveries, fewer CPRD events and less power (and probably an increasing number of regions with no valid observations)
- The recommended spatial structure may correspond to an assessment model that is exceedingly over-parameterized (e.g. the estimation of movement from tagging data with an unbalanced release design is problematic).

We recommend that additional work involving high resolution simulations should be undertaken to help understand the implications of tag mixing for current stock assessments. This should provide an indication of the implications of scale mismatches between natural processes and model structural limitations. This should lead to a more realistic understanding of the biases and uncertainties in these assessments, and might suggest some obvious paths for improvement if this proves necessary.

Acknowledgements

This work was funded by the Secretariat of the Pacific Community Oceanic Fisheries Programme (New Caledonia) and Commonwealth Scientific and Industrial Research Organization Wealth from Oceans National Flagship (Australia). John Hampton, Paige Eveson and Simon Nicol provided helpful comments on the manuscript. We are indebted to the countless people who have collected and managed the tagging data over the last several decades, and the visionary funding agencies that made it all possible.

References

Brownie, C., Anderson, D.R., Burnham, K.P., Robson, D.S., 1985. Statistical inference from band recovery data: a handbook. U.S. Fish and Wildlife Resource Publication 156.

Caillot, S., Leroy, B., Sanchez, C., Kolody, D. Nicol, S. Hampton, J. Lewis, A., Usu,T.,Kumasi, B. Kumoru, L., 2012. Pacific Tuna Tagging and PNG Tagging Project Progress Report and Workplan for 2012-2013. Western and Central Pacific Fisheries Commission submitted paper WCPFC-SC8-2012/RP-PTTP-02.

Davies, N., Hoyle, S., Harley, S., Langley, A., Kleiber, P., Hampton, J. 2011. Stock assessment of bigeye tuna in the western and central Pacific Ocean. Western and Central Pacific Fisheries Commission Scientific Committee working paper WCPFC-SC7-2011/SA-WP-02.

Dortel, E., Sardenne, F., Le Croizier, G., Million, J., Hallier, J.P., Morize, E., Munaron J.M., Bousquet, N., Chassot, E., 2012. A hierarchical Bayesian integrated model incorporated direct ageing, mark-recapture and length-frequency data for yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) of the Indian Ocean. Indian Ocean Tuna Commission 14th working party on tropical tunas working paper IOTC-2010-WPTT14-24(Rev1).

Eveson, J.P., Million, J., Herrera, M., 2012. Application of the Brownie-Petersen method for estimating mortality rates and abundance to Indian Ocean yellowfin tuna tag-recapture and catch data. Indian Ocean Tuna Commission 14th working party on tropical tunas working paper IOTC-2010-WPTT14-32 (Rev1).

Hallier, J.P. 2008. Status of the Indian Ocean Tuna Tagging Programme – RTTP-IO. Indian Ocean Tuna Commission Working Party on Tag Data Analysis Working Paper IOTC-2008 WPTDA-10.

Hampton, J., Fournier, D.A., 2001. A spatially-disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. Mar. Freshwater Res. 52, 937-963.

Hampton, J., Gunn, J. 1998. Exploitation and movements of yellowfin tuna (*Thunnus albacares*) and bigeye tuna (*T. obesus*) tagged in the north-western Coral Sea. Mar. Freshwater Res. 49, 475-89.

Hoyle, S., Kleiber, P., Davies, N., Langley, A. , and Hampton, J., 2011. Stock assessment of skipjack tuna in the western and central Pacific Ocean. Western and Central Pacific Fisheries Commission Scientific Committee working paper WCPFC-SC7-2011/SA-WP-04.

Ianelli, J., Maunder, M. Punt, A.E., 2012. Independent review of 2011 WCPO bigeye tuna assessment. Western and Central Pacific Fisheries Commission Scientific Committee working paper WCPFC-SC8-2012/SA-WP-01.

Kaltongga, B., 1998, Regional tuna tagging project : data summary. Noumea, New Caledonia: Secretariat of the Pacific Community. vi, 70 pp. (http://www.spc.int/coastfish/en/publications/digital-library.html)

Kearney, R.E., 1982, Methods used by the South Pacific Commission for the survey and assessment of skipjack and baitfish resources. Noumea, New Caledonia: South Pacific Commission. South Pacific Commission. 122 pp. (http://www.spc.int/coastfish/en/publications/digital-library.html)

Kolody, D., Hoyle, S. In review. Evaluation of tag mixing assumptions in western Pacific Ocean skipjack tuna stocks and assessment models. Submitted to Fish. Res. XXX

Kolody, D., Herrera, M., Million, J., 2010. Exploration of Indian Ocean Bigeye Tuna Stock Assessment Sensitivities 1952-2008 using Stock Synthesis (Updated to include 2009). Indian Ocean Tuna Commission 12th working party on tropical tunas working paper IOTC-2010-WPTT12-04(Rev1).

Latour, R.J., Hoenig, J.M., Olney, J.E., Pollack, K.H., 2001. A simple test for nonmixing in multiyear tagging studies: application to striped bass tagged in the Rappahhannock River, Virginia. Trans. Am. Fish. Soc. 130: 848-856.

Langley, A., Hoyle, S., Hampton, J, 2011. Stock assessment of yellowfin tuna in the western and central Pacific Ocean. Western and Central Pacific Fisheries Commission Scientific Committee working paper WCPFC-SC7-2011/SA-WP-03 (Rev.1).

Langley, A., Herrera, M., Million, J., 2012. Stock assessment of yellowfin tuna in the Indian Ocean using MULTIFAN-CL. Indian Ocean Tuna Commission 14th working party on tropical tunas working paper IOTC-2012-WPTT14-38 (Rev1).

Langley, A., Million, J., 2012. Determining an appropriate tag mixing period for the Indian Ocean yellowfin tuna stock assessment. Indian Ocean Tuna Commission 14th Working Party on Tropical Tuna Working paper IOTC-2012-WPTT14-31.

Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fish. Res. 142, 61-74.

Methot, R.D., 2000. Technical description of the stock synthesis assessment program. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-43, 46 p.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Sharma, R., Herrera, M., Million, J., 2012. Indian Ocean skipjack tuna stock assessment (stock synthesis). Indian Ocean Tuna Commission 14th working party on tropical tunas working paper IOTC-2012-WPTT14-29 (Rev1).

Table 1. The frequency with which different release areas are represented in the WCPO CPRD analyses. The same release event may be included more than once if it qualified for multiple comparisons. The same release area may be included either once or twice in the same CPRD, depending if the release events are separated in space or time.

Release Area (within recapture area)	n
Skipjack	
PH (SKJ-SA2) - Philippines	53
ID (SKJ-SA2) – Eastern Indonesia	184
BS (SKJ-SA2) – Bismarck Sea	228
SB (SKJ-SA2) – Solomon Islands	398
CS (SKJ-SA2) – Coral Sea	51
KI (SKJ-SA3) – Kiribati/Marshall Islands	26
t190 (SKJ-SA3) – Tau buoys	7
t205 (SKJ-SA3) – Tau buoys	0
FJ (SKJ-SA3) - Fiji	13
Yellowfin	
PH (SKJ-SA3) - Philippines	3
ID (SKJ-SA3) – Eastern Indonesia	52
BS (SKJ-SA3) – Bismarck Sea	126
SB (SKJ-SA3) – Solomon Islands	189
CS (Y/B-SA5) – Coral Sea	0
KI (SKJ-SA4) – Kiribati/Marshall Islands	0
t190 (SKJ-SA4) – Tau buoys	2
t205 (SKJ-SA4) – Tau buoys	0
FJ (SKJ-SA6) - Fiji	0
Bigeye	
PH (SKJ-SA2) - Philippines	0
ID (SKJ-SA2) – Eastern Indonesia	4
BS (SKJ-SA2) – Bismarck Sea	9
SB (SKJ-SA2) – Solomon Islands	3
CS (SKJ-SA2) – Coral Sea	0
KI (SKJ-SA3) – Kiribati/Marshall Islands	4
t190 (SKJ-SA3) – Tau buoys	14
t205 (SKJ-SA3) – Tau buoys	12
FJ (SKJ-SA3) - Fiji	0

Table 2. CPRD results for Western and Central Pacific Ocean SKJ, YFT and BET by recovery region.

Tag time-at-liberty (quarters)	N pairs of recovery events	Proportion CPRDs with incomplete mixing			Binomial *P*-Value NNP1
		Chi-square	NNP1	NNP2	
Skipjack tag recovery area = SKJ SA 2					
0	134	0.78	0.82	0.72	<0.001
1	135	0.74	0.64	0.58	<0.001
2	91	0.65	0.57	0.60	<0.001
3	59	0.54	0.49	0.49	<0.001
4	19	0.58	0.47	0.68	<0.001
5	15	0.47	0.60	0.67	<0.001
6	3	0.33	0.33	0.33	0.14
7	1	0.00	0.00	0.00	1.00
Skipjack Tag recovery area = SKJ SA 3					
0	11	0.82	0.73	0.82	<0.001
1	1	0.71	0.86	0.71	<0.001
2	2	0.50	0.25	0.50	0.19
3	3	0.00	0.00	0.00	1.00
Yellowfin tag recovery area = YFT/BET SA 3					
0	55	0.73	0.75	0.76	<0.001
1	54	0.61	0.46	0.57	<0.001
2	25	0.60	0.40	0.68	<0.001
3	18	0.44	0.33	0.50	<0.001
4	16	0.38	0.44	0.38	<0.001
5	10	0.30	0.50	0.40	<0.001
6	2	0.00	0.00	0.00	1.00
10	2	0.50	0.00	0.50	1.00
11	3	0.33	0.00	0.67	1.00
Yellowfin tag recovery area = YFT/BET SA 4					
0	1	0.00	0.00	0.00	1.00
Bigeye tag recovery area = YFT/BET SA 3					
0	3	0.68	0.67	0.33	0.007
1	4	0.75	0.75	1.00	<0.001
2	1	0.00	0.00	0.00	1.00
Bigeye tag recovery area = YFT/BET SA 4					
0	8	0.75	0.50	0.63	<0.001
1	3	0.33	0.00	0.33	1.00
2	3	0.33	0.33	0.33	0.14
3	1	0	0.00	0.00	1.00

Table 3. R Analysis of Deviance output from the WCPO logistic regression model.

```
Call:
glm(formula = (mixSig ~
    as.numeric(minMixTime) +
    as.numeric(dist) +
    as.numeric(minNTags) +
    as.factor(species), family = binomial(logit))

Coefficients:                Estimate  Std. Error  z value  Pr(>|z|)
(Intercept)                 -1.3494454  0.5109946  -2.641  0.00827 **
minMixTime                  -0.4481859  0.0668051  -6.709  1.96e-11 ***
dist                         0.0009038  0.0000934   9.677  < 2e-16 ***
minNTags                     0.0062967  0.0031648   1.990  0.04664 *
species=Skipjack             1.2659356  0.4991262   2.536  0.01120 *
species=Yellowfin            1.0319595  0.5169355   1.996  0.04590 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Table 4. The frequency with which different release areas are represented in the Indian Ocean CPRD analyses. The same release event may be included more than once if it qualified for multiple comparisons. The same release area may be included either once or twice in the same CPRD, depending if the release events are separated in space of time.

Release Area (within recapture area)	n
Skipjack	
OMN (whole IO) - Oman	0
TNK (whole IO) – East Africa	283
SEY (whole IO) - Seychelles	120
MLD (whole IO) - Maldives	105
MAD (whole IO) - Madagascar	18
Yellowfin	
OMN (SA1) - Oman	0
TNK (SA2) - East Africa	212
SEY (SA2) – Seychelles	8
MLD (SA2) - Maldives	20
MAD (SA3) - Madagascar	0
Bigeye	
OMN (whole IO) - Oman	0
TNK (whole IO) – East Africa	135
SEY (whole IO) - Seychelles	0
MLD (whole IO) - Maldives	3
MAD (whole IO) - Madagascar	0

Table 5. CPRD results for Indian Ocean tuna by recovery region and species.

Tag time-at-liberty (quarters)	N pairs of recovery events	Proportion CPRDs with incomplete mixing			Binomial *P*-Value NNP1
		Chi-square	NNP1	NNP2	
Skipjack tag recovery area = whole Indian Ocean					
0	76	0.76	0.83	0.48	<0.001
1	62	0.48	0.43	0.27	<0.001
2	40	0.18	0.18	0.20	0.003
3	33	0.18	0.30	0.06	<0.001
4	24	0.13	0.08	0.08	0.34
5	16	0.00	0.06	0.06	0.56
6	8	0.00	0.13	0.00	0.34
7	3	0.00	0.00	0.33	1.0
8	1	0.00	0.00	0.00	1.0
Yellowfin tag recovery area = SA2					
0	30	0.70	0.83	0.60	<0.001
1	14	0.36	0.29	0.14	0.004
2	15	0.20	0.27	0.07	0.005
3	16	0.00	0.00	0.13	1.0
4	13	0.08	0.00	0.07	1.0
5	7	0.00	0.29	0.00	0.044
6	8	0.00	0.00	0.00	1.0
7	3	0.00	0.00	0.00	1.0
8	2	0.00	0.00	0.00	1.0
9	0				
10	7	0.14	0.00	0.00	1.0
11	4	0.00	0.00	0.00	1.0
12	1	0.00	0.00	0.00	1.0
Bigeye tag recovery area = whole Indian Ocean					
0	11	0.64	0.73	0.27	<0.001
1	15	0.33	0.20	0.07	0.04
2	15	0.20	0.13	0.13	0.17
3	12	0.08	0.08	0.17	0.46
4	8	0.50	0.25	0.13	0.06
5	4	0.00	0.00	0.00	1.0
6	3	0.00	0.00	0.00	1.0
7	1	1.00	0.00	0.00	1.0

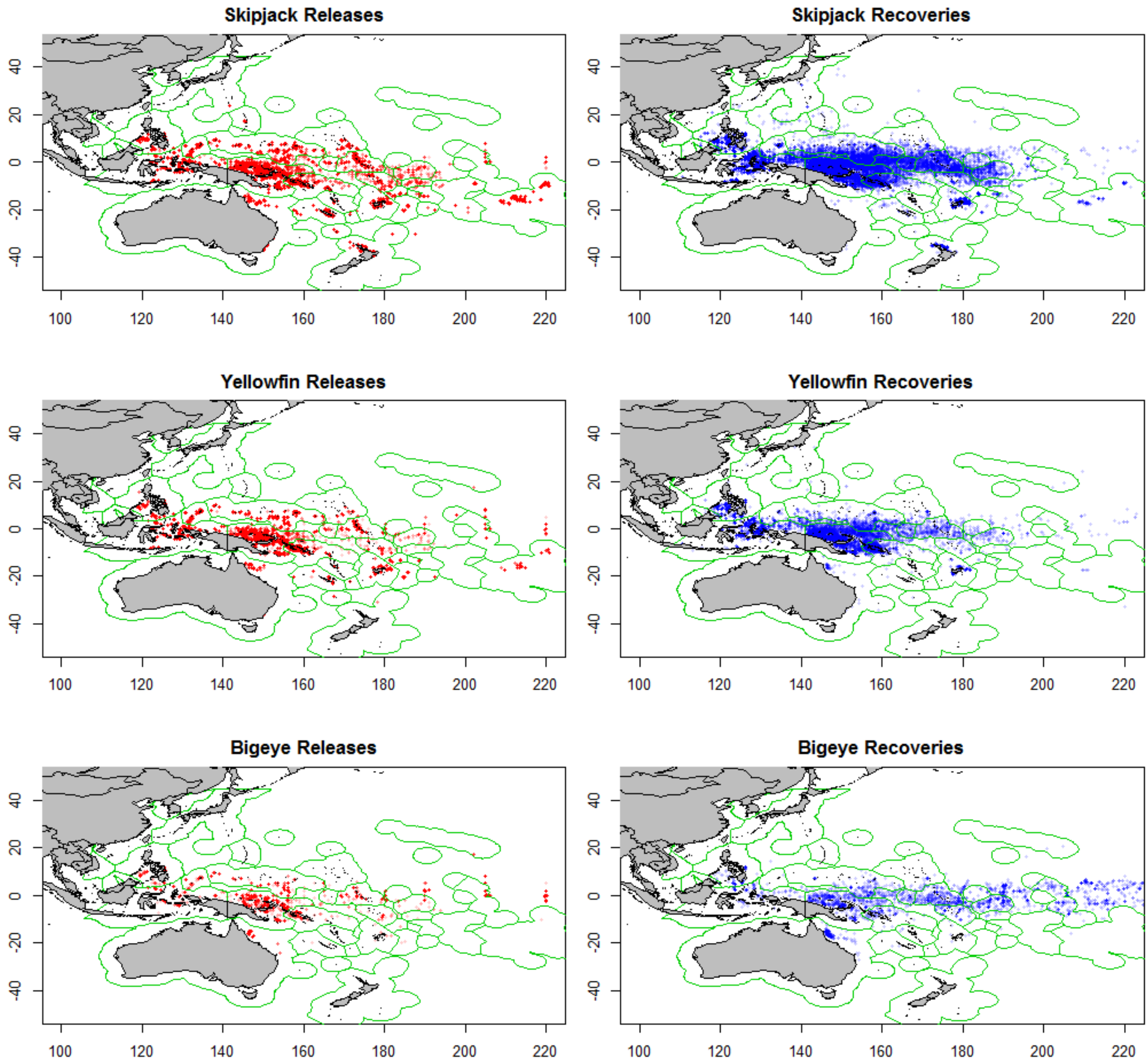
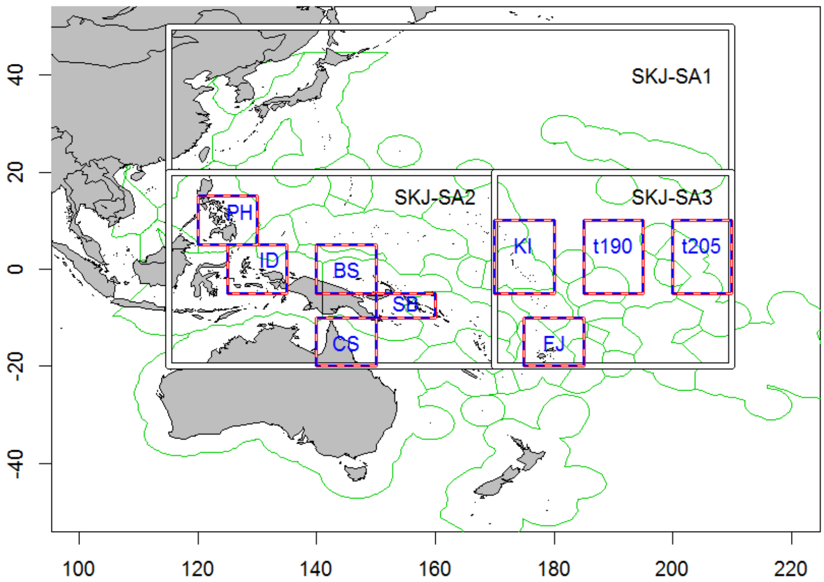


Figure 1. Release and recovery positions from the combined WCPO tagging programmes by species.

CPRD Skipjack Release and Recovery Areas



Yellowfin/Bigeye Release and Recovery Areas

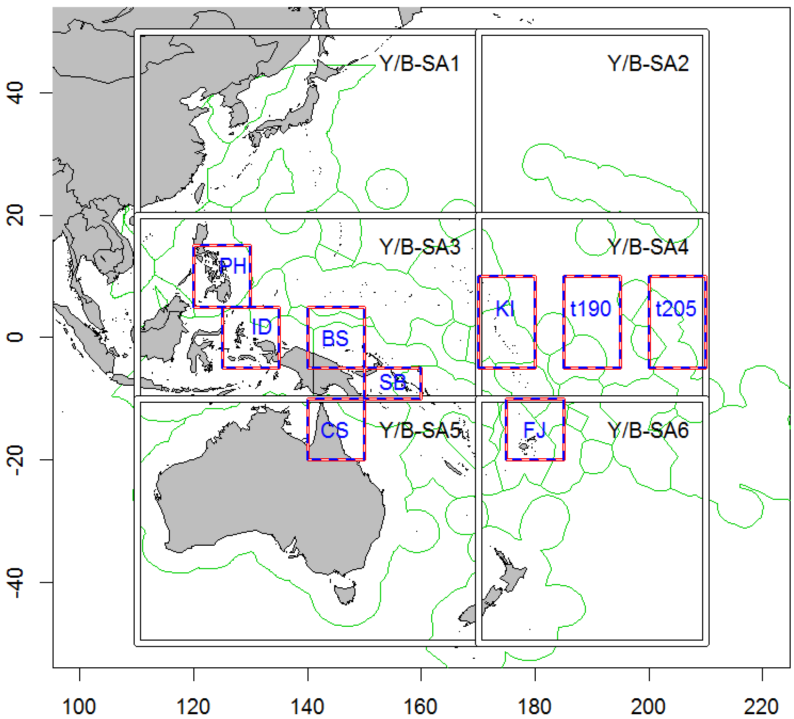


Figure 2. WCPO release and recovery areas defined for the CPRD analysis. Large black boxes represent the recovery areas as defined by the stock assessment spatial structure of skipjack (top) and bigeye and yellowfin (bottom). The small rectangles represent tag release areas defined for the CPRD analyses.

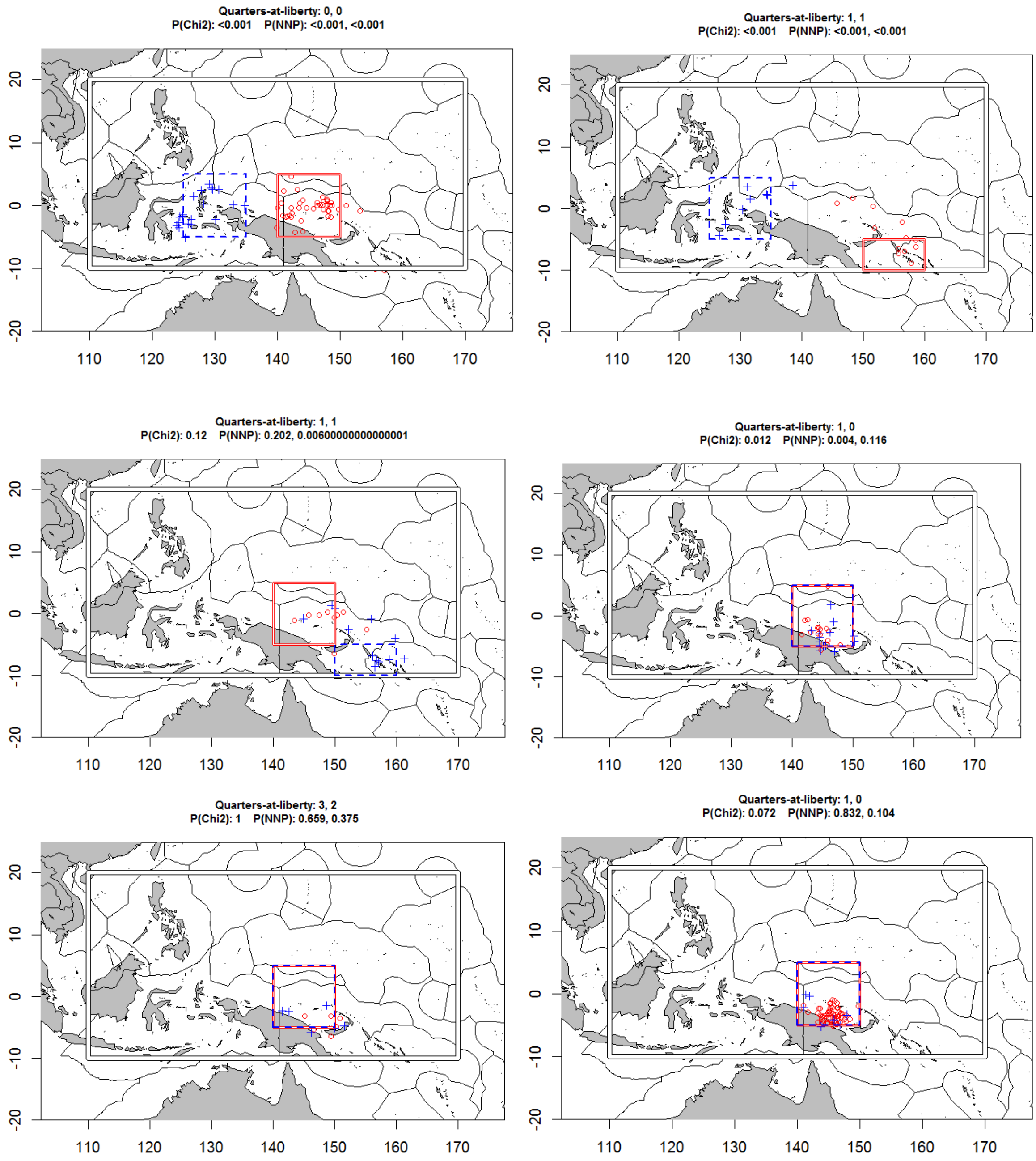


Figure 3. Example bigeye tuna CPRD events illustrating different degrees of tag mixing, and the corresponding P-values from the chi-square and nearest neighbour permutation tests (first P(NNP) value is the one summarized in the results). Large black boxes represent the recovery areas, small coloured boxes represent release areas, points represent (jittered) recovery locations with colours corresponding to the release areas.

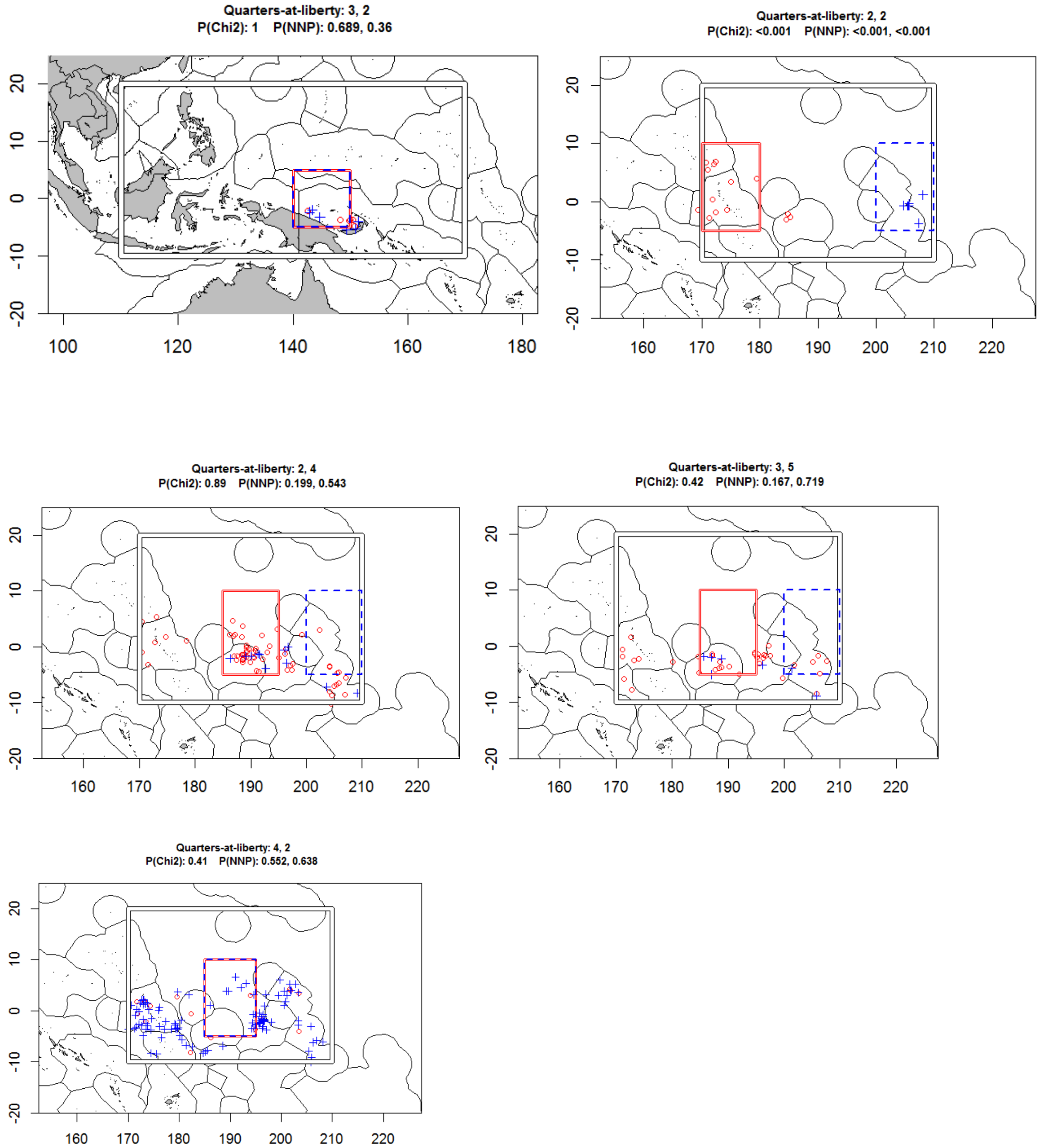


Figure 4. All bigeye tuna CPRD events with minimum time-at-liberty of 2 quarters (consistent with the assessment model mixing assumption). P-values from the chi-square and nearest neighbour permutation tests are shown (first P(NNP) value is the one summarized in the results). Large black boxes represent the recovery areas, small coloured boxes represent release areas, points represent (jittered) recovery locations with colours corresponding to the release areas.

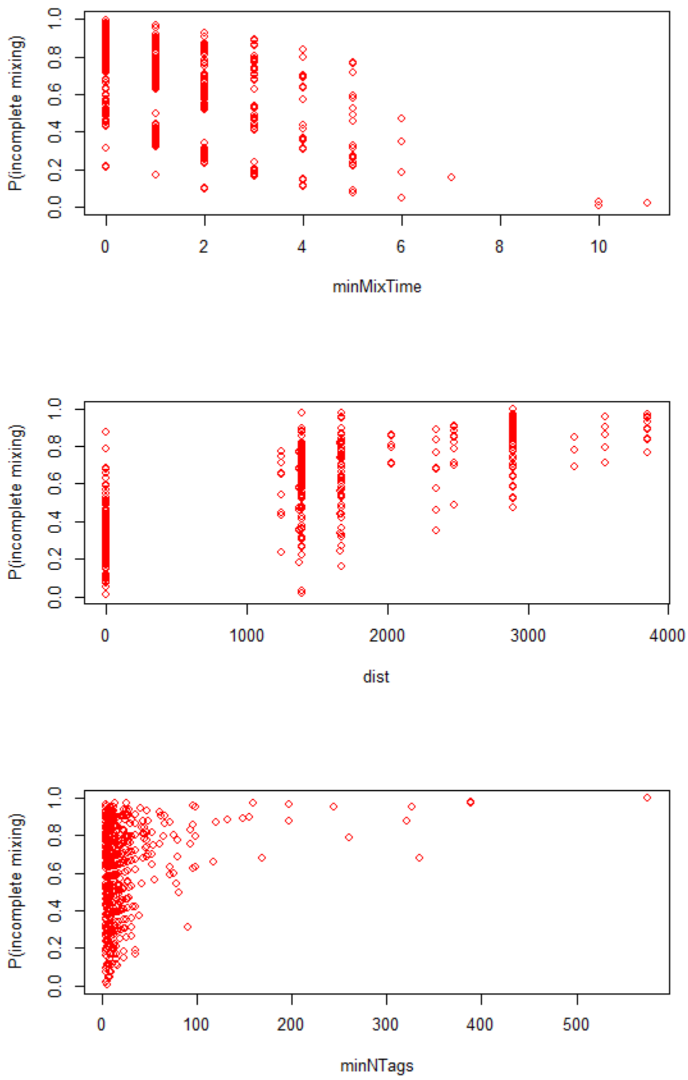


Figure 5. Predictions of the probability of identifying incomplete tag mixing as a function of the 3 main variables in the WCPO multiple logistic regression model.

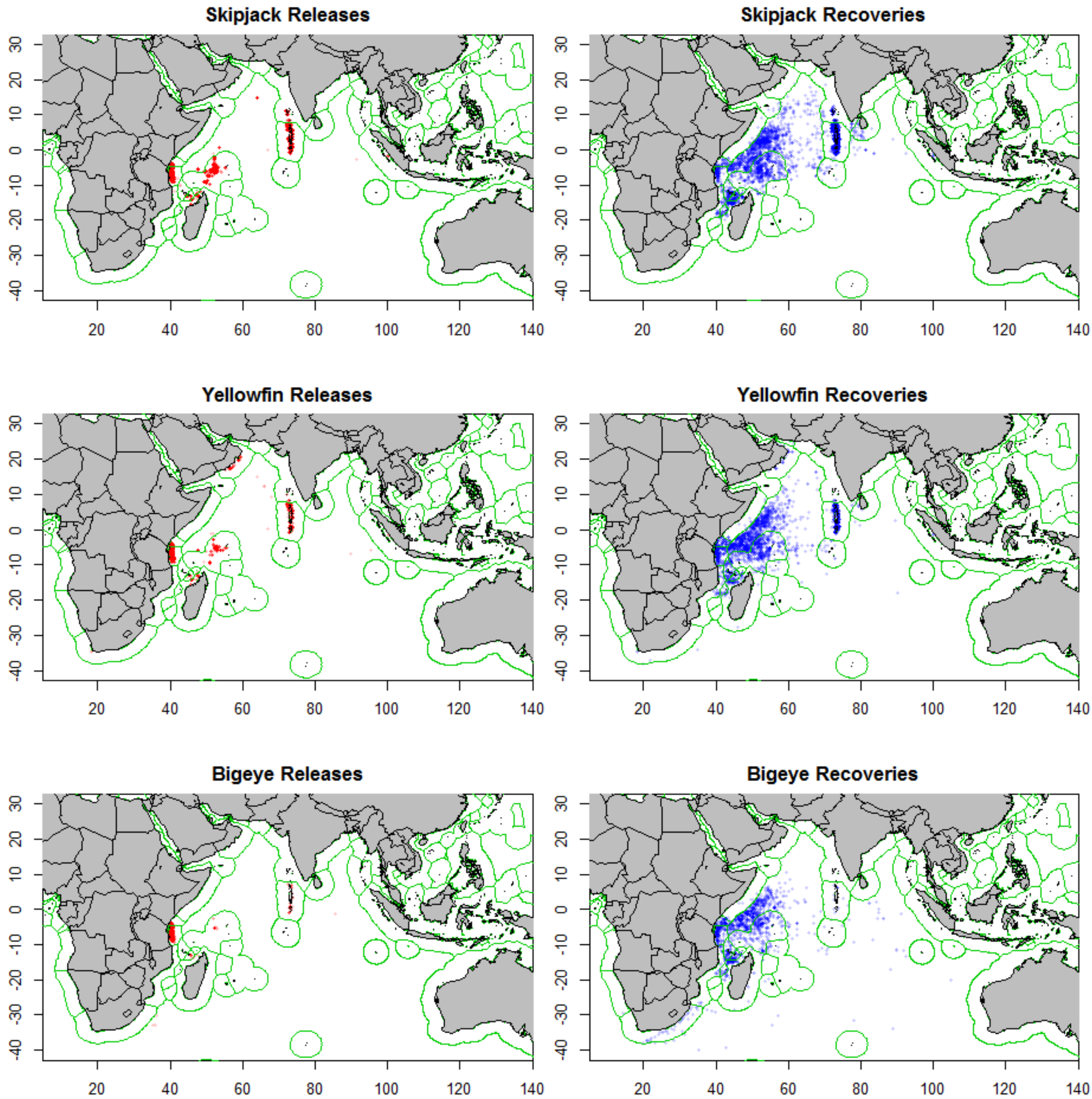


Figure 6. Release and recovery positions from the combined Indian Ocean tagging programmes by species.

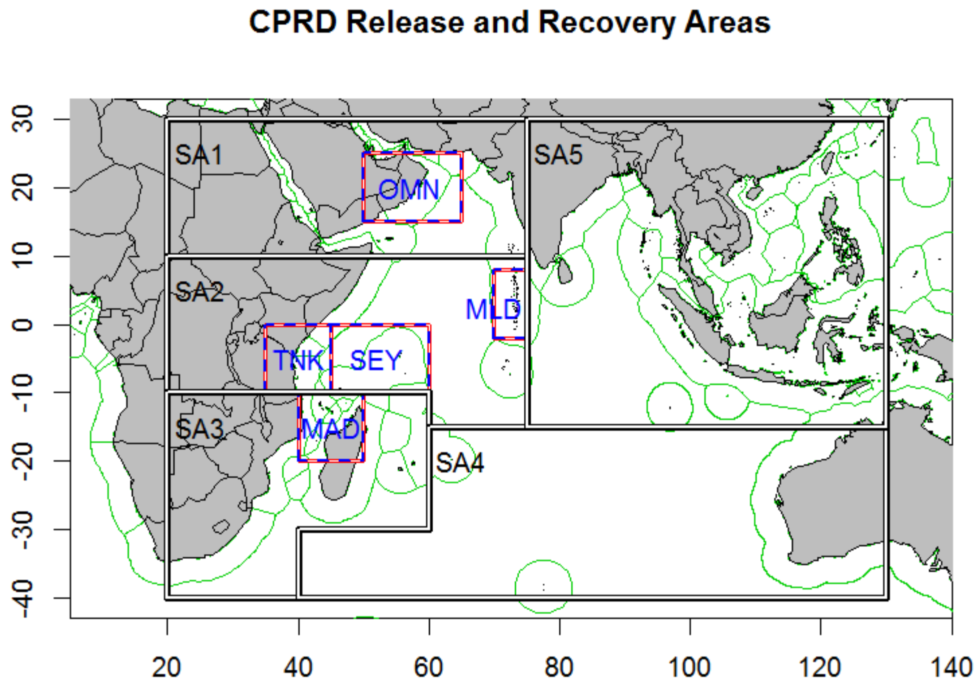


Figure 7. Indian Ocean release and recovery areas defined for the CPRD analysis. Large black boxes labelled SA1-SA5 correspond to the spatial structure of the yellowfin stock assessment, while skipjack and bigeye recovery areas were assumed to be the whole Indian Ocean. The small coloured rectangles represent core tag release areas.

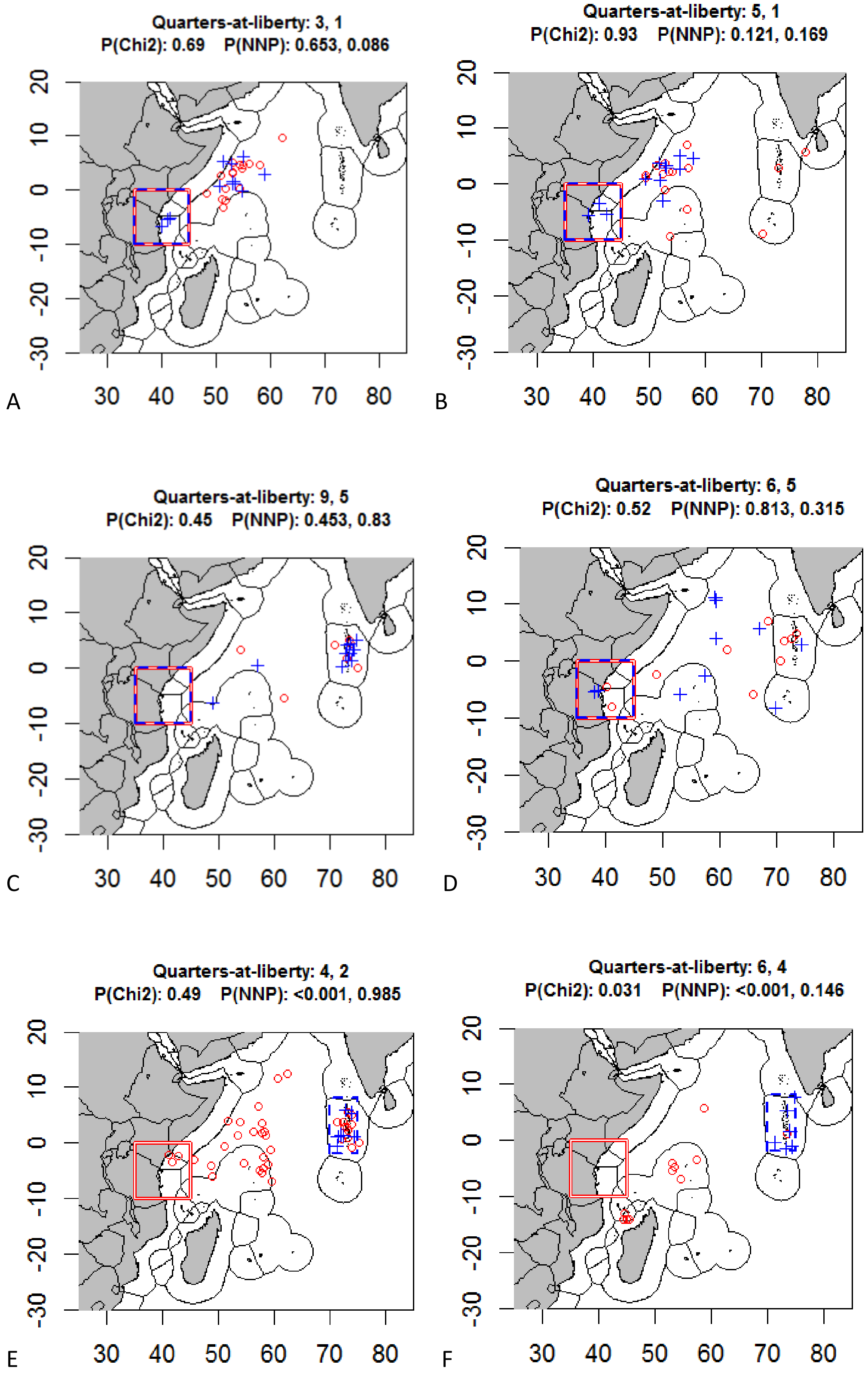


Figure 8. Some example Indian Ocean skipjack CPRD events. P-values from the chi-square and nearest neighbour permutation tests are shown (first P(NNP) value is the one summarized in the results). Small coloured boxes represent release areas, points represent recovery locations (jittered) with colours corresponding to the release areas (Red circle = release event 1). The whole Indian Ocean represents the recovery region.