**SCIENTIFIC COMMITTEE**
**SIXTH REGULAR SESSION**

10-19 August 2010
Nuku'alofa, Tonga

# BACKGROUND ANALYSES IN THE DEVELOPMENT OF THE 2010 WCPO BIGEYE TUNA ASSESSMENT

**WCPFC-SC6-2010/SA-WP-01**

Shelton Harley, Simon Hoyle, Peter Williams, and John Hampton

Oceanic Fisheries Programme, Secretariat of the Pacific Community, Noumea, New Caledonia

# BACKGROUND ANALYSES IN THE DEVELOPMENT OF THE 2010 WCPO BIGEYE TUNA ASSESSMENT

*SHELTON HARLEY, SIMON HOYLE, PETER WILLIAMS, AND JOHN HAMPTON*

## EXECUTIVE SUMMARY

This paper describes background analyses undertaken in the development of the 2010 bigeye tuna (BET) stock assessment for the western and central Pacific Ocean (WCPO). Most of the analyses focus on attempting to resolve the principle concern in the 2008 and 2009 assessments, namely the strong recruitment patterns estimated in recent BET assessments. We attempt to identify the key data inputs and model assumptions that are responsible for the trends and see if improvements can be made. The document draws on work undertaken for the SPC Pre-Assessment workshop held in Noumea in April 2010 and subsequent analyses.

There were three steps to the analyses described in this paper:

1. Exploration of the key data indicators (specifically catch, CPUE and size data) from the 2009 BET assessment;
2. Evaluation of the current grouping of longline fleets within fisheries in the MULTIFAN-CL assessment ; and
3. An initial suite of sensitivity analyses using the 2009 BET assessment.

These analyses indicate that:

- There is significant data conflict among the various data sources (e.g. size, catch, and CPUE data) in region 3 which provide the model conflicting signals regarding recruitment and ultimately stock status trends
- [Fleets flagged to FM, MH, and PH, previously included in the LL ALL fishery, have fishing patterns more similar to the TW OS fleet, and hence would be more appropriately modeled as a separate fishery (from LL-ALL) in future assessments.
- There are strong patterns within the JP length data which appear to be driven by spatial patterns in fish size
- JP length data may not be representative of the catches in some regions / years
- The CPUE trend in region 3 appears to be the primary factor driving the recruitment trend in that region and the relatively stable CPUE is strongly inconsistent with the increased catches seen in that region at the estimated biomass level.
- The CN/TW-OS size data are extremely influential on the assessment, in particular the early estimates of recruitment and growth
- The high (and incorrect) estimates of ID longline and small fish catches were resulting in a more pessimistic view of stock status.

# INTRODUCTION

One of the key features of recent bigeye tuna (BET) assessments for the western and central Pacific Ocean (WCPO) has been the strongly increasing trend in recruitment estimated for the western equatorial region of the assessment (Region 3) which leads to a strong increasing recruitment trend for the assessment as a whole (Figure 1). This trend leads to three important concerns for management advice: 1) estimates of steepness which are near 1, suggesting that the WCPO BET stock is extremely resilient to overfishing and that the level of spawning biomass that will support the MSY is only 20% of the unfished level; 2) large differences between estimates of 'virgin' biomass and the level of biomass predicted to occur in the absence of fishing (Figure 2); and 3) recent catches being much higher (around double) the estimated MSY (Figure 3).

The paper continues the investigations of Langley and Hoyle (2008) to determine the cause(s) of the recruitment trend and whether improvements to data inputs or model structures can result in more plausible recruitment trends and therefore a more credible stock assessment. There were three steps to this investigation:

1. Exploration of the key data indicators;
2. Evaluation of the current grouping of longline fleets within fisheries in the MULTIFAN-CL assessment; and
3. A set of sensitivity analyses using the 2009 assessment

When considering recruitment patterns from MULTIFAN-CL assessments which have regional substructure, it is important to note that movement and recruitment can often be correlated, and because movement is assumed to be time-invariant (i.e., the same age-specific, seasonal, movements occur every year), departures from this assumption can manifest in trends in recruitment. This is particularly the case for areas and periods for which juvenile BET are not observed in the catch. The model simply needs to ensure that the fish are there in time to be taken in longline fisheries – it can either move them there or they can recruit there. Therefore, it is the overall trend in recruitment which is more important than the regional trends for this investigation, though we do look closely at some regional patterns in this paper.

# EXPLORATION OF KEY DATA INDICATORS

There are five key data inputs to the bigeye assessment: 1) catches, 2) catch rates or CPUE, 3) length data, 4) weight data[1], and 5) tagging data and in this section we will examine the first four of these for each of the regions used in the assessment model to look for evidence of inconsistency or conflict between data sources. We do not look at the influence of the tagging data in this paper. These analyses were undertaken based on the dataset used for "run 10" from the 2009 assessment which used the lower estimates of purse seine catches determined from the grab sampling analysis.

[1] In the remainder of the paper whenever the phrase "size data" or "size frequency data" is used we are referring to both length and weight frequency data.

Most of the catch of BET is taken in a narrow band between 10S and 10N with much lower catches at higher latitudes (Figure 4). Purse seine catches are higher to the south and west of the equatorial region.

Over the period of the assessment (since 1952) there has been a steady increase in BET catches (Figure 3). Annual longline catches were relatively stable between 40,000 – 60,000 mt from 1960 to the late 1990s and increased a little further since then to around 70,000 mt. The has been an increasing trend since the 1970s in the catches of small BET from the domestic fisheries in Indonesian and the Philippines and a sharp increase in purse seine catches of BET since the expansion of the FAD fishery in the mid 1990s. In recent years the longline catches have been around 60% of the total catch, with equal contributions from purse seine and the ID/PH domestic fisheries[2].

Regional trends in catch, CPUE, and size data are presented in Figure 5 and Figure 6 and are described briefly here:

- For region 1 there has been a decline over time in total catches and longline CPUE and, with the exception of the length data from the 1970s, median length and weight has also declined over this period. It is interesting that the fish in region 1 were typically the smallest taken in the WCPO.
- For region 2 there are similar patterns to region 1 with declines in catch, CPUE, and fish size. The decline in fish size in region 2 since 1950 is the greatest of any of the regions.
- For region 3 there has been a gradual increase in longline catches, but a dramatic increase in overall removals, especially since the late 1990s. Longline CPUE has been variable through time, and the overall decline in CPUE is minimal. In contrast to the other regions, fish sizes show little change over the last 50+ years. Fish taken in region 3 were among the smallest fish in the 1950s, but are now the largest due to the declines in other regions.
- For region 4 there has been a gradual increase in longline catches through time and a slow, but consistent, decline in longline CPUE. There has been a considerable decline in the sizes of fish taken in this region.
- For region 5 the patterns are similar to region 4 with gradually increasing catches coupled with a consistent decline in CPUE. Fish sizes have also declined.
- For region 6 both catches and CPUE have been quite variable and recent CPUE is based on low levels of effort and is less reliable than trends in other regions. Fish sizes have declined through time.

For regions 1, 2, 4, 5, and 6 the trends between catch, CPUE, and size data are generally consistent and reflect the expected trends in fishery indicators in exploited fish stocks. However, the lack of a response in the CPUE data from region 4 to the increased longline and purse seine catches since the late 1990s should be expected to have some impact on the assessment, e.g. increased recruitment to fit to the CPUE data.

[2] The assumed purse seine catches here are about half the level estimated based on the spill sampling methodology and the catches for ID/PH domestic fisheries have all been revised down in 2010. There has also been a large reduction in the ID longline catches.

The patterns in the data for region 3 are very puzzling and these conflicts are undoubtedly playing a key role in some of the key parameter patterns (e.g. recruitment, biomass etc) estimated for that region. It is not surprising that the model estimates strongly increasing recruitment in region 3 in the face of increasing catches, but essentially stable CPUE and size data. Data for region 3 are further examined in the rest of this paper.

# EXPLORATION OF FLEET GROUPINGS

The second step involved reviewing the longline fishery definitions used in the 2008 and 2009 bigeye assessments (Table 1).

An important part any stock assessment is the definition of fisheries and this is particularly difficult for WCPFC assessments due to the large number of fleets that operate and the large spatial area covered. Theoretically a "fishery" is comprised of one of more fleets which operate in a similar way (i.e. operational characteristics) in a defined spatial region over a particular time period. In MULTIFAN-CL a fishery is defined to have a common estimate of selectivity, i.e. they target/catch fish of a particular size range; therefore it is necessary that size or weight frequency data is available for at least one fleet within a fishery in order to allow estimation of that selectivity curve. In the case of WCPO bigeye, we have many different fleets operating across the convention area and some of these have changed flag over time, and/or their targeting practices. For some fleets we have some information on their fishing practices (e.g. some off-shore fleets based in PICT EEZs), but for most we have none (e.g. DWFN fleets), and we do not have size/weight frequency data for all fleets. In some instances, the data provided do not differentiate between two fleets with clearly different operational modes, for example, there is no differentiation in the aggregate catch/effort data for what we term as the Chinese longline 'offshore' fleet, based in Micronesian countries, and the Chinese 'distant-water' longline fleet which fishes in the central and eastern tropical WCPFC Convention Area.

Since these fishery definitions were made, considerably more length and weight frequency data has become available for the various fleets in the WCPO so it is now timely to re-assess the previous groupings of fleets into fisheries. Previously it was necessary to group many fleets into a single fishery because no size frequency data was available, nor sufficient operational information to indicate the sizes of fish likely to be taken.

Incorrect combination of fleets, areas, or time periods (due to different fishing practices or sizes of fish) can cause considerable problems in stock assessments. In particular, assessments such as those conducted in the WCPO and EPO are sensitive to trends in the size frequency data for estimating important parameters such as time trends in recruitment. A recent review of the EPO bigeye assessment indicated that a change in fishing practices in a key fleet led to it catching larger fish which was incorrectly interpreted by the model as reflecting a strong increase in recruitment[3].

In this section we describe the current longline fisheries definitions used in the assessment and examine the available data to see which fleets are the most important for the different fisheries and

[3] http://iattc.org/PDFFiles2/BET-01-Meeting-report-ENG.pdf

what sizes of fish they catch. Using this information we will evaluate the current definitions to determine if there are other groupings of fleets which are more appropriate.

The data used in the assessment is from many sources and a summary for the major DWFN fleets is provided in Table 2. The key thing to note from this table are that catch and effort data for Japanese flagged vessels are not split between distant water (DW) and off-shore (OS) components, but that there is separate weight data for the OS component. Chinese catch and effort are also not separated and the current assessment only includes length and weight data for the OS component. All data from TW is split into DW and OS components, and Korea essentially only has a DW fleet.

## CURRENT LONGLINE FISHERY DEFINITIONS

The longline fishery definitions used in the 2008 and 2009 assessments are provided in Table 1 and the spatial regions used are presented in Figure 4.

In this section we will review data summaries for each of the mixed fleet longline fisheries defined in Table 1 (i.e. we will not focus on the single fleet fisheries). We then provide some suggestions for improvements to the fishery definitions or other analyses that should be undertaken. We focus on data post 1970 as there is very little non-Japanese data prior to this time in the assessment.

*Fishery 1: Longline fishing in Region 1(Figure 7)*

- The catch in this fishery is almost entirely dominated by JP vessels. The mix between DW and OS components is not known, but we might expect considerable OS effort in this area.
- Length frequency data is very patchy and mostly from JP-DW, but there is some questionable AU samples in 2007 and some TW-DW samples in 2005 (a year for which when there was also very little JP-DW data). The median length declined over 1970-1990 and has been stable / slightly increasing since then.
- There is a large amount of weight frequency data from JP-DW, but it is a little patchy since 2003. Median weight declines from 1970 through to the end of the time series.
- Recommendations for fishery definitions:
    - Keep current definition, but exclude all non-JP data.
    - Consider downweighting the length data.
- Recommendations for further analysis
    - Investigate validity of the AU length samples [immediately]
    - Investigate spatial patterns within JP-DW size data [longer term]

*Fishery 2: Longline fishing in Region 2 – excluding US (Figure 8)*

- The catch in this fishery is almost entirely dominated by JP vessels prior to 2002, but there is some VU, CN, and TW-DW since then though JP is still the dominant fleet. The mix between DW and OS components is not known, but we would expect less OS effort in this region when compared to region 1.
- There is a reasonable amount of length frequency samples and they almost all come from JP-DW. There are some years with limited JP-DW samples, e.g. 2005 when samples from TW-DW are the dominant data source. Median length was relatively stable around 120cm

from 1970-1982 and then it dropped over 30cm for five years before increasing again. Median length has averaged around 110cm since then.

- There is a large amount of JP-DW weighting frequency data, but the time series essentially stops in 2003. Median weight essentially halved from 1970-1980 and has been stable/slightly increasing since this time.
- Recommendations for fishery definitions:
    - Keep current definition and exclude all non-JP size data
    - Consider downweighting the length data
- Recommendations for further analysis
    - Investigate spatial patterns within JP-DW size data [longer term]
    - Consider splitting off JP-DW as a separate fishery if the other fleets continue significant activities in the area and size data are available for them [longer term]

*Fishery 4: Longline fishing in Region 3 – excluding Bismarck Sea, PG, CN, and TW-OS fleets (Figure 9)*

- This is one of the most important fishery definitions in the current assessment and represents a mix of fleets and data sources.
- Prior to 2004, JP was the dominant fleet with consistent, but smaller contributions from KR-DW and TW-DW. Since the mid-1990s there has also been a small but consistent contribution from FM vessels. From 2004 there was a large catch from ID vessels of the same order of the JP catchesFor JP, the mix between DW and OS components is not known, but we could expect considerable levels of OS effort in this region.
- The ID catch history, in particular the post-2004 numbers require urgent re-evaluation as they are likely influential in the assessment.
- Prior to 1990 only JP-DW length samples were available, but since then data has become available from a range of sources, but not consistently through time. These include: JP-OS, FM, KR, and PH. These other sources represent almost all data in most years since 1998.
- There are strong trends in the JP length data, in particular an abrupt decrease in median length since 1999. In general, many of the other fleets have median sizes greater than those from the JP-DW fleet, which is consistent with our current understanding that these other fleets fish in different ways and may target larger bigeye tuna. The JP-OS sizes were towards the small sizes of the other fleets.
- Almost all the weight frequency data is from JP vessels, but it is dominated by JP-OS vessels since 1990 and there are very few JP-DW samples after 2003. There are a few samples from FM vessel in recent years.
- The JP median weights are typically lower than those of the other fleets, and for years where the number of samples is sufficient, there appears to be little difference between the JP-OS and JP-DW median weights, but there is a sharp decline in median weights in the JP-DW data post 2000, similar to the pattern observed in the length data.
- There are at least two critical issues to be resolved with the size and weight frequency data currently used for this 'fishery':
    - There are strong latitudinal patterns in the sizes of fish recorded in the JP data with region 3 and it is possible that the data are being 'contaminated' by samples from

    regions where the fish are smaller, but there is little effort (e.g. outside the core fishery); and
  - There are several fleets that have fishing characteristic more aligned to the TW-OS vessels (see Fishery 5) and this is supported by the larger median lengths and weights of these fleets.
- Recommendations for fishery definitions:
  - Move PH, ID, FM, MH, GU to the current Fishery 5
  - Keep the remainder of the fleets in Fishery 4, but exclude all non-JP size data
  - Consider downweighting the length data
- Recommendations for further analysis
  - Further investigate spatial patterns within JP-DW data for contamination by samples outside the core BET fishing area –[immediately]
  - Investigate the large ID longline catches [immediately]

*Fishery 5: Longline fishing in Region 3 –CN and TW-OS fleets (Figure 10)*

- This fishery is dominated by TW-OS fishing with CN vessels (presumably OS, but the data is not available to make this separation) arriving since 1990.
- Length frequency data are only available since 1990 and more than 50% of the data are from CN in most years, though there was an approximately equal split of samples in 2006-07. Through to the early 2000's median lengths show a slow decline, but with median sizes much larger than those for the JP fleet (from Fishery 4). From the early 2000's the CN data show a continued decline, but the TW data show a strong increase. It is not known if the difference is due to a change of operational practices, ports of landing (where size data may or may be collected) or fishing areas.
- The majority of the weight data comes from the TW-OS fleet – in closer proportions to those of the catch. The same pattern seen in the median lengths is also seen in the weights with a significant difference in the patterns between CN and TW occurring in the early 2000's.
- Observer data suggest that the operational practices of the Chinese offshore fleet started to diverge from the TW-OS fleet in about 2004 (i.e. fishing deeper). The extent of this change throughout the fleet is currently not known.
- Recommendations for fishery definitions:
  - Include the other fleets from Fishery 4
  - Continue to assume asymptotic selectivity for this fleet
  - Consider downweighting the size data until an improved understanding of the operational and spatial aspects is available.
- Recommendations for further analysis:
  - Examine the recent spatial distributions of the CN and TW-OS samples to see if a reason can be found for the observed differences [immediately]
  - Further examine the CN data to see if it is possible to separate DW and OS components [longer term]
  - Investigate spatial patterns within the fleets that make up the fishery [longer term]

*Fishery 7: Longline fishing in Region 4 – except US, CN and TW-OS fleets (Figure 11)*

- JP and KR-DW are the dominant fleets in this fishery through time with fishing from TW-DW and other fleets since 2000. For JP, the mix between DW and OS components is not known, but we would expect far less OS effort in this region when compared to regions 1 and 3.
- Up until 2004 almost all the length samples came from JP-DW, but since 2005 there have been a very large number of samples from the TW-DW fleet and these fish are much larger (15-20cm) than the JP fish and cause an overall upward trend in fish sizes from 2005-07.
- All the historical weight data is from the JP fleet, but the number of JP samples is quite low after 1990 and stops around 2003. After this time there are a few samples from the FM fleet.
- Recommendations for fishery definitions:
    - Move the FM and MH fleets to Fishery 8 (TW-OS and CN)
    - Keep the remainder of the fleets in Fishery 4, but exclude all non-JP size data
    - Consider downweighting the length data as a sensitivity analysis
- Recommendations for further analysis:
    - Determine if any size data exist for the KR fleet and then compare with JP data. Consider splitting KR from JP is suitable size and CPUE data exist [longer term]

*Fishery 8: Longline fishing in Region 4 –CN and TW-OS fleets (Figure 12)*

- This fishery is currently dominated by CN as there is very little TW-OS fishing in region 4. The catch has only been at a high level since 2004. It is likely that most of the CN effort is DW rather than OS (opposite from the pattern in region 3), but there is no size data for CN-DW.
- The length and weight data is almost entirely from CN. The fish appear to be slightly larger than those taken in Fishery 7, but the data comes from the OS fleet rather than the DW fleet which is responsible for most of the catches. There is insufficient evidence to assume that the selectivity in this fishery is asymptotic, let alone the same as Fishery 5.
- Recommendations for fishery definitions:
    - Include additional fleets from Fishery 7
    - Exclude size data as likely to be almost entirely CN-OS while the catch is likely to be CN-DW.
    - Link the selectivity to fishery 7 for the interim.
- Recommendations for further analysis
    - Further examine the CN data to see if it is possible to separate DW and OS components [longer term]
    - Investigate spatial patterns within the fleets that make up the fishery [longer term]

*Fishery 10: Longline fishing in Region 5 – All except AU(Figure 13)*

- JP-DW is the dominant fleet in this fishery and prior to 1990 was comprised almost entirely by JP-DW and TW-DW catches. Since then numerous fleets have entered the fishery.
- There is only very limited length data for this fishery and up until the late 1990s it all came from JP. Since then there has been little JP length data even though the fishery has remained. Since the late 1990s there has been much more length data for this fishery and it comes from NC, SB, and TW-OS. There is considerable variation among years, fleets, and

years within fleets in the median length, but there is general agreement of declining sizes over the last 20 years.

- More historical weight samples are available than length samples and again these are from JP. As was seen with the length samples, the JP samples decrease considerably after the late 1990s and samples are then available from NC and SB. All the fleets show a decrease through time in median size.
- While there is a huge mix of fleets in this fishery, it is outside the main fishing area, and the size trends are not really inconsistent across fleets.
- Recommendations for fishery definitions:
  - Leave as is – but consider downweight the length data for sensitivity
- Recommendations for further analysis:
  - None at this time, consider further after analyses for regions 1-4 are complete.

*Fishery 12: Longline fishing in Region 6 –excluding PICT fleets (Figure 14)*

- This fishery is comprised of catches from JP-DW, KR, TW-DW, and CN in vary proportions through time. CN only arrived from 2002 and prior to this the other fleets 'took turns' at being the dominant fleet in this fishery.
- The amount of length data is very low, especially historically when it was only available for JP (even though the other fleets were taking much of the catch). Since 2004, TW-DW and CN are the main sources of length data. The patterns in median lengths are messy.
- There is slightly more JP-DW weight data than length data, but it is also very variable across years and stops in 1998.
- Recommendations for fishery definitions:
  - Leave as is, but consider downweighting the length data as a sensitivity
- Recommendations for further analysis:
  - If available, use TW-DW CPUE to calculate standardised effort for this fishery

*Fishery 13: Longline fishing in Region 6 –PICT fleets (Figure 15)*

- This fishery is a mixture of numerous fleets which developed since 1990. FJ is the dominant fleet.
- Length data are available for most of the fleets, but there is a higher proportion of TO length compared to their catches. The median lengths are much larger for NZ than the other fleets, but this has little impact as the overall median track FJ quite closely.
- Weight data became available after length data and mostly come from FJ and TO and NZ (only in the last two years).
- Recommendations for fishery definitions:
  - Leave as is, but consider downweighting the size data
- Recommendations for further analysis
  - Consider a separate fishery just for FJ [longer term]

*Fishery 23: Longline fishing in the Bismarck Sea (Region 3) –same fleets as fishery 4 (Figure 16)*

- Up until the mid-1990s catches for this fishery were dominated by JP, but TW-DW was significant in some years. Since the decrease in JP catches, SB catches increased. Since 2004 there are large ID catches, but as mentioned previously, these are uncertain.

- Length data are relatively limited for this area and we solely from JP up until 1990, there were some KR samples in 1992 and since 1996 the SB samples dominate. The patterns in the median lengths are messy, but suggest an increase over time which appears to be driven by changes in the fleets from which the data come more than anything else.
- The patterns for the weight frequency data are similar to that for length data and there are no weight frequency samples after 2000. The trend in median weight is flat / declining over the period.
- <u>Recommendations for fishery definitions:</u>
    - <u>Exclude all fleets which were moved to fishery 5</u>
    - <u>Exclude all non-JP size data</u>
    - <u>Consider downweighting the length data</u>
- <u>Recommendations for further analysis</u>:
    - <u>None at this time</u>

### *RESPONSE TO URGENTLY NEEDED ANALYSES*

In the course of the detailed examination of the data described above, several issues of immediate importance were encountered and these are described here:

1. The misplaced positional information for the AU length frequency samples collected by observers were reassigned to the correct hemisphere and therefore removed from fishery 1. Further, some samples from TO and FJ were found to have been wrongly allocated and these errors were fixed in the database.

2. The large increase since 2004 in ID longline catches was of considerable concern. A dedicated annual catch estimates workshop was conducted in early 2010 between WCFPC, SPC and ID fisheries officials which led to an improved time series of longline catches which will be used in the 2010 assessments. The impact of these revised catches will be examined in the sensitivity analyses described later in this paper.

3. The location of the JP length samples in region 3 was further examined and some interesting patterns were discovered. Figure 17 presents the location of samples and resulting length frequency distribution for region 3 in five year blocks throughout the time series. The figure includes the 10th, 50th, and 90th percentiles for the length samples for both that five year period and the overall time series. In periods when many samples are taken from the northern or eastern part of region 3 there are far more small fish in the length samples. A comparison to the locations for the weight frequency samples (Figure 18) indicates a far more consistent sampling pattern for the weight data. Given that the length samples are not taken directly from the fishing fleet, it seems likely that they are less representative of the catch than the weight frequency samples. Preliminary analysis that used a GLM to examine patterns in the length and weight frequency data indicate that for the length data the fish are much small in the north and east (Figure 19). The longitudinal pattern is less clear for the weight data, but weights are also much smaller in the north. <u>Therefore, further</u>

examination of the spatial aspects of the JP length and weight sampling data is strongly recommended for all regions, but with an initial focus on regions 3 and 4.

4. There was also a troubling divergence of the trends in median length and weight for CN and TW-OS in fishery 5. In Figure 20 was can clearly see that in recent years the samples from these fleets are coming from completely different parts of region 3. It seems likely that the fishing locations are leading to differences in the sizes of fish taken, though there could also be differences in fishing patterns. With the addition to this fishery of further data from FM and MH vessels, some detailed analysis of these data is warranted.

# SENSITIVITY ANALYSES USING THE 2009 ASSESSMENT

Due to delays in the provision of data by many WCPFC members, it was necessary to undertake this investigation of data sources / model assumptions using the 2009 assessment.

The main focus of these model runs was to investigate the major problem identified in previous bigeye assessments, namely the strongly increasing trend in recruitment estimated by the model and the poor fit to the Japanese longline CPUE in region 3. For these examinations we used the 2009 model run which used the spill sampling estimates of purse seine catch (run 14). Run 14 was used for two reasons: 1) these purse seine estimates were considered more plausible than the previous estimates based on grab sampling and logsheets, and 2) this model had one of the largest trends in recruitment.

## MODEL RUNS

Twenty-three model runs were undertaken to investigate five main areas of uncertainty in the assessment and these are summarized in Table 3 and described in greater detail below.

### *TIDY-UP*

Some of the observed data are associated with large residuals in the MULTIFAN-CL assessment and while it cannot be determined if these data are valid or not, it is important to examine their impact on the assessment.

There are a small number of fishing events where there is effort but no catch. MFCL cannot handle zero catches so they are typically set to a small arbitrary number (0.1 in the case of the bigeye assessment). In the model these are often associated with large negative effort deviates and two examples are provided in Figure 22. In model 2 we excluded the fishing events with zero catch for several fisheries – this represented 31 observations. We did not exclude those records for the Japanese coastal pole and line and purse seine fisheries in region 1 as these observations actually provide important information of the seasonality of the availability of juvenile bigeye in these areas.

Residual plots of the fit to the Japanese length frequency data in regions 1-4 all showed a similar pattern with a large number of small bigeye in the catches from 1955-1960 that the model was unable to predict (Figure 23). The model also shows, particularly in regions 2, 3, and 4, that the model is unable to predict all the large fish in the early 1950s. Dominance of the catch by large fish seems more likely than small fish during the first ten years of a fishery. These same patterns were not as clear in the weight data. At this time we have not determined if these data are valid or whether there may have been some problems with the sampling. This should be addressed, but in the interim we excluded these data from model 3 to determine if they were having an impact on the model, in particular the earlier recruitment estimates.

### *Size data*

Size data (length and weight observations) carry considerable weight in the overall objective function for the stock assessments undertaken using MULTIFAN-CL. These data should provide the model the information it needs to estimate selectivity curves for each fishery, and assist in the estimation of growth and recruitment when there is good modal progression in the data. Previous assessments have indicated some evidence of conflict between the size data and other data available to the model. This is most easily demonstrated when alternative weightings are assumed for the size data (see run 11 in Harley et al. (2009)). Further, Langley and Hoyle (2008) showed that when the weighting assumed for the CN/TW-OS longline fisheries was reduced, the recruitment trend was reduced. We continue that line of examination here with models 4-7 where we reduce the weights to the size data of the important longline fisheries in regions 3 and 4.

Another issue regarding size data relates to the potential bias in the length frequency samples from the tropical purse seine fisheries. Lawson (2009) showed that there is strong evidence that the implementation of the grab sampling method to estimate the length composition of purse seine catches may result in over-estimation of the sizes of fish caught. Specifically it seems likely that bigeye under 50cm are underrepresented in the grab samples while bigeye over 50cm are over-represented (Figure 24). There are not yet sufficient data to formally revise the length frequency samples used in the assessment, but to get an indication of the possible consequence of this bias; we reduced the lengths of bigeye from the purse seine fishery by 4 cm in model 8.

### *Biology*

Difficulties in ageing large tropical tunas and the typical problems encountered in estimating natural mortality, especially age/length specific natural mortality, result in uncertainty in these processes in the stock assessment.

The bigeye model is sensitive to estimates of the sizes of the oldest individuals in the model and estimates of variability in length at age often appear quite small – and there is no ageing data and little information in the length data to estimate this. In model 9 we take the mean growth curve from model 1, but arbitrarily increase the variability in length at age (Figure 25). This curve is not necessarily more realistic, but is useful to examine the importance of more accurate estimation of the variability around the growth curve.

Models 10 and 11 relate to uncertainty in adult and juvenile natural mortality respectively. In model 10 natural mortality increases slightly at the older ages rather than declining; in model 11 we assume the same levels of natural mortality that are assumed for yellowfin tuna for the first three quarters (Langley et al. 2009) (Figure 26). Some biologists have previously questioned the large difference assumed in the natural mortality at the youngest ages for these two species.

In addition to these runs, we did have an initial attempt to look for evidence of different growth rates in the different regions (as is most likely the case). This was done by reducing the weight of the size data outside the region of interest. These results were very preliminary and require further consideration and therefore were not included here in detail.

### *CATCHES ESTIMATES FOR INDONESIA AND THE PHILIPPINES*

For several years the Scientific Committee has been increasingly aware of the importance of obtaining more accurate estimates of the catches from the fisheries in the Philippines and Indonesian region and operated by their flagged vessels. Several programmes have been initiated and some real progress has been made in collaboration with the fisheries officials in these countries. The 2009 assessment incorporated revised catches from the domestic fisheries of the Philippines and these were much smaller than the previous estimates and alleviated the previous problem of a large "step" in the catch levels (see Figure 11 in Harley et al. (2009)). However, the large increase in both longline and small-fish domestic catches for Indonesia remained unresolved. As the result of considerable effort over several years, we now have available refined time series of catches for both Indonesia and the Philippines. A comparison of the assumed catches used in the 2009 assessment and the new estimates are provided in Figure 27. We investigate the impact of these new catch estimates in a two-step process, in model 12 we incorporate the new catches for the small fish fisheries (fisheries 18 and 24) and in model 13 we include the revised catch estimates for the longline fisheries (fisheries 4 and 23).

### *CPUE / CATCHABILITY*

Estimated CPUE series for the key longline fleets and assumptions about catchability for both these and other fleets in the model provide the model with important information on relative abundance trends. Further there is also information on the distribution of bigeye biomass across regions through the assumption of shared catchabilities for the regional LL-ALL fisheries. In the current assessment we include standardized CPUE for the main LL-ALL fleets and assumed fixed catchability, while for all other fleets we allow time-varying catchability to reflect our lesser confidence that CPUE for these fleets provide information on relative abundance.

There have been dramatic changes in the purse seine fishery over the past ten years, even if you only consider the fishing effort related to FADs. There have been considerable advances in FAD technology with new FADs less vulnerable to detection and 'use' by others and the introduction of sonar technology on FADs which can provide some information on whether there are fish associated with a FAD. Further, many of the purse seine vessels are newer, larger, and more efficient than those in operation ten years ago. Some of these changes have been sudden, and none of them have been accounted for in the estimates of effort assumed in the model. In model 1 we assumed that purse seine catchability can change every two years and then amount of change was penalized with a CV of 0.4. It has been suggested that rapid increases in purse seine catchability at shorter time steps, has the potential be wrongly interpreted as increased recruitment rather than increased fishing efficiency. Three model runs were undertaken to investigate this issue. In model 14 we increased the frequency of catchability changes to every quarter. This results in over 400 more parameters to estimate (i.e. catchability deviates). In model 15 we kept the frequency of changes at two years, but increased the amount that catchability could change by increasing the CV to 0.7. This allows more freedom in catchability, but without extra parameters. Finally in model 15 we increased both the frequency of changes and the CV together.

In addition to deriving CPUE indices, the Japanese CPUE data are also used to estimate the relative biomass in each model region. This is described by Hoyle (2009) and essentially involves looking at the absolute levels of CPUE across the WCPO during the period of most consistent fishing effort (1960-86). One troubling aspect of this analysis is the relatively high weight estimated for region 2 compared to that for region 3 which is in the core of the fishery. This weighting, in combination with other data sources results in biomass estimates at the start of the model being higher in region 2 than region 3. In model 17 we arbitrarily examine the potential impacts of the weights being wrong by halving the weighting for region 2 and doubling it for region 3.

As noted in the exploration of the key data indicators, one of the troubling conflicts in the assessment data is the relatively stable CPUE in region 3 in the face of increasing catches. We were interested in how the model would respond if the region 3 CPUE series was declining through time instead and this was done in model 18. Rather than fabricating a CPUE series we simply took the yellowfin index that Hoyle (2009) calculated for the 2009 yellowfin assessment. The index was rescaled so that the regional weighting for region 3 was maintained[4]. The annualized indices for model 1 and model 18 are shown in Figure 28.

A potential bias in the CPUE series due to the inability to account for changes in fleet catchability due to changes in the composition of the fleet was estimated by Hoyle (2009). When this was incorporated into the 2009 bigeye assessment (run 15 of the 2009 assessment) the regional weighting was not maintained. Incorporating the increase in effective effort had the unintentional effect of reducing the regional weight for region 3. In model 19 we implement the 0.47% per year increase in region 3 effective effort while maintaining the regional weights. The difference in the effect on CPUE is shown in Figure 28.

### *COMBINATION RUNS*

The 2010 pre-assessment workshop indicated a desire for these investigations to consider some of these factors together to look at the combined effects. While a full cross of all possible combinations of the factors would do this fully, the computational demands and difficulty in interpretation meant that this approach has not been taken here. We have undertaken just three model runs here that combine some of the potentially important factors.

In model 20 we take the new ID/PH data (model 13), the spill sampling 'corrected' purse seine length frequency data (model 8) and downweight the size data for fisheries 5 and 8 (model 5). In model 21 we also add the increased juvenile natural mortality to the changes in model 20. The examination of size data indicated the potential for non-Japanese data to be introducing bias into the trends in median size in some of the LL-ALL fisheries. In model 22 we take the new ID/PH catch estimates and exclude all non-Japanese size data from fisheries 1, 2, 4, and 7. Model 23 builds on

[4] The weighting was incorrectly calculated over the period 1976-86 instead of 1960-86, however, upon further examination this made little difference.

model 22 with the additional feature that the last five years size data from fishery 5 are excluded on the basis of the uncertainty over those data noted previously.

## MODEL OUTPUT INDICATORS

There are several indicators that we choose to calculate in order to simplify the comparisons across the twenty-three model runs and these are provided below.

The first set of indicators relate to average levels of overall recruitment in the early and later part of the model time series. We also visually inspected some of the regional recruitment trends, in particular regions 3 and 4.

- $\bar{R}_{1952-1980}$
- $\bar{R}_{1981-2008}$
- $\bar{R}_{1981-2008}/\bar{R}_{1952-1980}$

The 2009 assessment displayed a conflict between recent size data and the CPUE indices for the LL-ALL fleet in region 3. This was evident in the strong trend in the effort deviates, with the model estimating negative deviates since 1990. The slope of the effort deviates for the period 1970-2008 was calculated as one model indicator.

We also considered some of the standard reference point outputs and biological parameters:

- $MSY$
- $C_{current}/MSY$: average annual catch over a recent period relative to $MSY$
- $F_{current}/F_{MSY}$: average fishing mortality-at-age for a recent period relative to $F_{MSY}$
- $SB_{current}/SB_{MSY}$: average spawning biomass for a recent period relative to $SB_{MSY}$
- $B_0$: equilibrium unexploited total biomass
- $h$: the estimated steepness of the Beverton-Holt spawner recruitment curve
- $L_{max}$: estimated length at age 40 quarters

We also examined estimated selectivity curves, growth, curves, and residual plots of fits to size frequency data for many of the model runs.

## MODEL RESULTS AND DISCUSSION

Estimates of the key performance indicators for all twenty-three model runs are provided in Table 4, the annual recruitment estimates in Figure 29, and the effort deviates for fishery 4 (LL-ALL in region 3) in Figure 30. We describe the results first by group of sensitivity analysis and then provide further analysis of a restricted set of model runs.

The "Tidy-up" model runs (2-3) had little impact on the important performance indicators, though the exclusion of the late 1950s Japanese size data did slightly reduce the recruitment trend by slightly increasing early recruitment and also slightly reduced the trend in effort deviates. Further examination of the Japanese size data in the early years of the fishery is warranted to determine if they are being interpreted / modeled correctly at present.

All the "size" runs had considerable impact on the key performance indicators, with all model runs resulting in improvements in the recruitment and effort deviate trends as well as estimated stock status. Even though these data are only available since the 1990s, the data for both CN/TW-OS fleets are having considerable influence over the stock assessment. They are suggesting smaller maximum sizes and much lower recruitment during the first half of the time series (e.g. the recruitment that occurs prior to the fishery commencing). Down-weighting these data leads to considerable improvement in the recruitment and effort deviate trends and stock status. The improvement in stock status is even more remarkable given that this model estimates a lower value for steepness. It is interesting that their impact on the latter recruitment estimates in much less than the impact on the earlier part of the series. The impact is greater than that for the LL-ALL size data suggesting something specific to these data.

Elsewhere in this paper we have mentioned that there appears to be strong spatial patterns in the sizes of fish taken in the CN/TW-OS fisheries and also that there are many fleets that could comprise this fishery and that there are likely changes through time in these fleets fishing strategies. While this data set will provide considerable information to inform future assessments, given these current uncertainties, it is a concern that these data are having such an impact on the assessment. It is recommended that the size data for the CN/TW-OS be down-weighted until they can be better understood and modeled appropriately.

The spill sample length frequency "correction" led to a significant improvement in the key performance indicators. It increased early recruitment and decreased late recruitment, but there was little impact on the estimated growth curve. Fishing mortality is decreased 15% by this change alone. Quantification of the potential bias in the length frequency samples taken by the grab sampling procedure should be a high priority for further field work and analysis.

The alternative biological assumptions made in "biology" runs were not that influential. There was little effect on the estimated MSY, but some improvements, mostly small, in the performance indicators. Increasing the variance in length at age had the greatest reduction in the recruitment trend and slope. Given the likelihood that it is likely that MULTIFAN-CL is not able to accurately estimate the length, and variation in length, at the oldest ages it is recommended that research be directed in this area for tropical tunas. Until such estimates are available, a meta-analysis of the relationship between lengths-at-age and associated variation be undertaken for those easier to age species (e.g. other teleosts).

The increase of natural mortality of the older fish had very little impact on any of the indicators. Unless there are plausible curves which have considerably higher natural mortality than that assumed in model 10, this is lower priority research area at this time. Of the biological model runs, this increase juvenile mortality run (model 11) had the greatest reduction in fishing mortality. It had a moderate impact on the recruitment trend, but little impact on the trend in effort deviates. Further consideration of the relative levels of natural mortality of juvenile bigeye and yellowfin tuna is warranted given the large differences currently assumed.

Uncertainty in the catches from the fisheries of Indonesia and the Philippines has been of concern for several years and considerable progress is being made in this area and the "IDPH" model runs presented here indicate that the stock assessment are sensitive to these data. The changes examined here mostly relate to changes to the catch estimates of Indonesia and the large changes to the Philippines estimates occurred last year. The revised domestic small-fish catches (model 12) resulted in a 10% reduction in fishing mortality and a 10% reduction in recent average recruitment. This change had little impact on the effort deviate trend or growth. When combined with the revised longline catches estimates (model 13) this resulted in a further reduction in recent recruitment and the recruitment ratio. The trend in effort deviates was greatly reduced suggesting that the incorrectly high longline catches for Indonesia were causing a conflict in the model. There was a 17% decline in fishing mortality compared to model 1. Given the impact on the assessment, continuation of the good progress being made in refining the estimates of catches from the fisheries of Indonesia and the Philippines would be one of the higher priorities for future efforts.

Several model runs were undertaken in the "CPUE/catchability" group and the effect of these varied. The evaluation of alternative assumptions regarding purse seine catchability had next to no impact on any of the key performance indicators. Under models 1 and 14, there were large positive effort deviates estimated for 2009. Allowing larger jumps (model 15) reduced this somewhat, but it was only when the two were combined that the effort deviate pattern was more reasonable (model 16). While these patterns, and the potential fixes, had little impact on the performance indicators, they might have an impact when projections are undertaken as it is the catchability estimate (without the effort deviate) which is used in the projections. Of course there is considerable uncertainty in the estimates of the effort deviates in the final year (due to correlation with the estimates of recruitment) so it is not clear if this is in fact a problem. It is recommended that the frequency and magnitude of purse seine catchability changes be examined in the context of undertaking projections, but at a minimum the CV on catchability changes for the purse seine fishery be increased to 0.7 for the current assessment.

Models 17 and 18 involved the fabrication of the longline CPUE indices to test the sensitivity of the assessment to the current indices. The doubling of the assumed relative abundance in region 3 and halving that in region 2 led to a far more optimistic assessment and improvements in the recruitment and effort deviate trends. The current regional weighting analyses suggest that biomass in region 2 is about two-thirds that in region 3 for the period 1975-1986. With the CPUE trend for region two, this results in initial biomass levels in region 2 being much greater than the core part of the fishery. The model results here suggest that the current estimates of regional weightings may not be appropriate and introducing bias into the assessment so further work is required in this area, perhaps using methods that do not rely on CPUE data.

Introduction of a CPUE trend with a continual decline had a dramatic impact on the stock assessment (model 18). This run provided similar improvements in stock status to model 5 (down-weighted CN/TW size data), but additionally resulted in the most significant reduction (or flattening) of the recruitment trend, of any of the model runs. It reduced the conflict between the size and CPUE data as seen in the estimated effort deviates. It did this without increasing the maximum length as seen in model 5. The remarkable stability of the LL CPUE in the face of increasing catches was a feature discussed earlier in the paper. Either the catch or CPUE series for region 3 is significantly incorrect with respect to their trends and the CPUE indices are the most likely candidate. Further research into the estimation of abundance indices from longline data should be the highest priority for the assessment. The initial collaborations on the analysis of operational level data is exciting and should be continued, but the types of CPUE analyses being undertaken need to be expanded to approaches that better account for the large changes in the spatial coverage of the fishery that is evident over the time series (see Harley 2009 and Hoyle et al. 2010 for details of the expansion and contraction of the Japanese longline fleet).

The adjustment for increases in efficiency due to fleet composition had only a minor impact on any of the key performance indicators, but as noted above, there are renewed questions over whether the abundance indices themselves are an index of abundance.

It was hoped that the model runs with combinations of the individual changes would assist in identifying necessary improvements to the assessment which might result in increased plausibility of the performance indicators, but this was not the case. When compared to model 5, the estimates of fishing mortality from models 20 and 21 were comparable, but the improvement to the recruitment and effort deviate trends were not as good. The exclusion of the non-JP size data generally made things worse rather than better (model 22) which isn't entirely surprising given some of the variable patterns within the JP data (e.g. region 3). Excluding the recent fishery 5 size data during the period of large spatial differences in the locations of the fleets did improve things a little (model 23).

We then examined the residuals from the fit to the length frequency data for fishery 4 (LL-ALL in region 3) to see if any of the models had resulted in an improvement in the residual patterns (Figure 32). Unfortunately the patterns seen in model 1 persisted, and only when the non-Japan data were excluded did we see some improvement in the fit, but only a reduction in the large residuals for large fish at the end of the time period which were due to wrongly assigned fleets.

While we had focused on the overall recruitment trend we also looked at the trend for region 3 for some of the models which had the greatest impact on the overall recruitment trend (Figure 33). While model 5 resulted in a small flattening in the region 3 recruitment trend, it was only the model which included the YFT CPUE index (model 18) which resulted in a very significant flattening of the region 3 recruitment trend. This reinforces the hypothesis that it is the mismatch between the catch series and the CPUE which is one of the major conflicts in the current assessment.

## CONCLUSIONS

This paper describes a detailed analysis of the input data used in the 2009 bigeye assessment and a wide-range of sensitivity analyses to see what might be causing some of the main conflicts in the assessment such as the strong positive recruitment trends.

A recurring theme throughout the paper is the apparent conflict in region 3 between the longline CPUE series and the catch history. The stability of this CPUE in the face of increasing catches is causing problems for the assessment. We have discovered that some of the fleets for which some size data exist have been wrongly assigned to fisheries and this was causing some trends in the size data for fisheries 4 and 7 (LL-ALL in regions 3 and 4). New fisheries definitions are proposed to address this. We also discovered some errors in the size database itself with some samples being allocated to the wrong area – these have been corrected and will no longer be an issue for the 2010 bigeye assessment. With these two changes made, there are still some troubling patterns in the Japanese length frequency data that warrant further attention. In particular, the length data collected during the period 1955-1965 show large proportions of small fish in the catch. The Japanese size data for region 3 contain troubling 'steps' that appear to be due to spatial structure in the size of fish taken and non-representative sampling. There are also some strong trends in the region three China / TW-Offshore size data which appear to be attributable to changes in the locations where the fishery is operating. These both require immediate examination.

There are numerous recommendations throughout this paper (underlined) which will form the basis for the 2010 bigeye assessment and the direction of research activities for the short to medium term.

## References


Harley, S.J., Hoyle, S.D., Langley, A., Hampton, J., and Kleiber, P. 2009. Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC SC5 SA-WP-4. Port Vila, Vanuatu, 10-21 August 2009.

Hoyle, S. 2009. Standardized CPUE for bigeye and yellowfin tuna. WCPFC-SC5 SA-WP-01. Port Vila, Vanuatu, 10-21 August 2009.

Langley, A., Hoyle, S. 2008. Report from the stock assessment preparatory workshop, Noumea, February 2008. WCPFC SC4 SA IP-5. Port Moresby, Papua New Guinea, 11-22 August 2008.

Langley, A., Harley, S.J., Hoyle, S.D., Davies, N., Hampton, J., and Kleiber, P. 2009. Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC SC5 SA-WP-3. Port Vila, Vanuatu, 10-21 August 2009.

Lawson, T. 2009. Selectivity bias in grab samples and other factors affecting the analysis of species composition data collected by observers on purse seiners in the Western and Central Pacific Ocean. WCPFC SC5 ST WP-3, Port Vila, Vanuatu, 10-21 August 2009.

**Table 1: Fishery definitions used in the 2008 and 2009 WCPO bigeye assessments.**

| Fishery | Code | Nationality | Gear | Region |
|---|---|---|---|---|
| 1 | LL ALL 1 | Japan, Korea, Chinese Taipei | Longline | 1 |
| 2 | LL ALL 2 | Japan, Korea, Chinese Taipei | Longline | 2 |
| 3 | LL HW 2 | United States (Hawaii) | Longline | 2 |
| 4 | LL ALL 3 | All excl. Chinese Taipei & China | Longline | 3 |
| 5 | LL TW-CH 3 | Chinese Taipei and China | Longline | 3 |
| 6 | LL PG 3 | Papua New Guinea | Longline | 4 |
| 7 | LL ALL 4 | Japan, Korea | Longline | 4 |
| 8 | LL TW-CH 4 | Chinese Taipei and China | Longline | 4 |
| 9 | LL HW 4 | United States (Hawaii) | Longline | 4 |
| 10 | LL ALL 5 | All excl. Australia | Longline | 5 |
| 11 | LL AU 5 | Australia | Longline | 5 |
| 12 | LL ALL6 | Japan, Korea, Chinese Taipei | Longline | 6 |
| 13 | LL PI 6 | Pacific Island Countries/Territories | Longline | 6 |
| 14 | PS ASS 3 | All | Purse seine, log/FAD sets | 3 |
| 15 | PS UNS 3 | All | Purse seine, school sets | 3 |
| 16 | PS ASS 4 | All | Purse seine, log/FAD sets | 4 |
| 17 | PS UNS 4 | All | Purse seine, school sets | 4 |
| 18 | PH MISC 3 | Philippines | Miscellaneous (small fish) | 3 |
| 19 | PH HL 3 | Philippines, Indonesia | Handline (large fish) | 3 |
| 20 | PS JP 1 | Japan | Purse seine | 1 |
| 21 | PL JP 1 | Japan | Pole-and-line | 1 |
| 22 | PL ALL 3 | Japan, Solomons, PNG | Pole-and-line | 3 |
| 23 | LL BMK 3 | All, excluding PNG | Longline, Bismarck Sea | 3 |
| 24 | ID MISC 3 | Indonesia | Miscellaneous (small fish) | 3 |
| 25 | HL HW 4 | United States (Hawaii) | Handline | 4 |

**Table 2: Current status of DWFN longline catch, effort and size data holdings as used in the stock assessments.**

| Flag | Aggregate Catch/Effort data | Length | Weight |
|---|---|---|---|
| China | Provided by China.<br>Data provided do not distinguish between offshore and distant-water fleets, so an attempt has been made to separate these data using 180° for the tropical fishery and assuming the vessels in Fiji are offshore. Single estimate of catch provided.<br>Issue of off-shore fleet catches and charter arrangements etc. remains | Bigeye length data recently provided for 2009 only for DWFN fleet (2cm intervals – yet to be reviewed and imported).<br>Longer time series of length data available for the offshore fleet based on Micronesia (Palau, FSM, RMI), but weight data from Guam (for example) more comprehensive. | No weight data provided for DWFN fleet.<br>Weight data available for the offshore fleet based on Micronesia (Guam, Palau, FSM, RMI).<br>For 2009, Luenthai have provided all weight data (100%) for vessels landing in Palau, FSM and RMI. |
| Japan | Provided by Japan.<br>No differentiation between DW and OS in aggregate data.<br>No aggregate catch/effort data available for Coastal fleet. | Confirmation on source of data has been sought from Japan. We understand that some of the data come from training vessels (DW class) and not necessarily in the same areas as the fishery. | From fishing vessels and split into DW and OS components – from JP? |
| Korea | Provided by Korea<br>No offshore fleet - only DW vessels. | Only DW vessels<br>Usable Bigeye length data provided for recent years only (2006-2009) | No weight frequency data |
| Chinese Taipei | Catches split into DW, OS/PIC and OS/domestic components.<br>DW and OS/domestic provided by Chinese Taipei and OS/PIC provided by PIC countries (thru SPC) where these vessels are based.<br>Issue of charter arrangements. | Length samples split, but some data without fleet assumed to be DW.<br>Usable data from Chinese Taipei cover DW and OS/domestic – essentially recent years only (2005-2009).<br>Longer time series of OS/PIC fleet data provided thru PIC monitoring programmes (1991-2009) | No weight data provided for DW fleet but DW and OS weight data collected at PIC ports of unloading.<br>Weight data available for the offshore fleet based on Micronesia (Guam, Palau, FSM, RMI) |

**Table 3: Summary of model runs undertaken using the spill sampling run from the 2009 bigeye assessment as a starting (run 14). Unless otherwise stated, the models represent one-off changes from model1.**

| Model | Code | Description | Type |
|---|---|---|---|
| 1 | Run 14 | Run 14 from the 2009 assessment | Base run |
| 2 | Run 14g | Exclude fishing events with low effort and zero catch for fisheries 6, 8, 11, 14, 15, 16, and 17 | Tidy-up |
| 3 | Run 14k | Exclude Japanese length frequency data for the period 1955-1965 | Tidy-up |
| 4 | Run 14e | Downweight fishery 5 (CN and TW-OS in region 3) size data | Size-data |
| 5 | Run 14e2 | Downweight fishery 5 and 8 (CN and TW-OS in region 4) size data | Size-data |
| 6 | Run 14l | Downweight fishery 4 (LL-ALL in region 3) size data | Size-data |
| 7 | Run 14l2 | Downweight fishery 4 and 7 (LL-ALL in region 4) size data | Size-data |
| 8 | Run 14i | Reduce sizes of purse seine length samples by 4cm to attempt to approximate the nature of the grab sampling bias | Size-data |
| 9 | Run 14d | Increase variability around the growth curve from model 1 | Biology |
| 10 | Run 14f | Increase natural mortality for the older ages | Biology |
| 11 | Run 14f2 | Increase natural mortality for juvenile bigeye to make it consistent with the estimates for yellowfin tuna | Biology |
| 12 | Run 14A | Revised catches for the Indonesian and Philippine domestic "small-fish" fisheries (fisheries 18 and 24) | IDPH catch |
| 13 | Run 14B | As in model 12 plus revised longline catches for Indonesia and the Philippines (fisheries 4, 19, and 23) | IDPH catch |
| 14 | Run 14a | Catchability for tropical purse seine fisheries (fisheries 14-17) can change every quarter rather than every two years | CPUE / catchability |
| 15 | Run 14b | Reduced penalty on the amount that purse seine catchability can change every two years (CV increased from 0.4 to 0.7) | CPUE / catchability |
| 16 | Run 14b2 | Time period for catchability changes as in model 14, but with the CV from model 15 | CPUE / catchability |
| 17 | Run 14h | Half the regional weight for region 2 and double that in region 3 | CPUE / catchability |
| 18 | Run 14j | Replace the non-declining BET CPUE for region three with the YFT index which has a consistent decline – regional weights retained | CPUE / catchability |
| 19 | Run 14n | Incorporate the Hoyle (2009) estimate of a 0.47% per year adjustment of CPUE in region 3 for fleet composition – regional weights retained | CPUE / catchability |
| 20 | Run 14o | New ID/PH data (from model 13) plus the "spill sample" length frequency data for the purse seine fisheries (model 8) and down-weighted size data for fisheries 5 and 8 (from model 5) | Combination |
| 21 | Run 14p | Model 20 plus the increased natural mortality for juvenile bigeye (model 11) | Combination |
| 22 | Run 14C | New ID/PH data and non-Japan size data excluded for fisheries 1, 2, 4, and 7 | Combination |
| 23 | Run 14D | Model 22 and with the last five years of size data from fishery 5 excluded | Combination |

**Table 4: Estimates of key performance indicators from the model runs described in Table 3.**

| | $MSY$ (MT) | $C_{current}$ $/MSY$ | $F_{current}$ $/F_{MSY}$ | $B_0$ | $SB_{current}$ $/SB_{MSY}$ | Steep. ($h$) | $L_{MAX}$ | Early Rec. | Late Rec. | Rec. ratio | Effdev. slope |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Model-1: run14 | 67,800 | 2.57 | 2.01 | 927,800 | 1.04 | 0.99 | 179.60 | 13.24 | 33.22 | 2.51 | -1.69 |
| Model-2: run14g | 67,840 | 2.57 | 1.99 | 930,400 | 1.07 | 0.99 | 178.60 | 13.43 | 33.33 | 2.48 | -1.64 |
| Model-3: run14k | 67,840 | 2.57 | 1.95 | 928,300 | 1.09 | 0.99 | 177.74 | 13.66 | 33.22 | 2.43 | -1.56 |
| Model-4: run14e | 74,480 | 2.35 | 1.72 | 1,061,000 | 1.19 | 0.98 | 186.12 | 16.41 | 32.89 | 2.00 | -0.75 |
| Model-5: run14e2 | 84,480 | 2.05 | 1.35 | 1,178,000 | 1.58 | 0.96 | 186.28 | 20.08 | 32.97 | 1.64 | -0.41 |
| Model-6: run14l | 73,280 | 2.36 | 1.91 | 913,600 | 1.00 | 0.99 | 180.01 | 13.09 | 31.10 | 2.38 | -0.99 |
| Model-7: run14l2 | 81,320 | 2.14 | 1.80 | 1,001,000 | 1.05 | 0.98 | 186.83 | 14.91 | 29.61 | 1.99 | -0.73 |
| Model-8: run14i | 73,000 | 2.39 | 1.71 | 949,800 | 1.45 | 0.97 | 177.46 | 14.78 | 31.44 | 2.13 | -1.13 |
| Model-9: run14d | 68,800 | 2.59 | 1.93 | 963,700 | 1.17 | 0.98 | 179.60 | 14.51 | 33.05 | 2.28 | -1.44 |
| Model-10: run14f | 67,760 | 2.57 | 1.97 | 914,100 | 1.08 | 0.99 | 178.61 | 13.49 | 33.25 | 2.47 | -1.65 |
| Model-11: run14f2 | 68,680 | 2.54 | 1.88 | 912,300 | 1.15 | 0.99 | 177.43 | 26.62 | 62.95 | 2.36 | -1.60 |
| Model-12: run14A | 69,120 | 2.31 | 1.81 | 892,400 | 1.13 | 0.99 | 178.50 | 13.33 | 30.61 | 2.30 | -1.64 |
| Model-13: run14B | 70,200 | 2.19 | 1.66 | 883,000 | 1.25 | 0.98 | 178.24 | 13.79 | 29.04 | 2.11 | -1.62 |
| Model-14: run14a | 67,760 | 2.57 | 2.01 | 927,500 | 1.04 | 0.99 | 179.60 | 13.22 | 33.24 | 2.51 | -1.71 |
| Model-15: run14b | 68,080 | 2.56 | 2.00 | 928,000 | 1.04 | 0.99 | 179.60 | 13.25 | 33.19 | 2.50 | -1.68 |
| Model-16: run14b2 | 68,080 | 2.56 | 2.00 | 927,200 | 1.04 | 0.99 | 179.60 | 13.24 | 33.18 | 2.51 | -1.70 |
| Model-17: run14h | 79,560 | 2.19 | 1.48 | 934,100 | 1.57 | 0.96 | 178.89 | 15.70 | 28.01 | 1.78 | -1.03 |
| Model-18: run14j | 86,880 | 2.04 | 1.35 | 1,199,000 | 1.53 | 0.96 | 175.31 | 23.48 | 32.52 | 1.39 | 0.19 |
| Model-19: run14n | 68,640 | 2.54 | 1.98 | 951,600 | 1.06 | 0.99 | 178.51 | 13.98 | 33.62 | 2.41 | -1.61 |
| Model-20: run14o | 76,200 | 1.99 | 1.37 | 1,055,000 | 1.47 | 0.97 | 186.67 | 17.30 | 30.36 | 1.76 | -0.51 |
| Model-21: run14p | 75,920 | 2.00 | 1.34 | 1,032,000 | 1.51 | 0.97 | 185.50 | 33.54 | 58.60 | 1.75 | -0.54 |
| Model-22: run14C | 67,760 | 2.26 | 1.80 | 873,500 | 1.08 | 0.99 | 178.44 | 13.19 | 29.74 | 2.25 | -1.30 |
| Model-23:run14D | 71,360 | 2.14 | 1.72 | 879,400 | 1.08 | 0.98 | 180.21 | 13.57 | 28.64 | 2.11 | -1.15 |

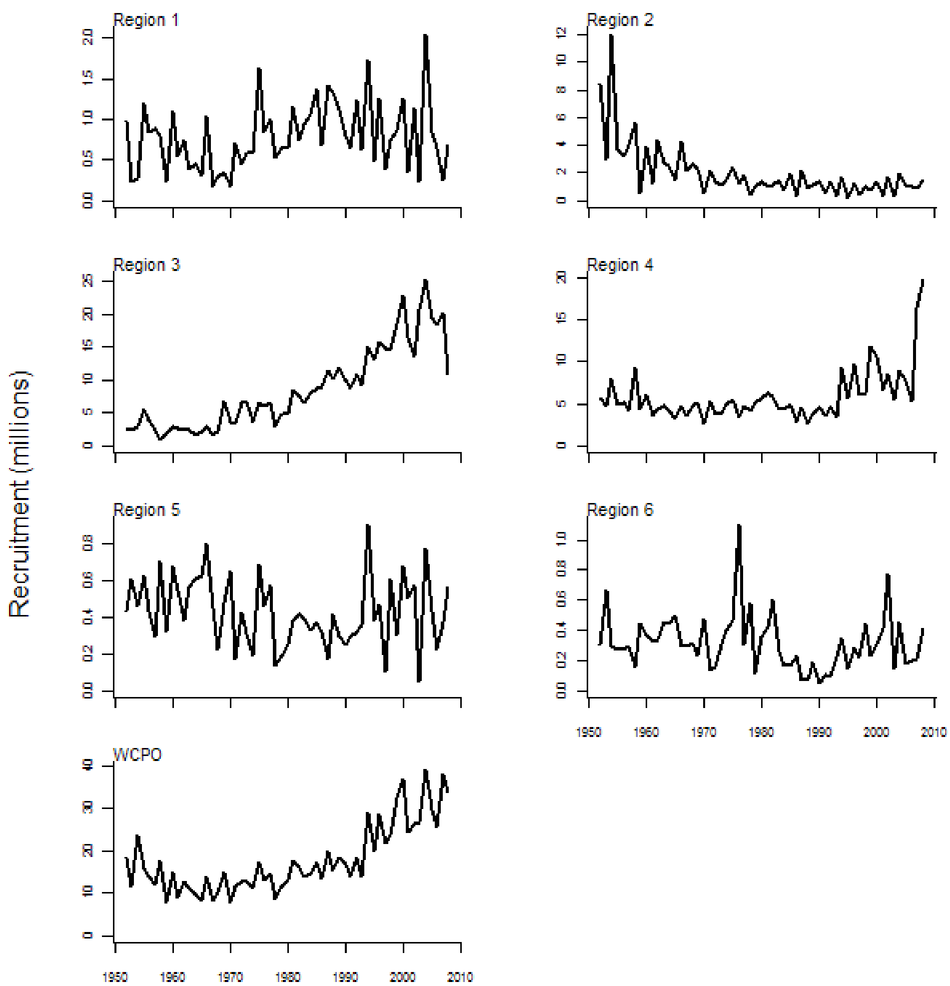


**Figure 1: Regional and overall estimates of recruitment for bigeye tuna in the WCPO from run 10 from the 2009 assessment of Harley et al. (2009).**

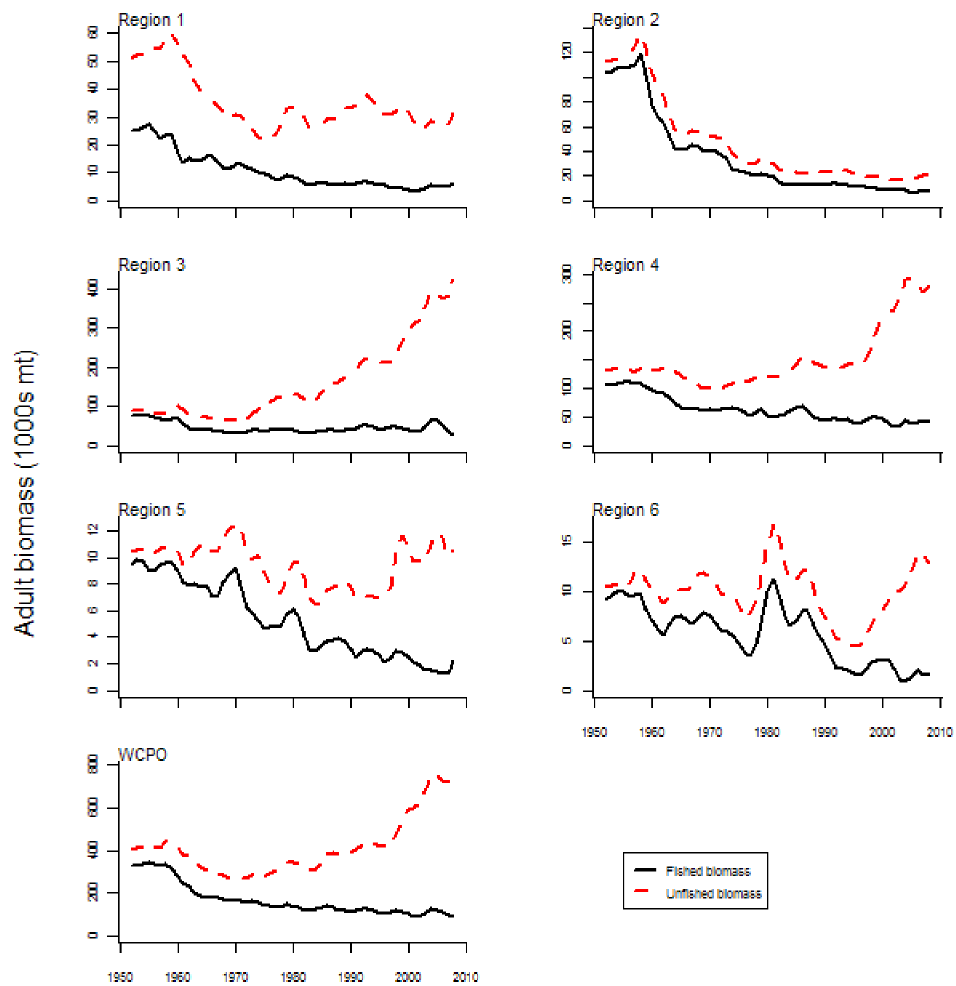


**Figure 2: Regional and overall estimates of fished and unfished spawning biomass for bigeye tuna in the WCPO from run 10 from the 2009 assessment of Harley et al. (2009).**

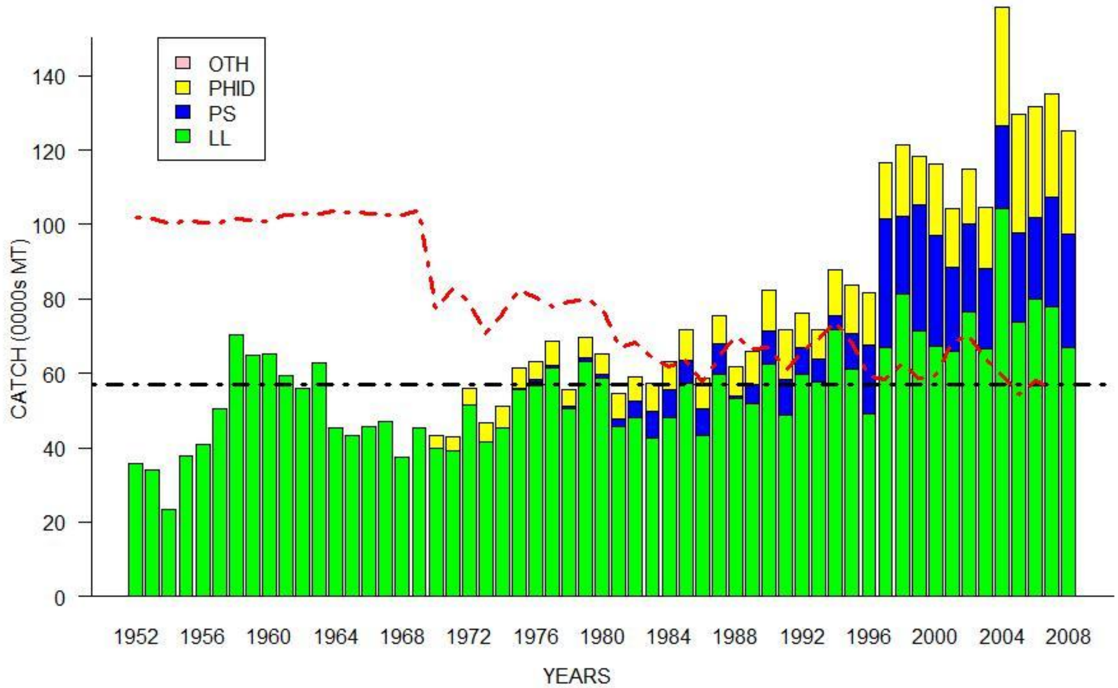


**Figure 3: Annual catches for bigeye tuna from the WCPO with estimates of deterministic (black dashed line) and dynamic (red dashed line) MSY overlaid. From run 10 from the 2009 assessment of Harley et al. (2009).**

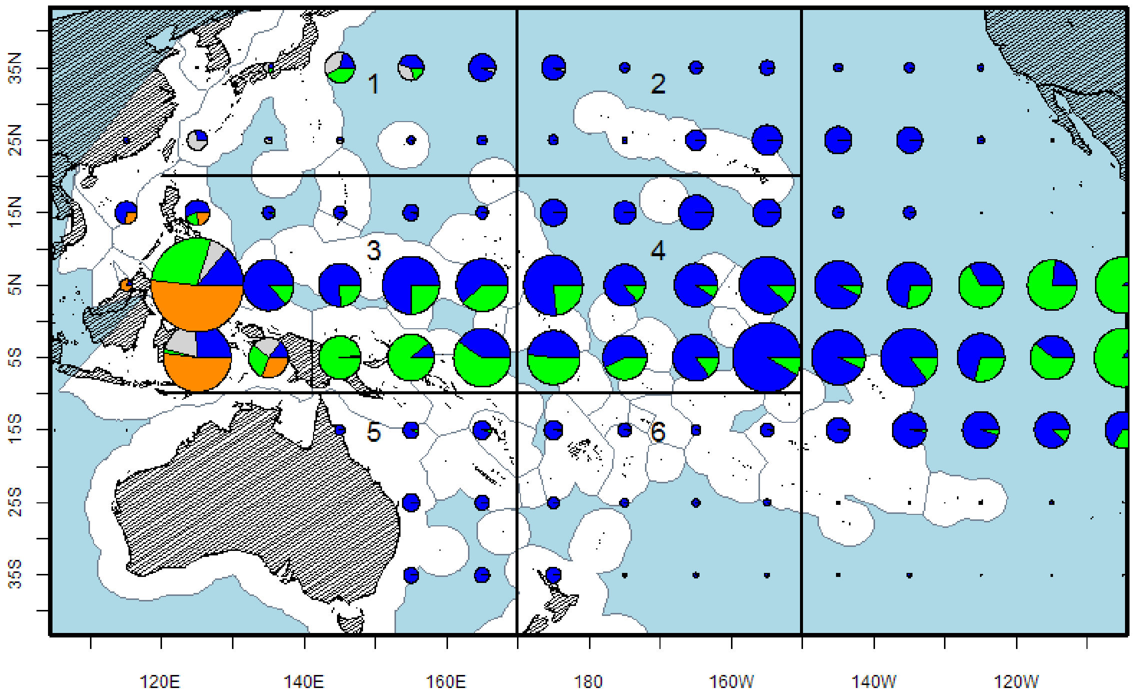


**Figure 4: Regions used in the 2008 and 2009 bigeye assessments for defining fisheries. The pies show the distribution of cumulative bigeye tuna catch from 1998-2007 by 10 degree squares of latitude and longitude and fishing gear; longline (blue), purse-seine (green), pole-and-line (grey) and other (dark orange). The maximum circle size represents a catch of 75,000 mt.**

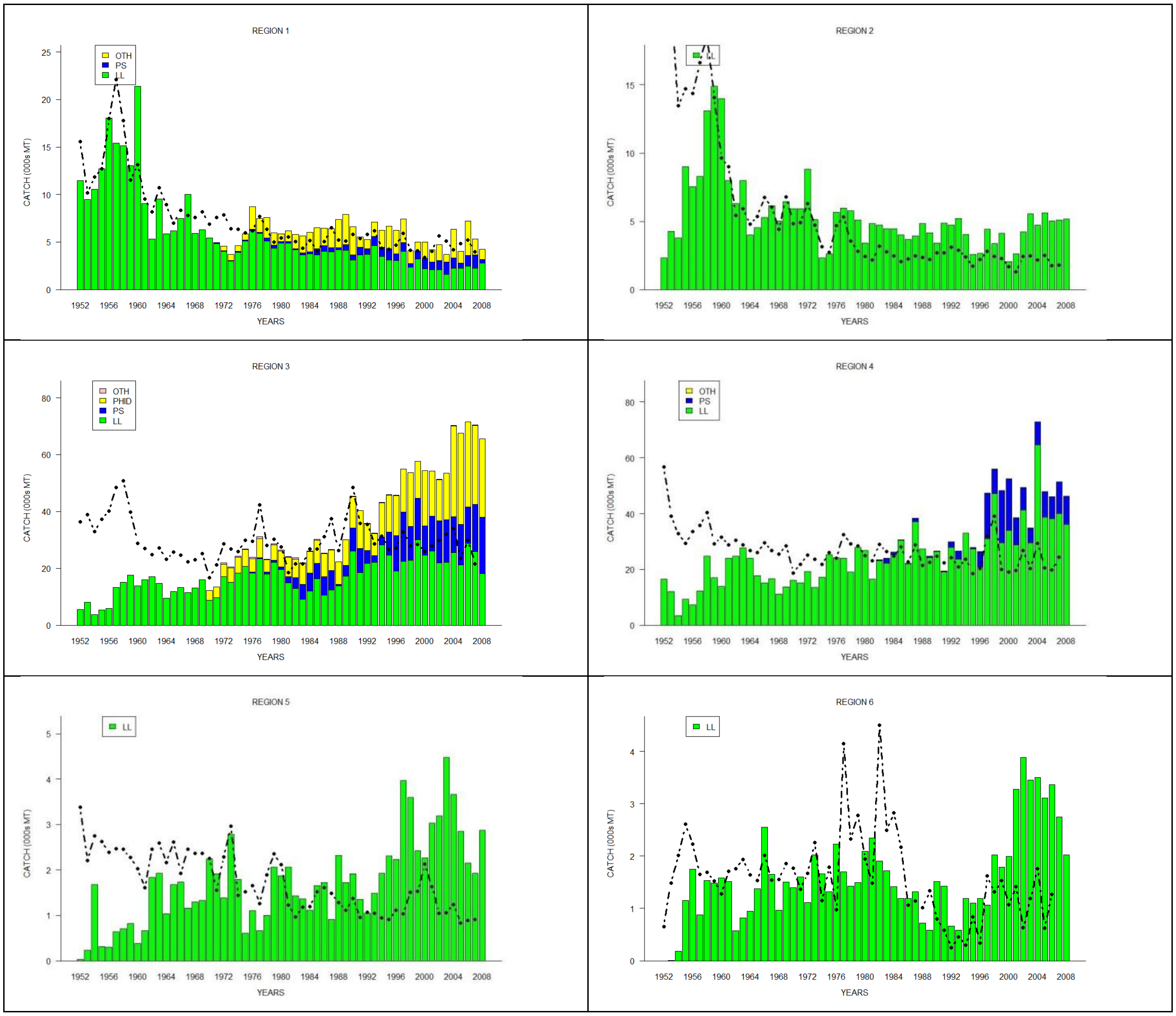


**Figure 5: Total BET catches by gear category and MFCL model region with the LL-ALL CPUE series overlaid.**

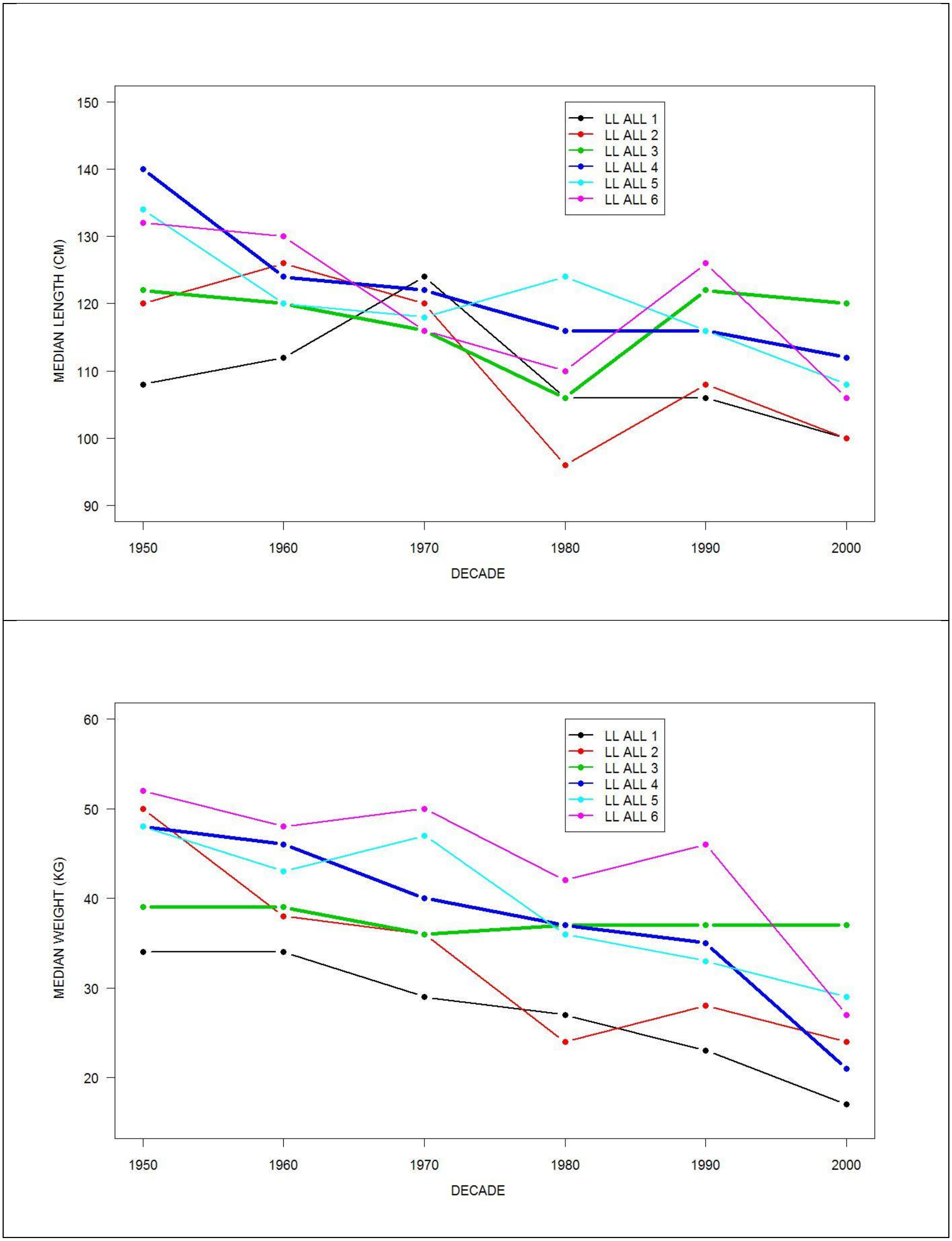


**Figure 6: Trends in median lengths (top) and weights (bottom) by decade for the LL-ALL regional fisheries.**

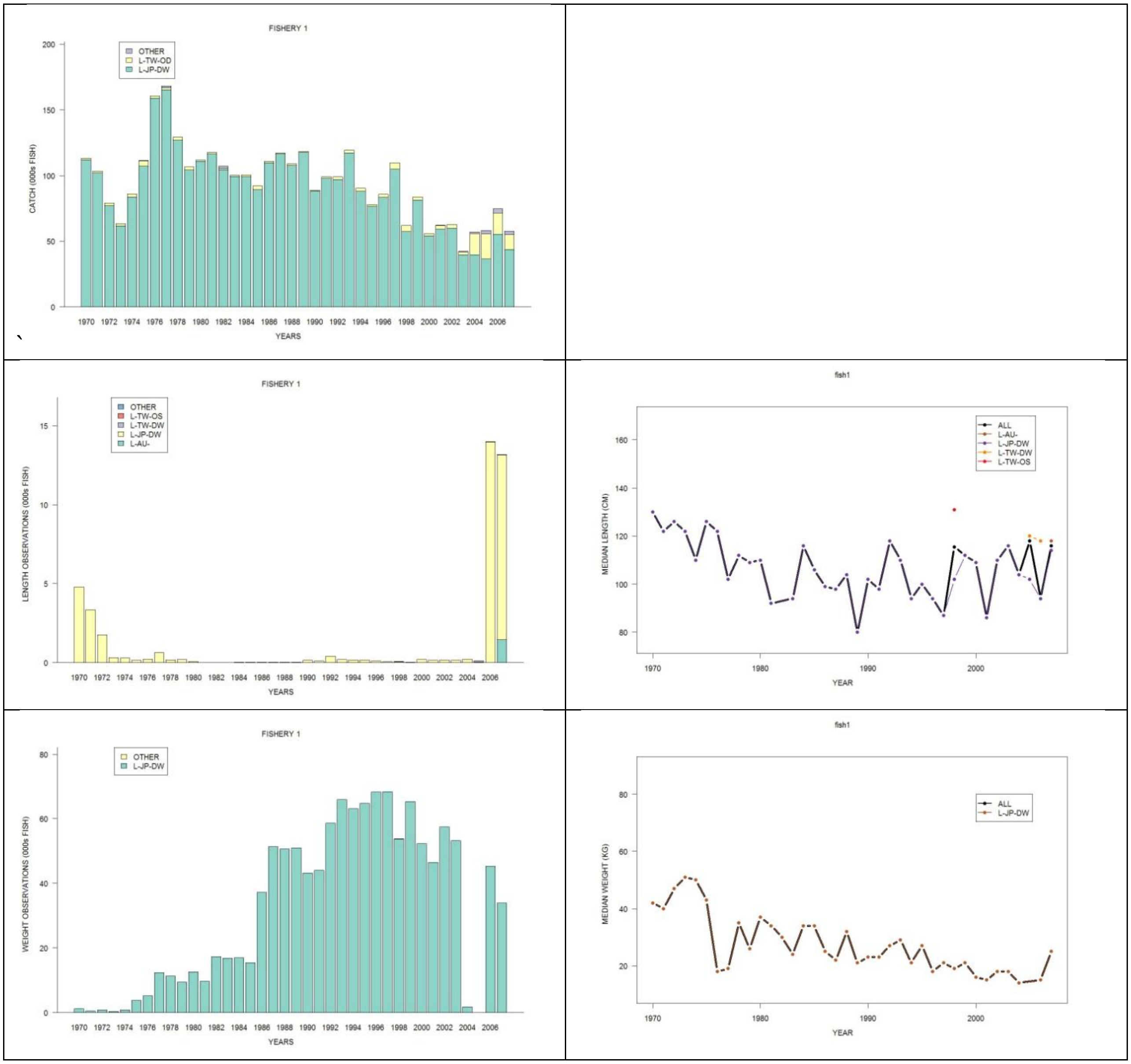


**Figure 7: Annual summaries of catch and size frequency data for Fishery 1 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**


**Figure 8: Annual summaries of catch and size frequency data for Fishery 2 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

**Figure 9: Annual summaries of catch and size frequency data for Fishery 4 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

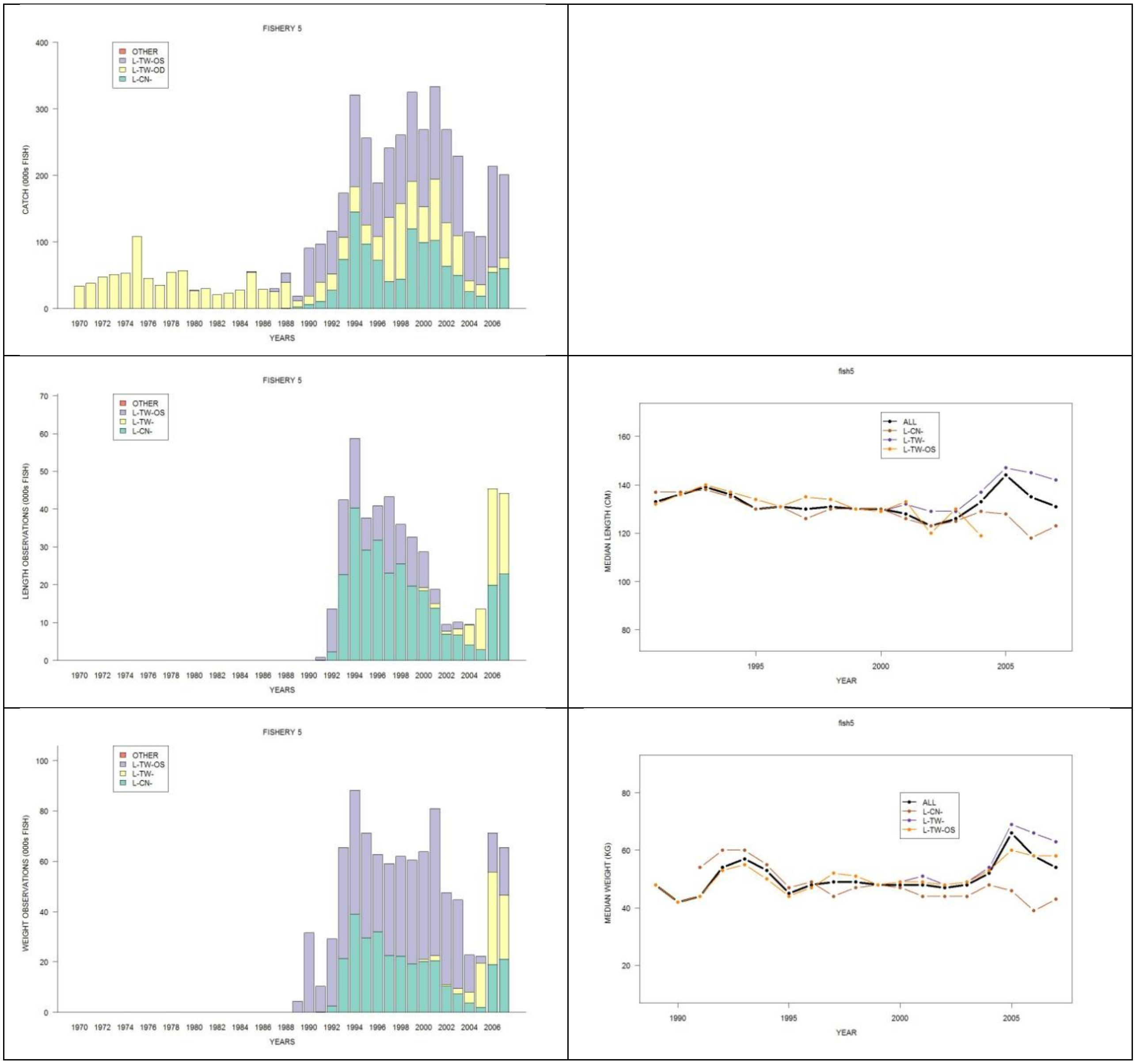


**Figure 10: Annual summaries of catch and size frequency data for Fishery 5 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**


**Figure 11: Annual summaries of catch and size frequency data for Fishery 7 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

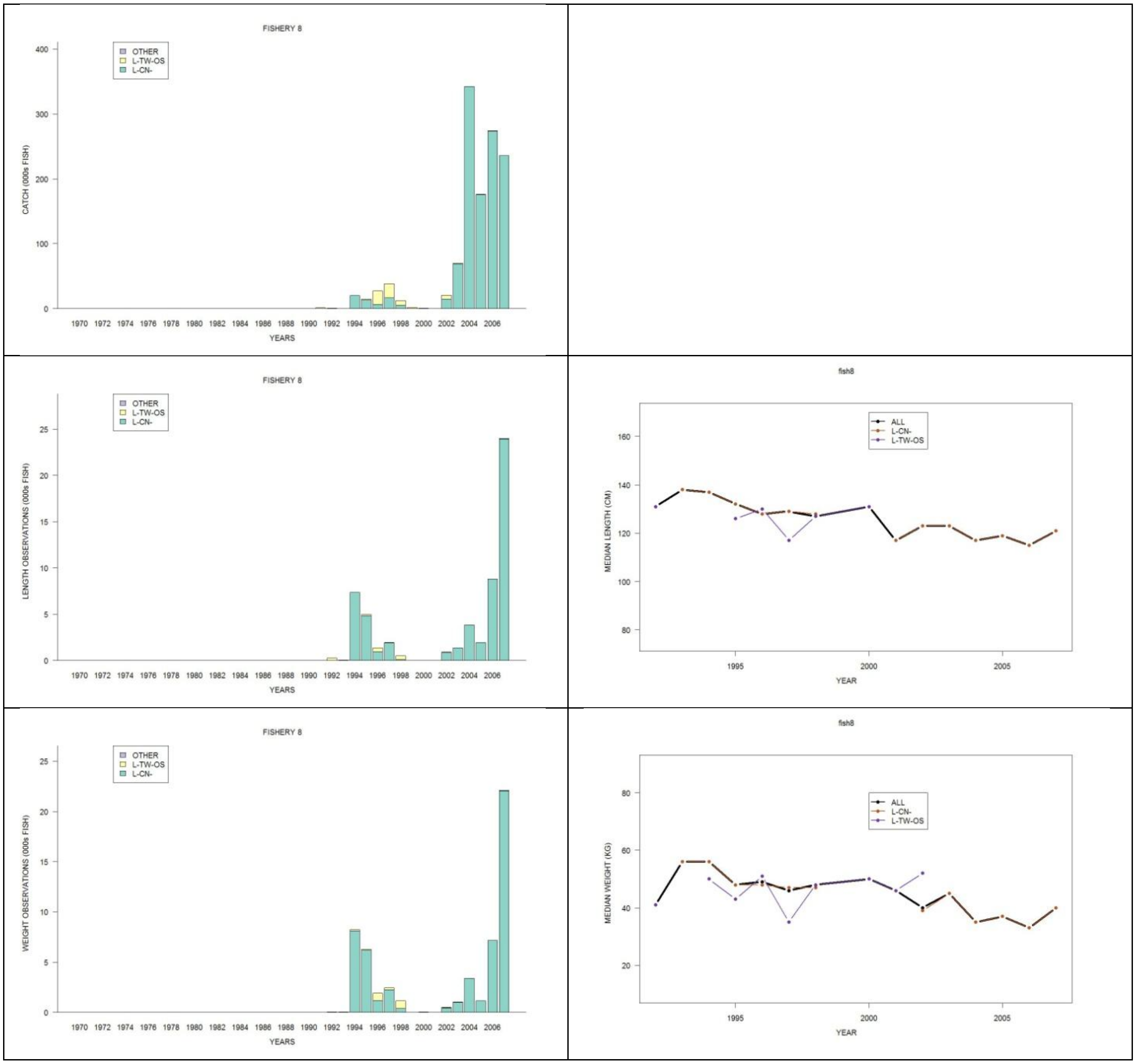


**Figure 12: Annual summaries of catch and size frequency data for Fishery 8 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

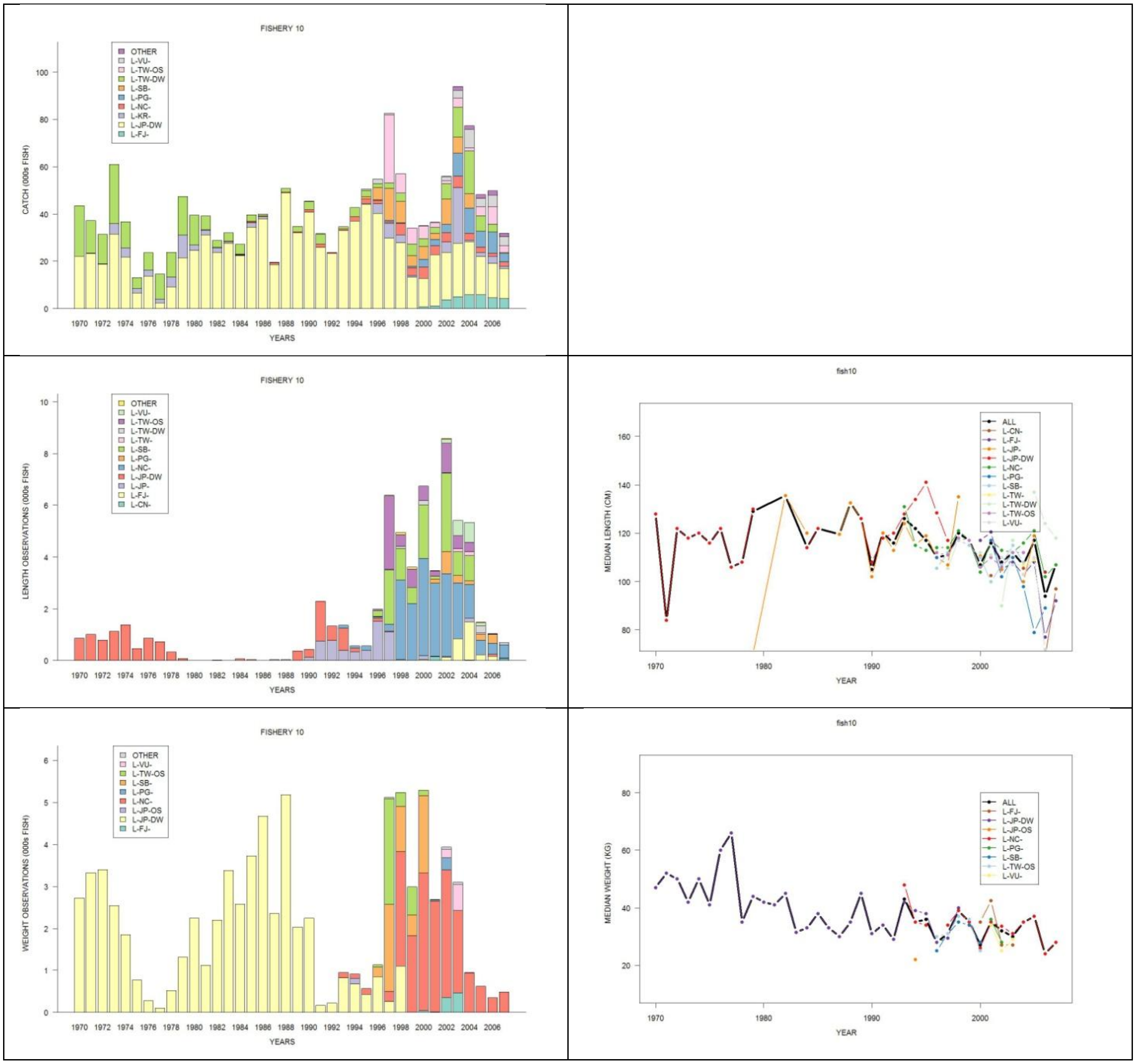


**Figure 13: Annual summaries of catch and size frequency data for Fishery 10 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**


**Figure 14: Annual summaries of catch and size frequency data for Fishery 12 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

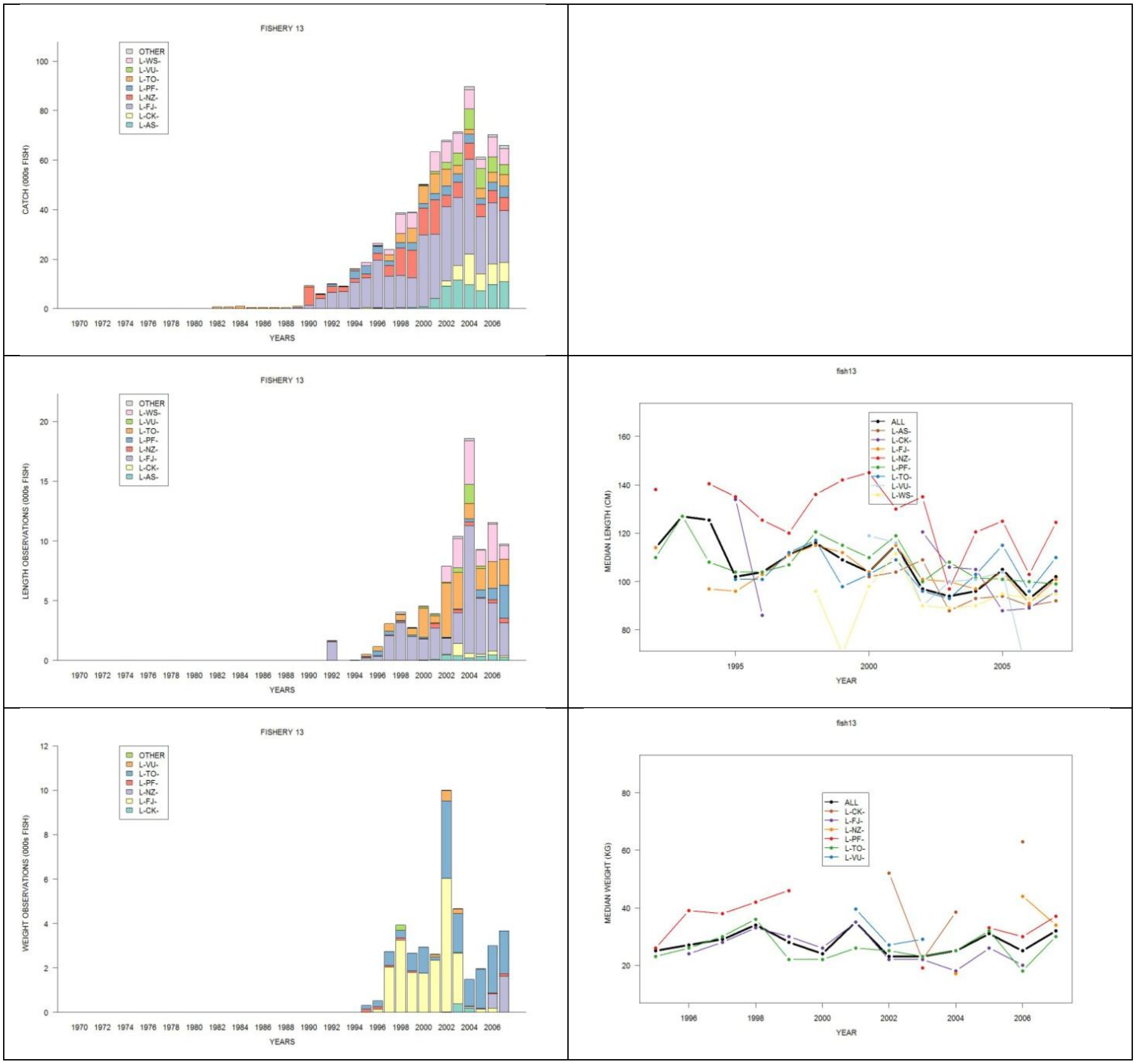


**Figure 15: Annual summaries of catch and size frequency data for Fishery 13 in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

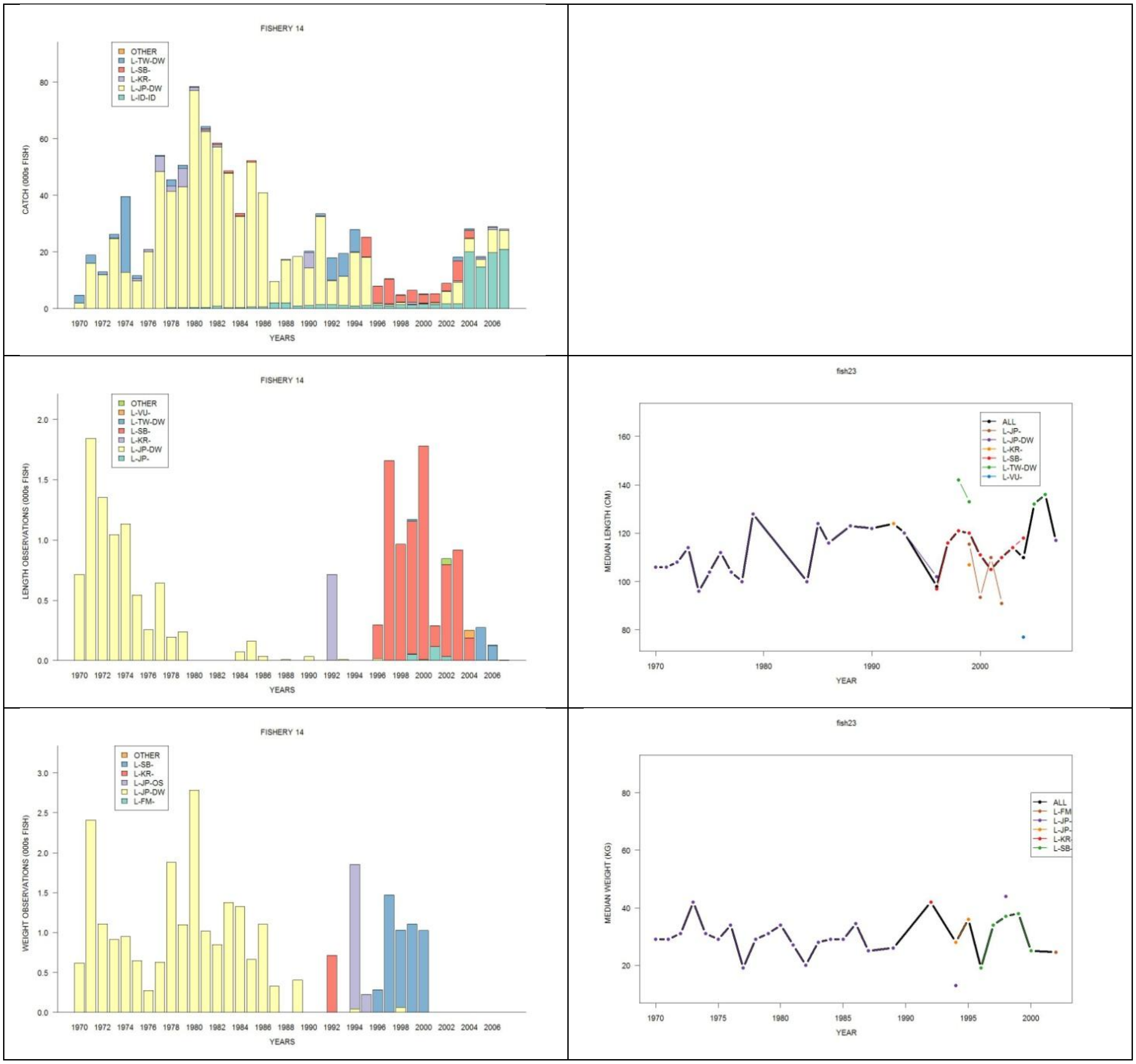


**Figure 16: Annual summaries of catch and size frequency data for Fishery <u>23</u> (not fishery 14) in the current BET assessment; catch by fleet (top left); number of length samples and median lengths (middle); and number of weight samples and median weights (bottom). Please note that the colours used for each fleet vary across panels.**

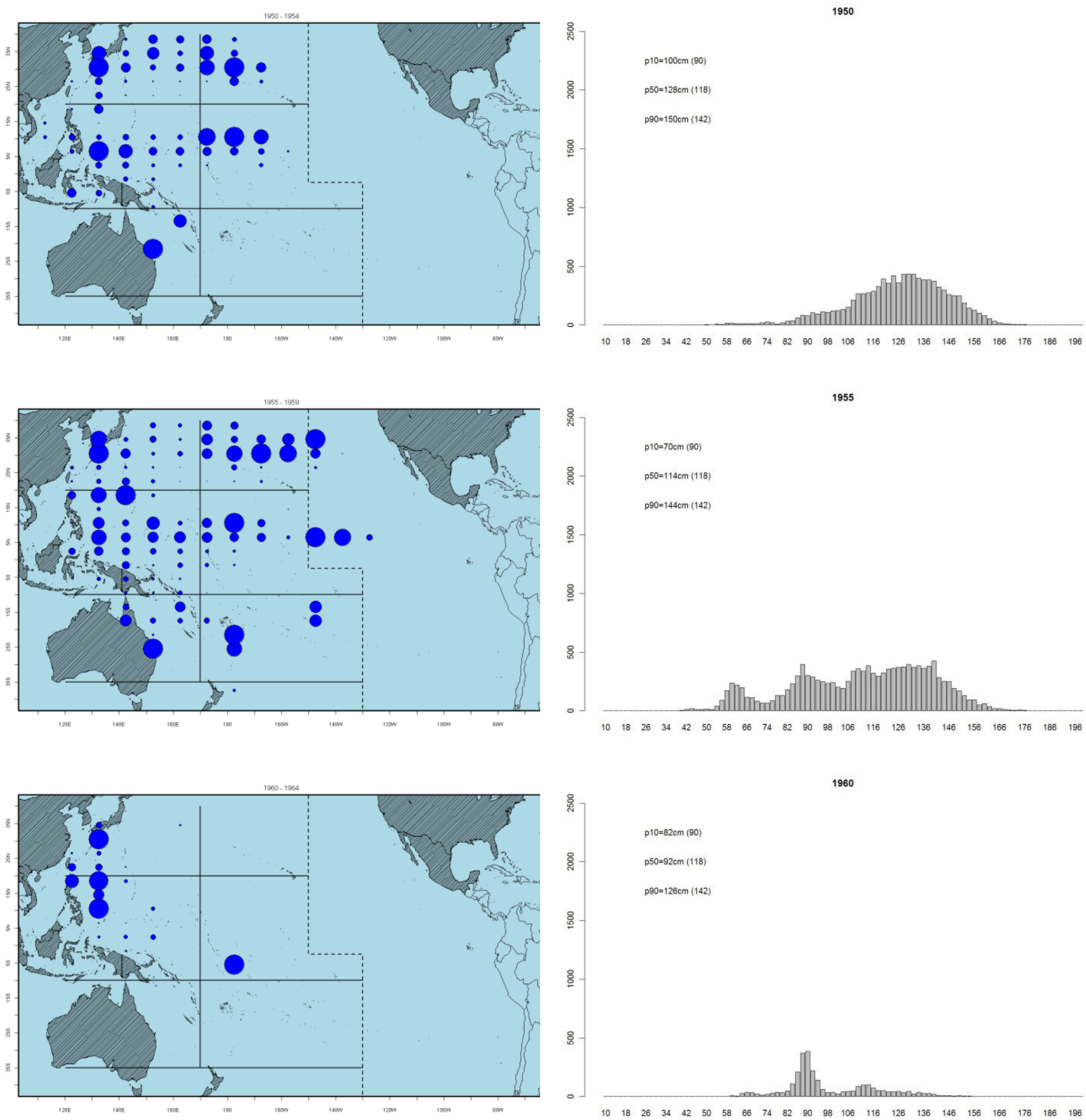


**Figure 17: Spatial distribution of Japanese length frequency samples (left) and length frequency distribution (right) in five yearly intervals from 1950-54 to 2005-09. The 10th, 50th, and 90th percentiles of the lengths for the five year period and the entire time period (in parentheses) are provided.**

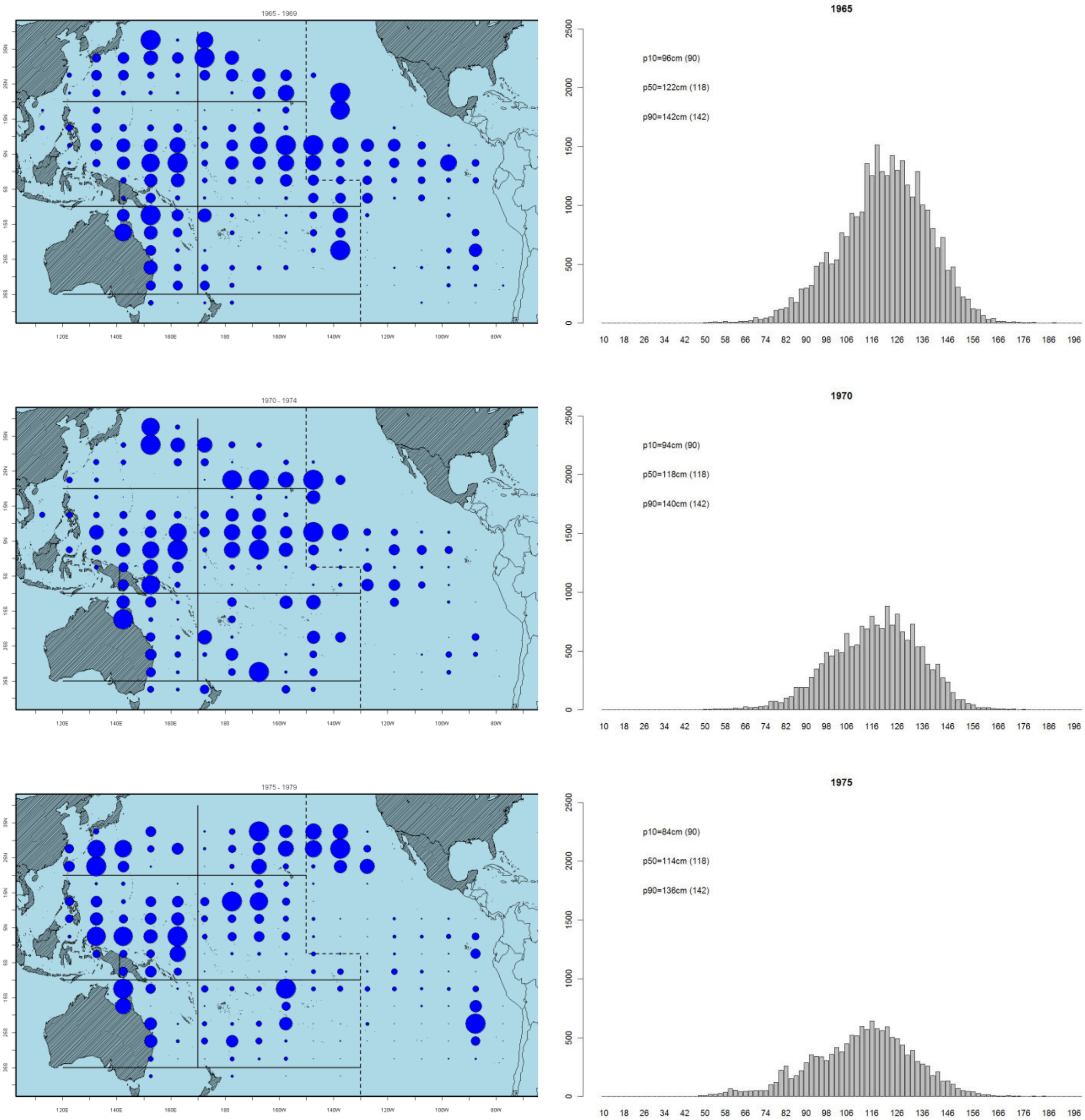


**Figure 17 continued.**

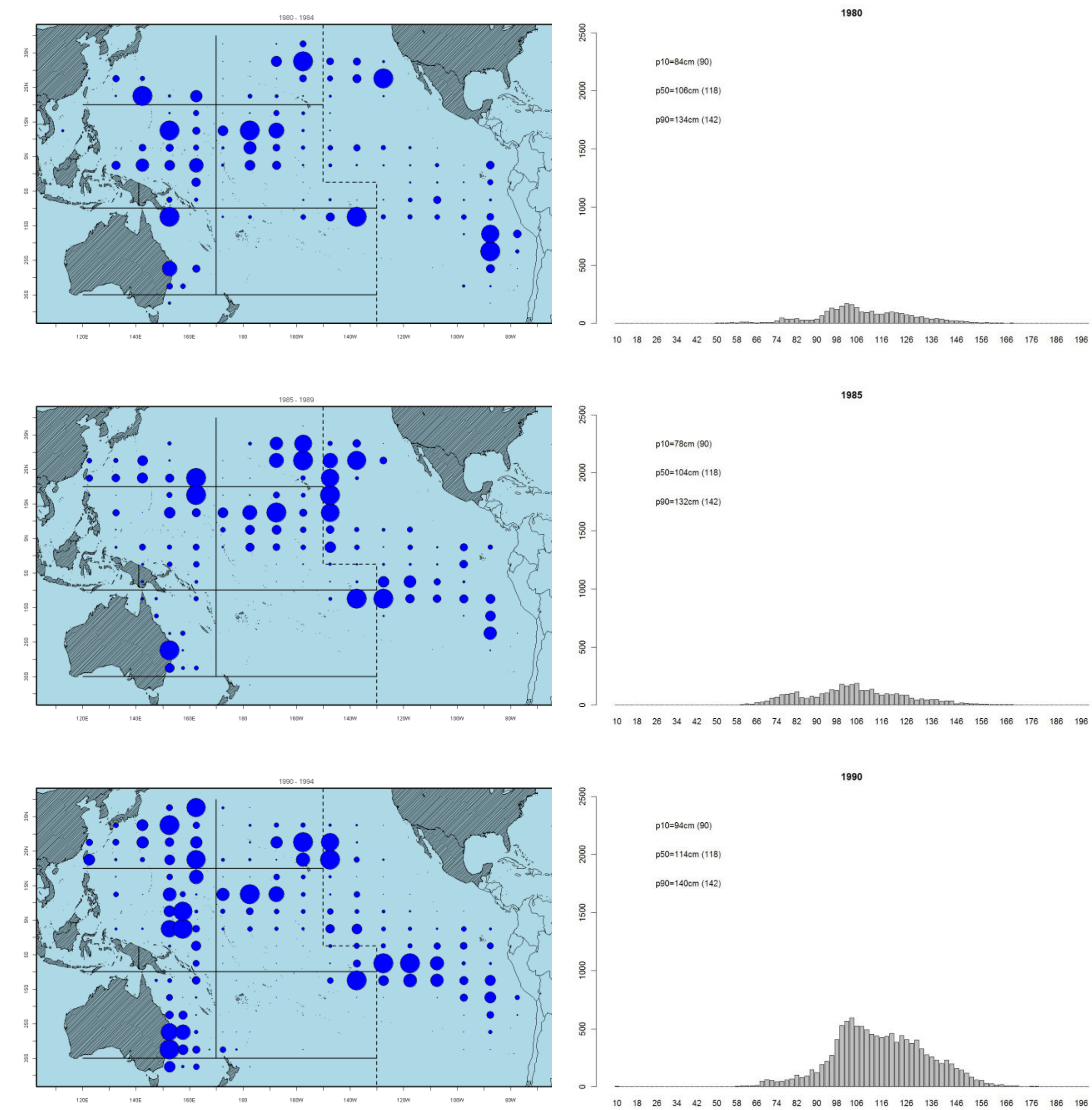


**Figure 17 continued.**

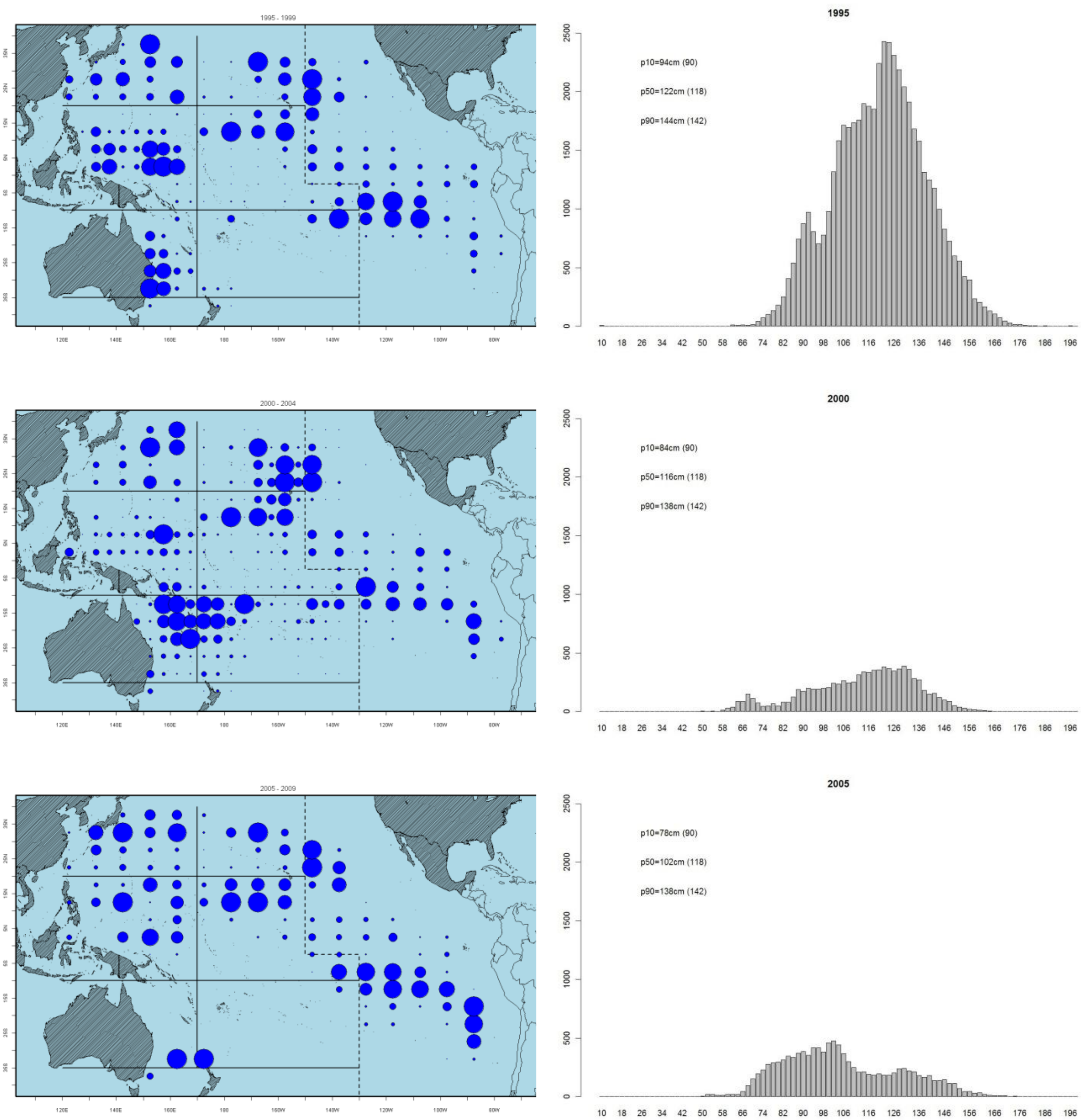


**Figure 17 continued.**

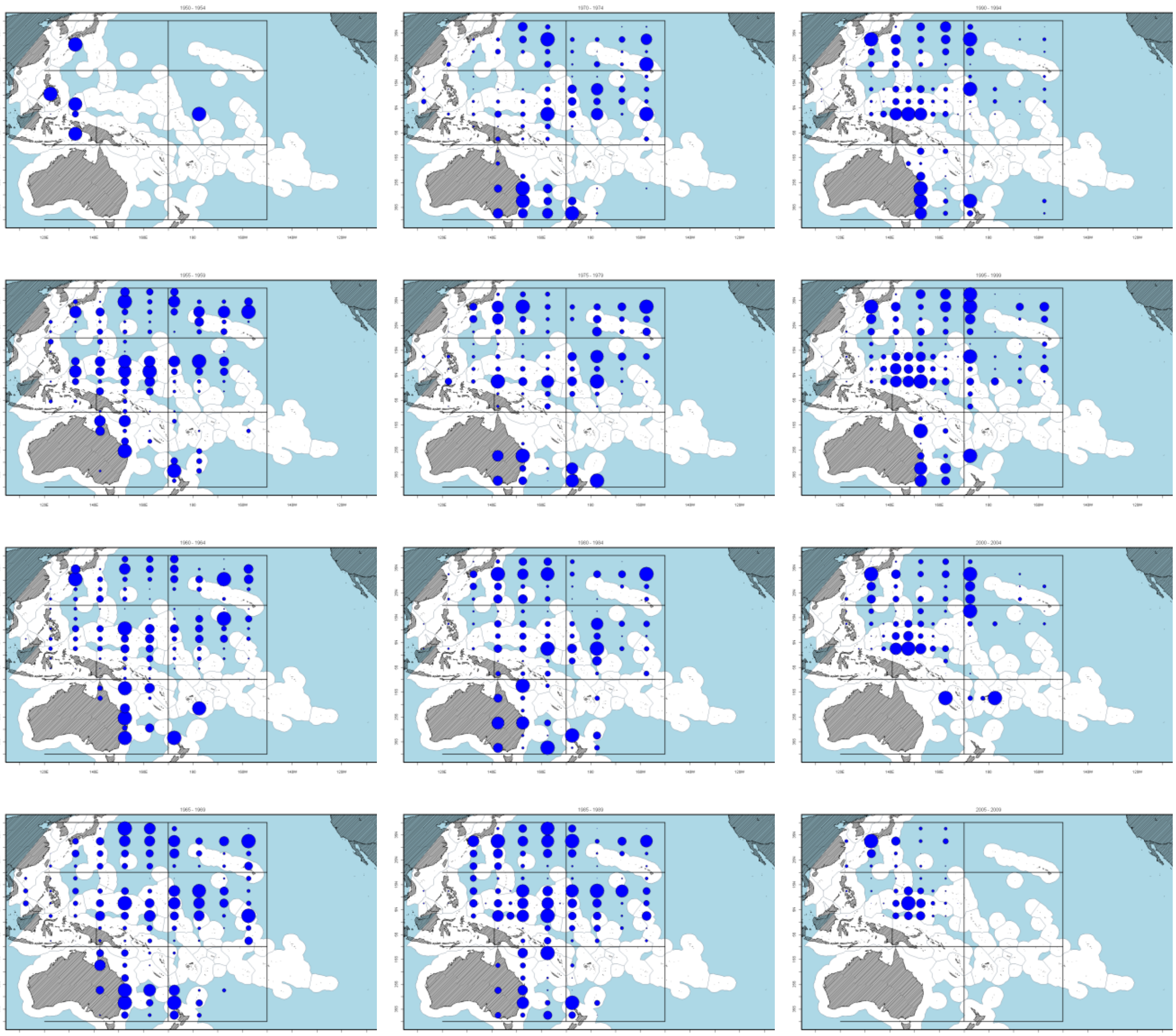

**Figure 18: Spatial distribution of weight frequency samples for the Japanese fleet in five yearly intervals from 1950-54 to 2005-2008. The figure includes data from both the distant water and off-shore components for the Japanese fleet.**

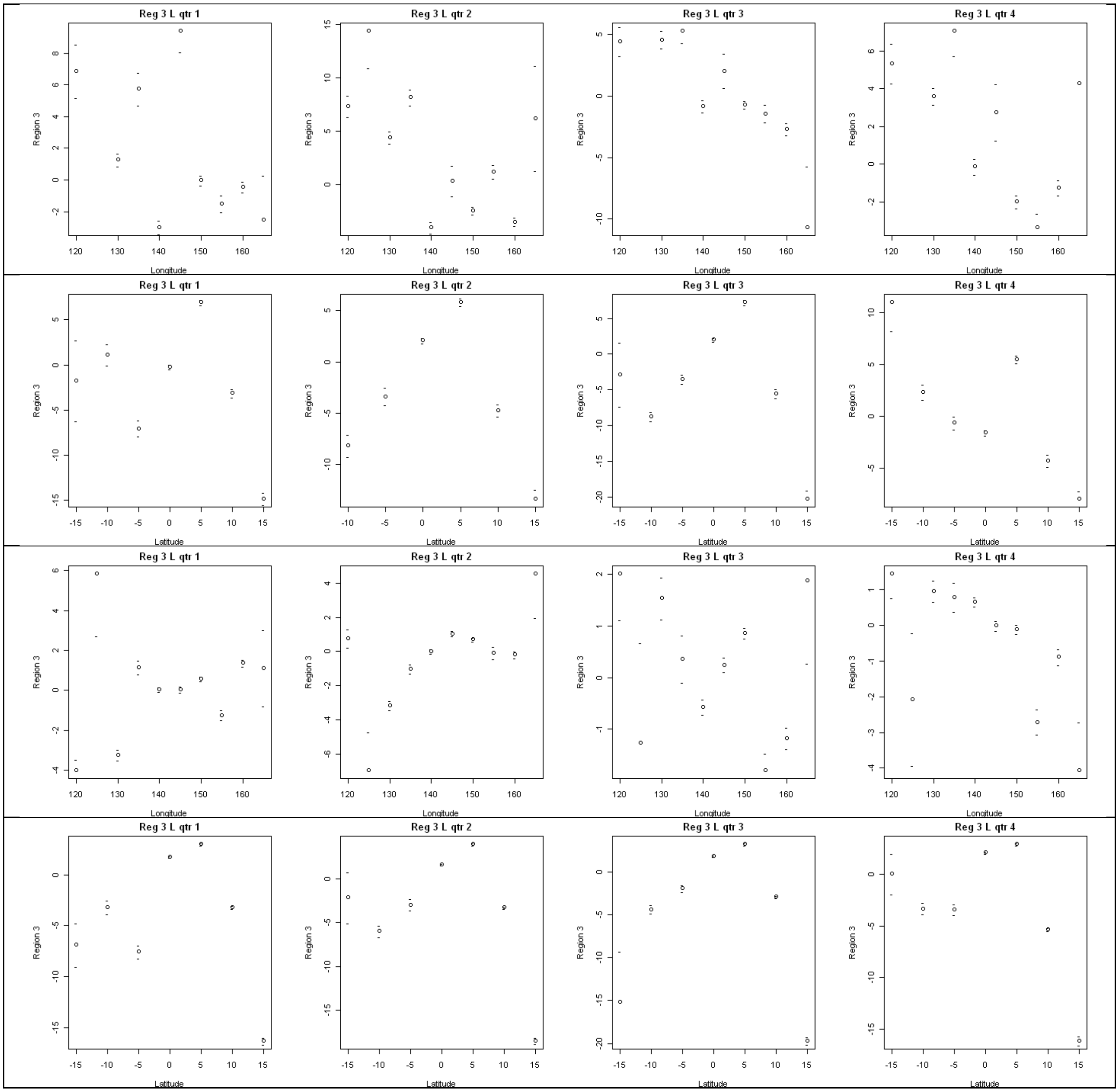


**Figure 19: Estimated longitudinal and latitudinal effects by quarter for fish length (top two panels) and fish weight (bottom two panels) for Japanese samples from region 3. The y-axes represent the difference in cm or kg from the overall mean that was estimated for that latitude or longitude bin, e.g. fish taken at 15N during the first quarter are typically 15 cm less than the overall average length.**

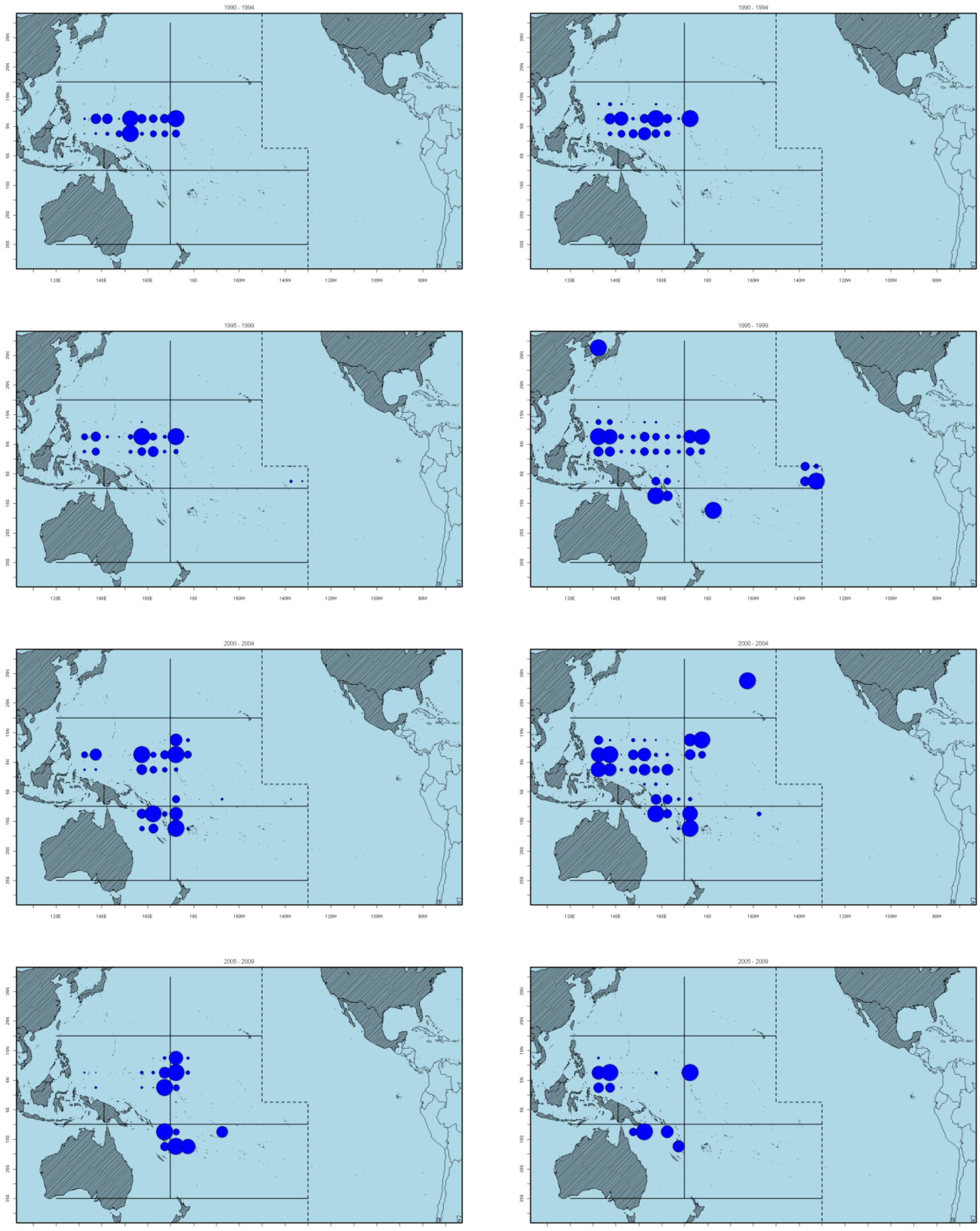


**Figure 20: Spatial distribution of length frequency samples for the Chinese Off-shore (left) and Taiwanese Off-shore (right) fleets in five yearly intervals from 1990-95 to 2005-2009.**

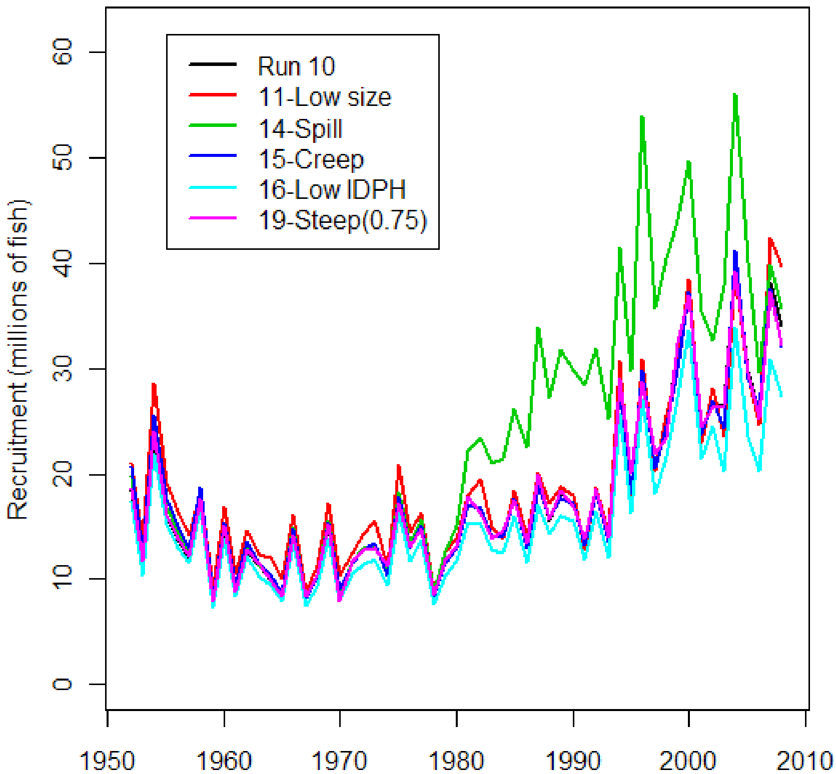


**Figure 21: Annual trends in overall recruitment from the 2009 assessment for the main model runs.**

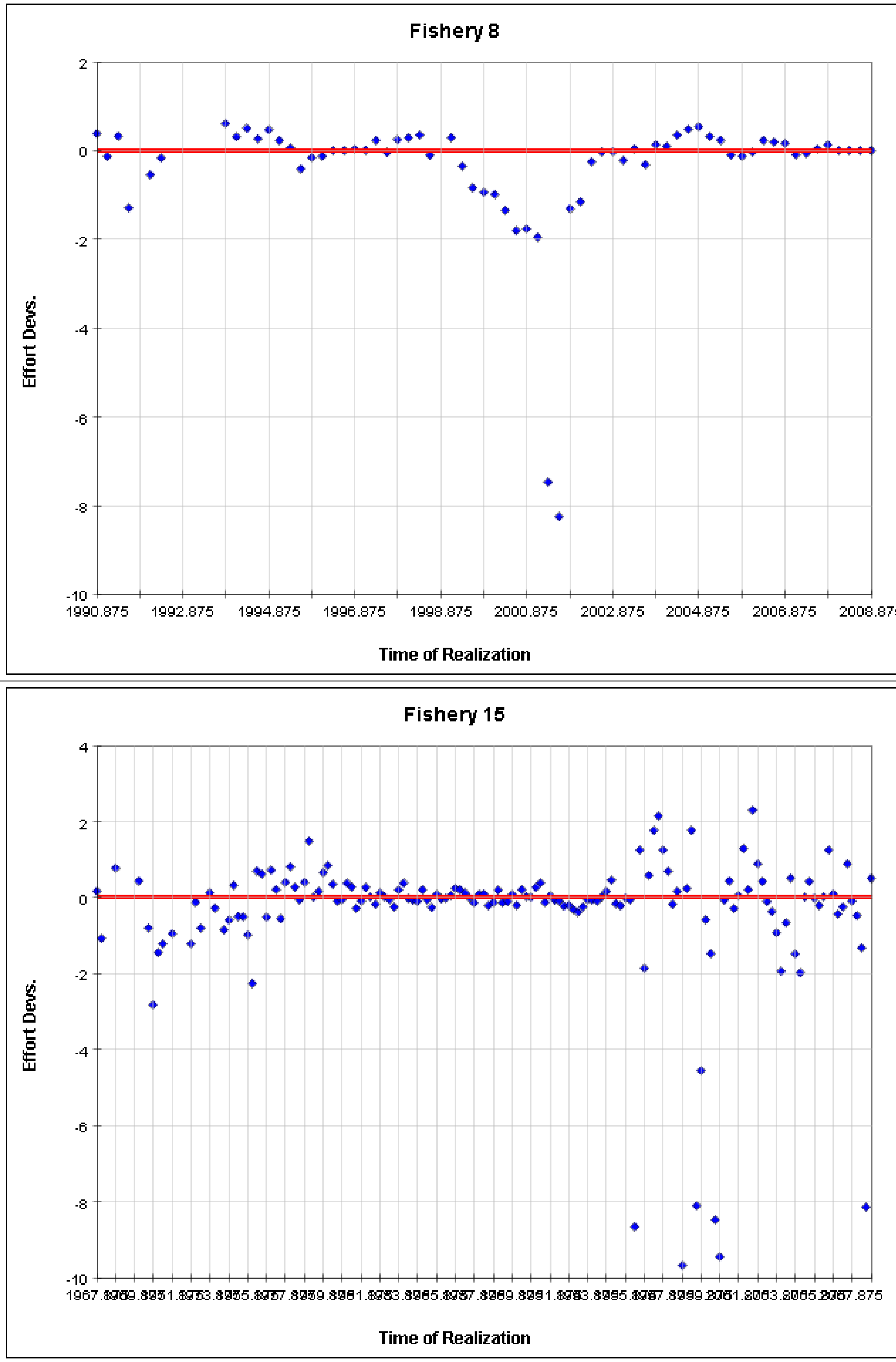


**Figure 22: Estimated effort deviates for the CN/TW-OS longline fishery in region 4 (Fishery 8 - top) and the unassociated school purse seine fishery in region 3 (Fishery 15 - bottom).**

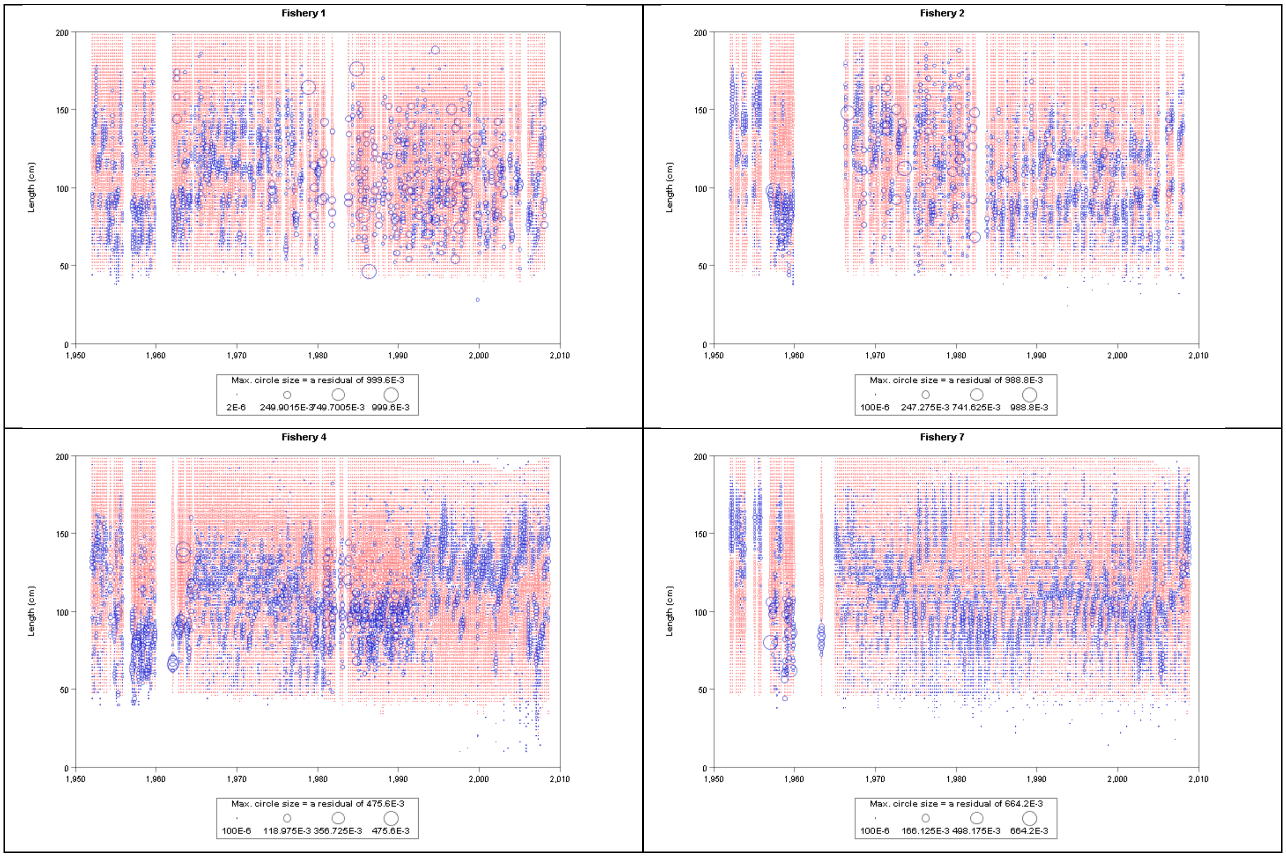


**Figure 23: Residuals from the fit to length frequency data for model 1 for the LL-ALL fisheries in region 1 (fishery 1), 2, (fishery 2), region 3 (fishery 4), and region 4 (fishery 7). Large blue circles indicate instances where there were much more fish in the catch of a given size than was predicted by the model.**

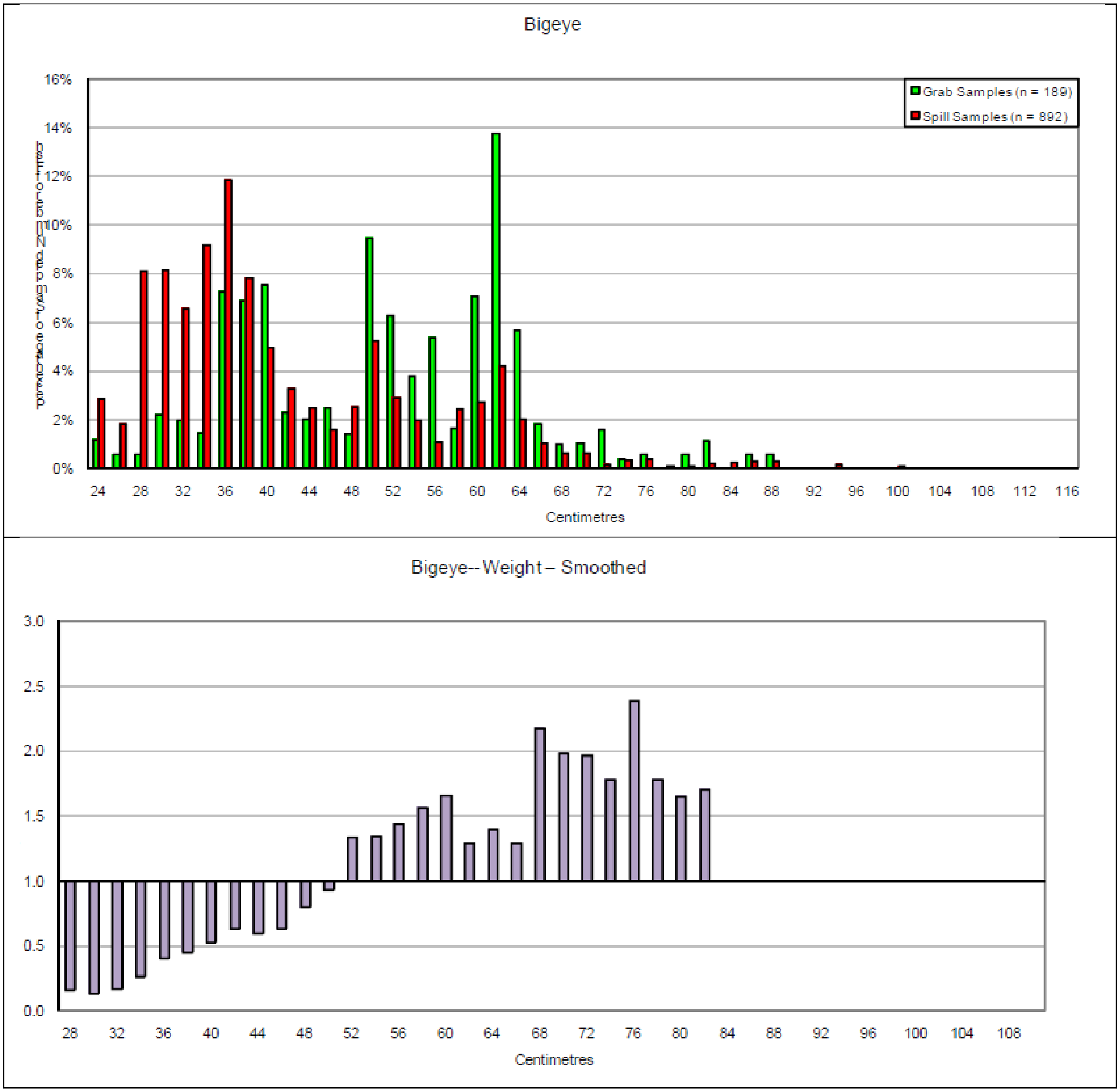


**Figure 24: (top) Comparison of length frequencies for paired spill and grab samples in terms of number of fish (from Lawson (2009; Figure 7)), and (bottom) grab selection bias in terms of weight (from Lawson (2009; Figure 9).**

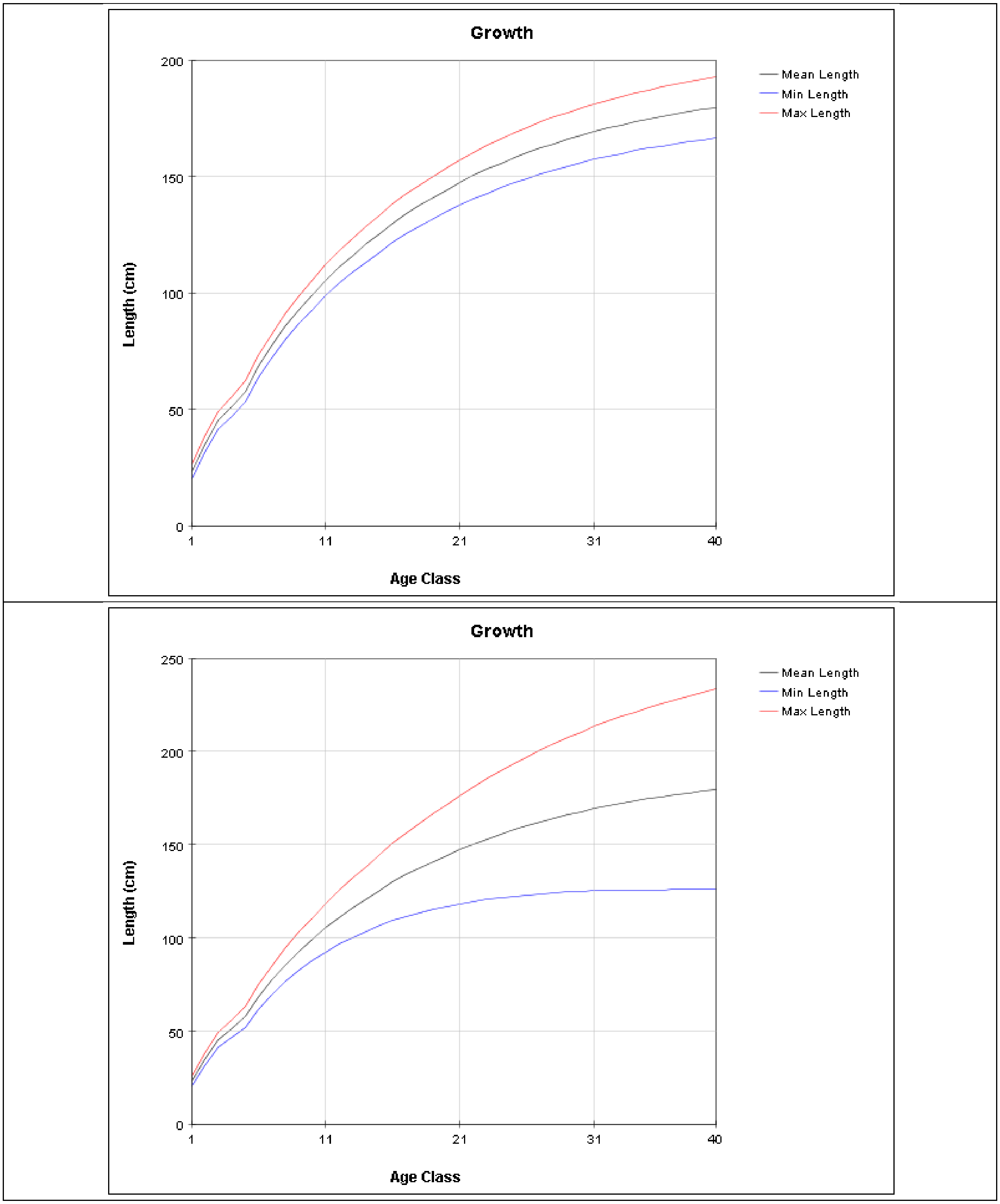

**Figure 25: Estimated growth curve from model 1 (top) and growth curve assumed in model 9 (bottom) which follows the same mean curve as model 1, but with much greater variability in length at age.**

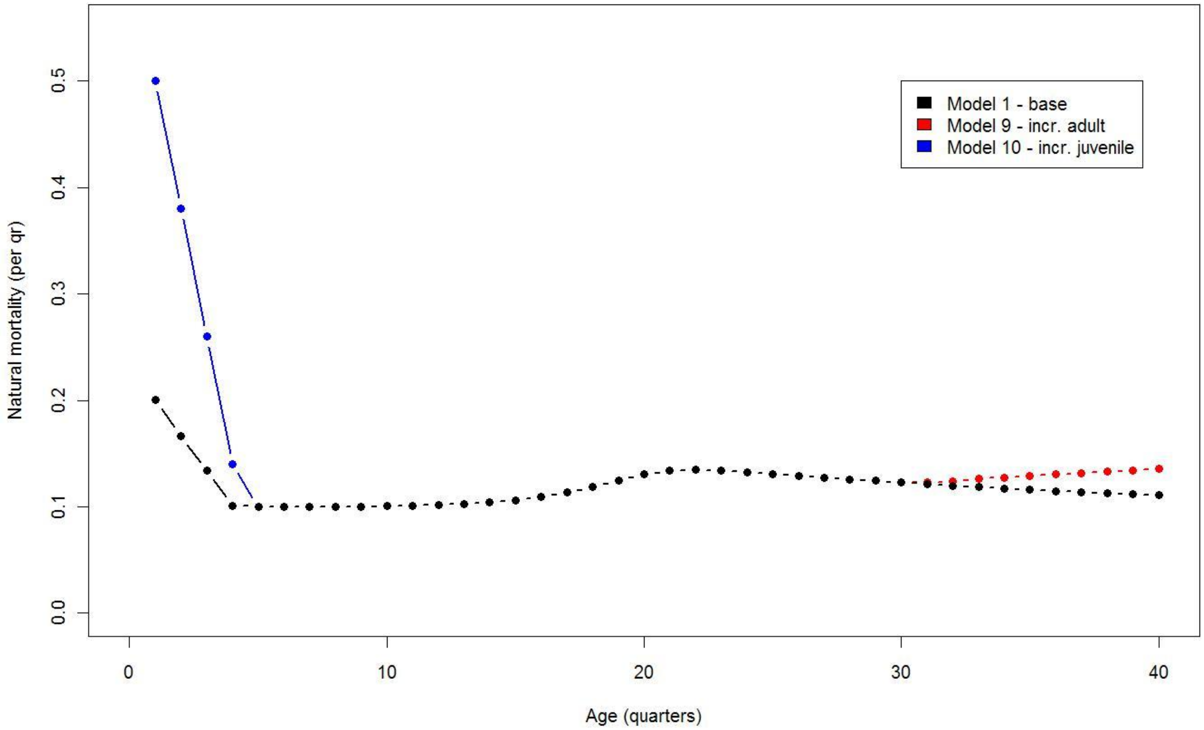


**Figure 26: Assumed values of quarterly natural mortality at age in models 1, 9, and 10.**

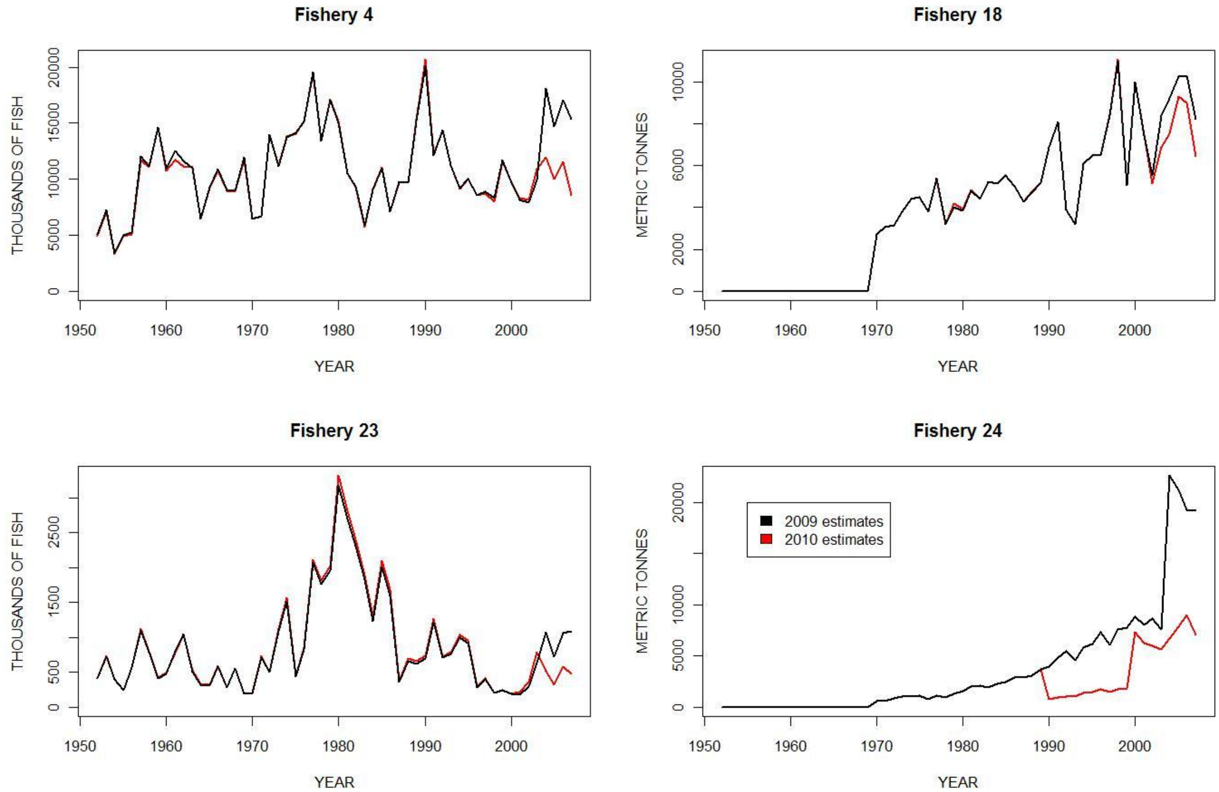


**Figure 27: Comparison of the catches assumed in the 2009 assessment for those fisheries that include the Philippines or Indonesia and the new catches that incorporate the revised catch estimates. See Table 1 for further details of each fishery.**

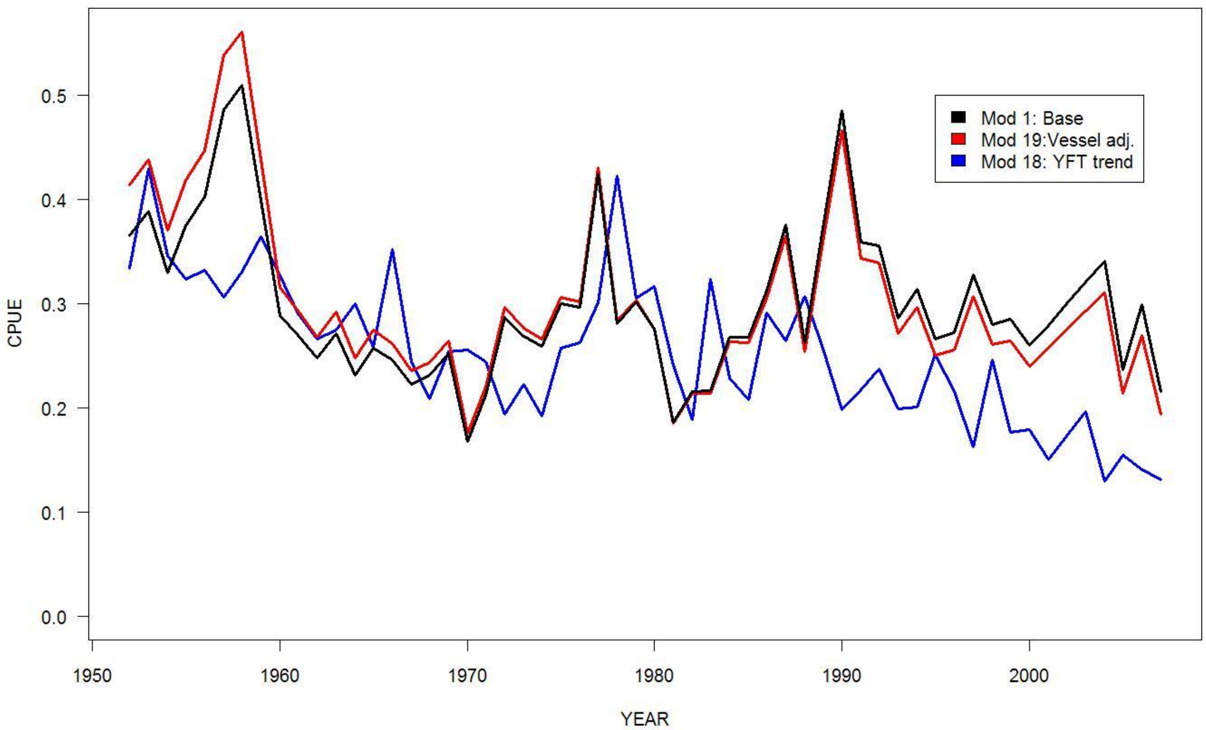


**Figure 28: The three CPUE series assumed in models 1, 18, and 19.**


**Figure 29: Annual recruitment trends for the six groups of model runs.**

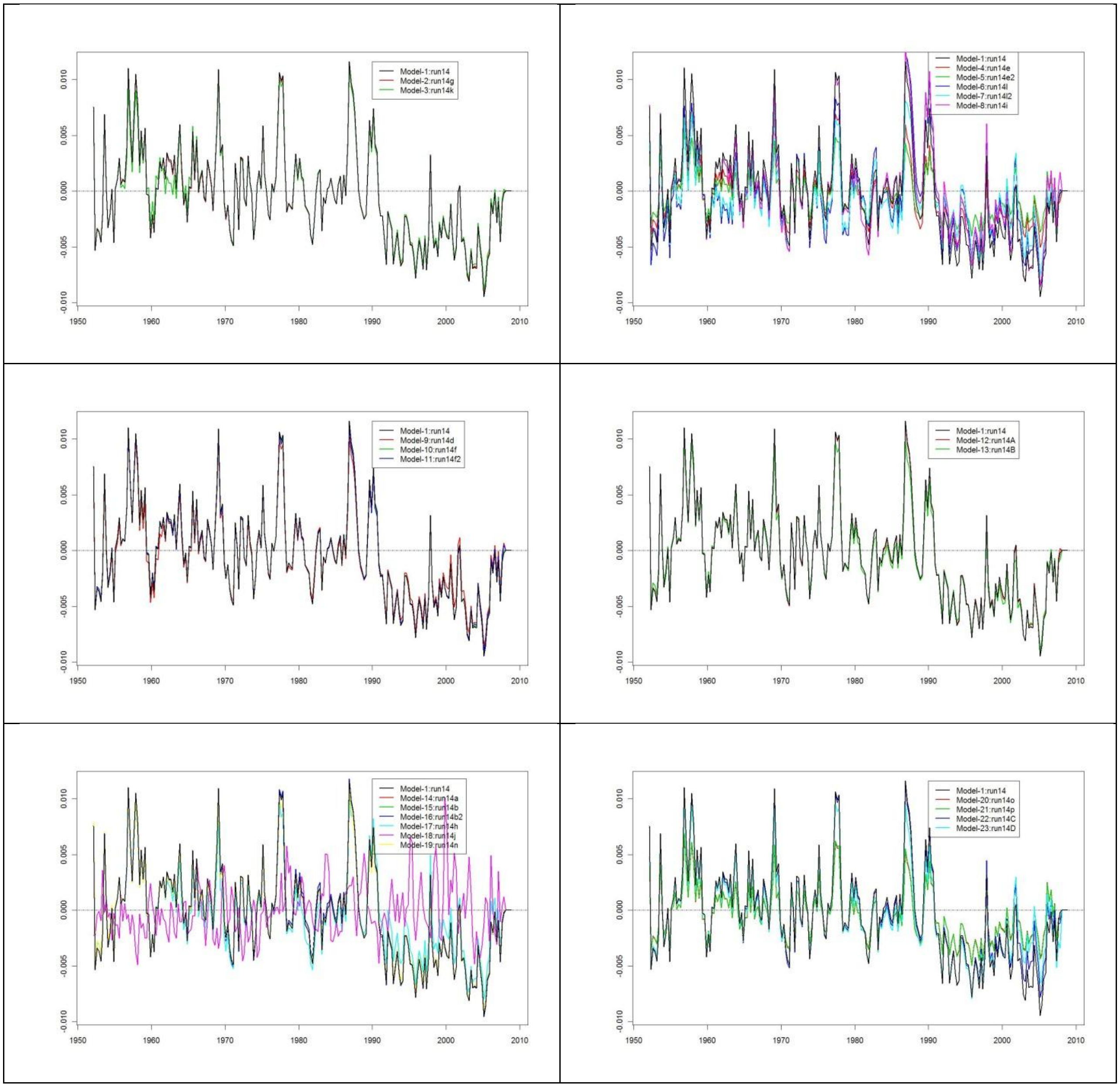


**Figure 30: Trends in effort deviates from fishery 4 (LL-ALL 3) for the six groups of model runs.**

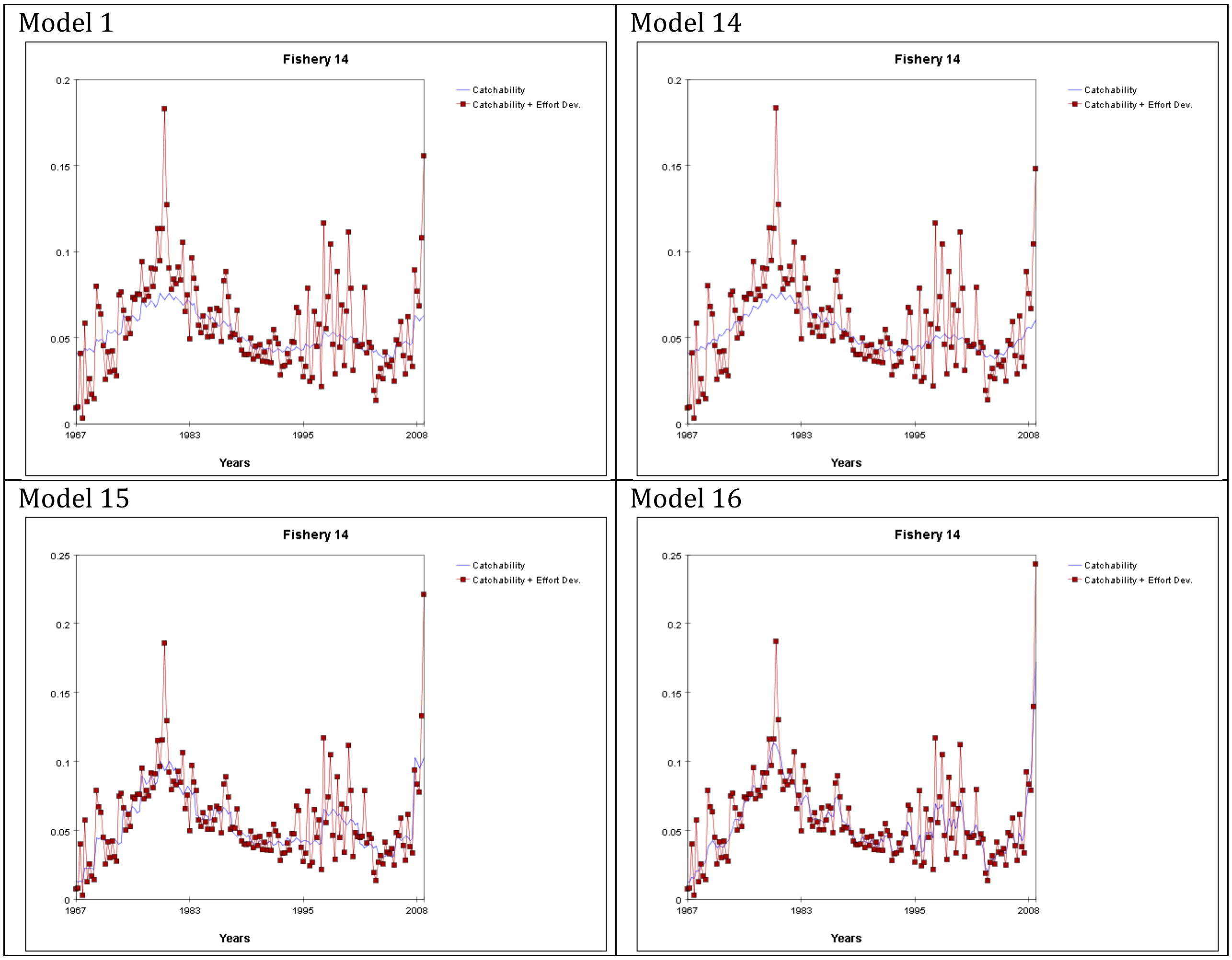


**Figure 31: Fishery 14 (FADs in region 3) catchability (blue line) and catchability plus effort deviate (red line with squares) trends. Note that the y-axis is not the same on all plots.**

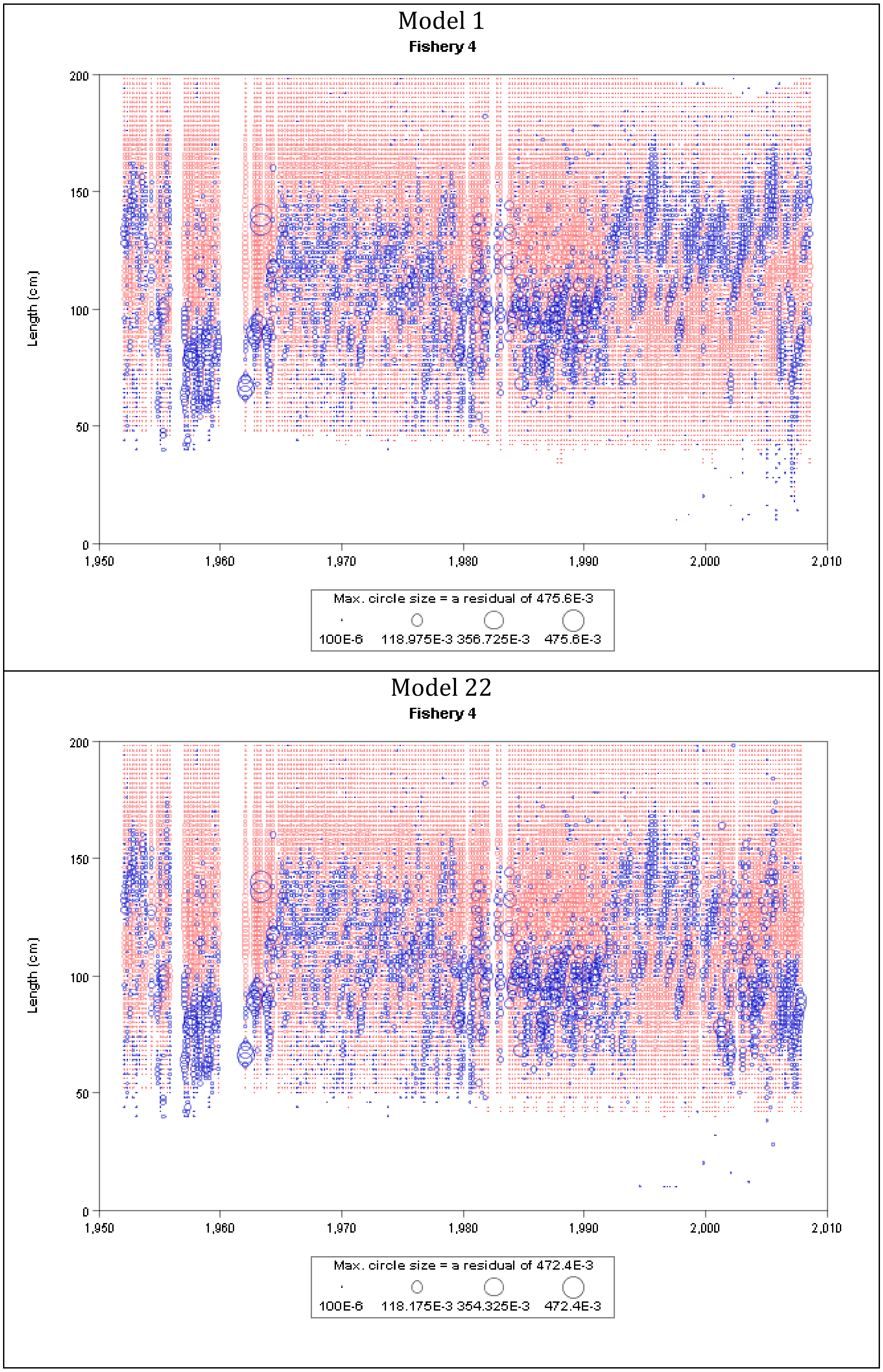


**Figure 32: Residual plots for the fit to length data for fishery 4 (LL-ALL in region 3) for models 1 and 22.**

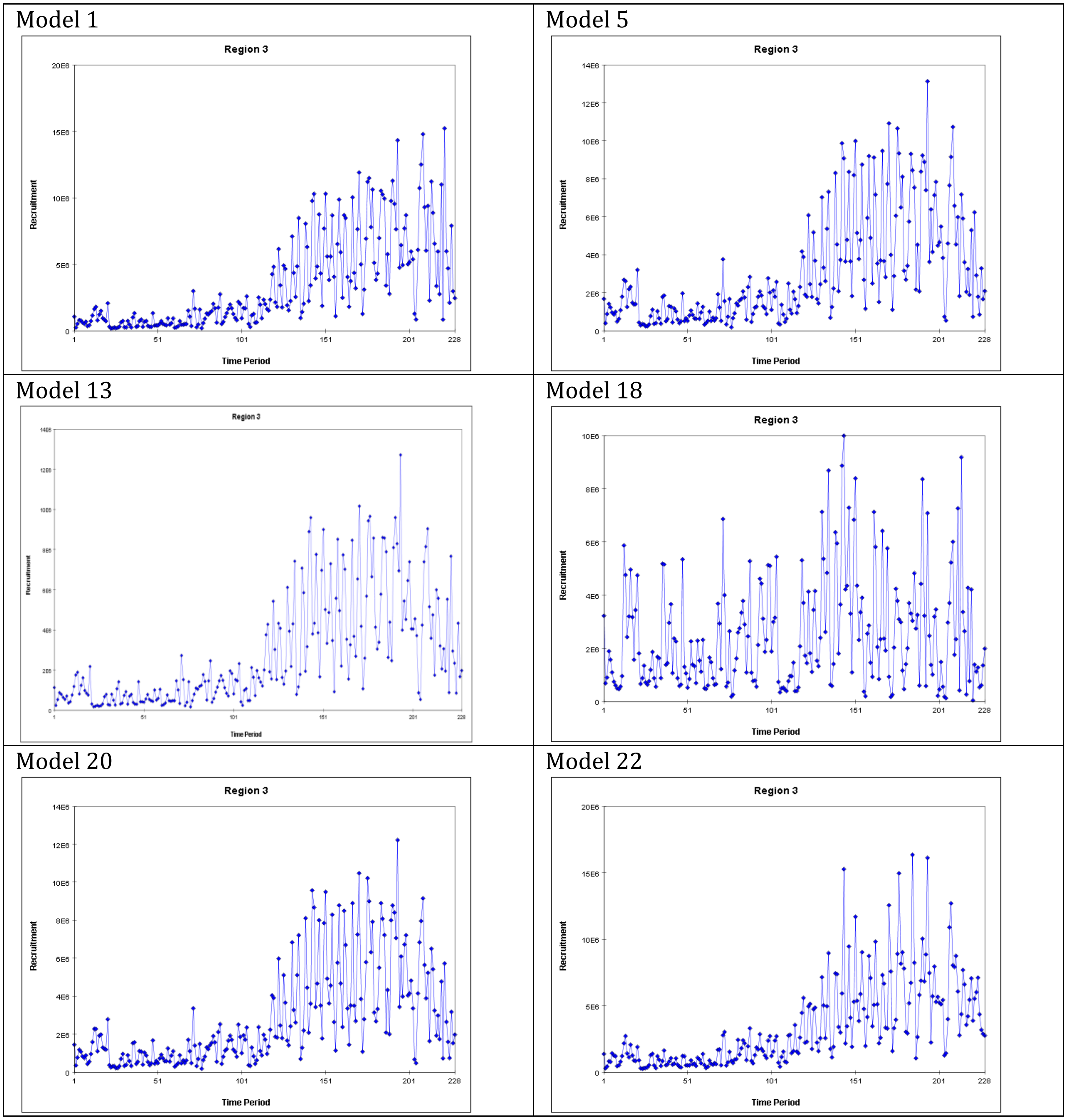


**Figure 33: Region three recruitment trends for six model runs. Note that the y-axis is not the same on all plots.**