# Collaborative study of tropical tuna CPUE from multiple Indian Ocean longline fleets in 2016.

Simon D. Hoyle[1], Doo Nam Kim[2], Sung Il Lee[2], Takayuki Matsumoto[3], Kaisuke Satoh[3], and Yu-Min Yeh[4].

[1] ISSF consultant. Email: simon.hoyle@gmail.com.

[2] National Institute of Fisheries Science, Republic of Korea. Email: doonam@korea.kr, and k.sungillee@gmail.com.

[3] National Research Institute of Far Seas Fisheries, Japan. Email: kstu21@affrc.go.jp and matumot@affrc.go.jp.

[4] Nanhua University, invited Taiwanese expert . Email: ymyeh@nhu.edu.tw

# Contents

COLLABORATIVE STUDY OF TROPICAL TUNA CPUE FROM MULTIPLE INDIAN OCEAN LONGLINE FLEETS IN 2016. ...... 1

**EXECUTIVE SUMMARY ........................................................................................ 10**

**INTRODUCTION ............................................................................................... 12**

TERMS OF REFERENCE ......................................................................................... 12

**METHODS ........................................................................................................ 13**

DATA CLEANING AND PREPARATION ..................................................................... 13

*Data ................................................................................................................ 13*

CLUSTER ANALYSIS ............................................................................................. 16

*Selecting the number of groups ........................................................................ 17*

*Plotting and data selection ................................................................................ 17*

CPUE STANDARDIZATION, AND FLEET EFFICIENCY ANALYSES ..................................... 18

*Distributions ..................................................................................................... 18*

*Models and datasets .......................................................................................... 19*

*Data periods ...................................................................................................... 19*

*Covariate effects ................................................................................................ 20*

*Indices of abundance ......................................................................................... 20*

**RESULTS .......................................................................................................... 21**

CLUSTER ANALYSIS ............................................................................................. 21

CPUE INDICES .................................................................................................... 22

**DISCUSSION .................................................................................................... 23**

**ACKNOWLEDGMENTS ....................................................................................... 26**

**REFERENCES .................................................................................................... 27**

**TABLES ............................................................................................................ 29**

**FIGURES ........................................................................................................... 42**

## Tables

Table 1: Analysis approaches for addressing the discontinuity in availability of vessel identity. ........19

Table 2: Data format for Japanese longline dataset. ............................................................29

Table 3: Number of available data by variable in the Japanese longline dataset. ..............................30

Table 4: Data format for Taiwanese longline dataset.............................................................34

Table 5: Tonnage as indicated by first digit of TW callsign. ...................................................35

Table 6: Codes in the Remarks field of the TW dataset, indicating outliers. ......................................35

Table 7a: Taiwanese data sample sizes by variable. ............................................................36

Table 8: Korean data description. ..............................................................................38

Table 9: Comparison of field availability among the three fleets. ................................................39

Table 10: Numbers of clusters included in clustering of each region and fishing fleet. .......................40

Table 11: Clusters included in indices for each fleet and region. ..................................................41

## Figures

Figure 1: Maps of the regional structures used to estimate bigeye (top left), yellowfin (top right) and albacore (lower) CPUE indices. ........ 42

Figure 2: Sets per day by region for the Japanese fleet in yellowfin regional structure. ........ 43

Figure 3: Sets per day by region for the Taiwanese fleet in the yellowfin regional structure........ 44

Figure 4: Sets per day by region for the Korean fleet in the yellowfin regional structure. ........ 45

Figure 5: Proportions of sets retained after data cleaning for analyses in this paper, by region and year-quarter, for Japanese (top left), Taiwanese (top right), and Korean (bottom left) data........ 46

Figure 6: Proportions of Taiwanese catch in number reported as bigeye, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of bigeye. Contour lines occur at 5% intervals. ........ 48

Figure 7: Proportions of Taiwanese catch in number reported as yellowfin tuna, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of yellowfin. Contour lines occur at 5% intervals........ 50

Figure 8: Proportions of Taiwanese catch in number reported as 'other' species, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of 'other' species. Contour lines occur at 5% intervals........ 51

Figure 9: Taiwanese catch rates per hundred hooks of oilfish, sharks, skipjack, and other tunas, by region and year-quarter........ 52

Figure 10: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 1 (western tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ........ 53

Figure 11: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 2 (eastern tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ........ 54

Figure 12: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 3 (western temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ........ 55

Figure 13: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 4 (eastern temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ........ 56

Figure 14: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 2 (western tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ........ 57

Figure 15: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 3 (western temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ...................................... 58

Figure 16: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 4 (eastern temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ...................................... 59

Figure 17: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 5 (eastern tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right). ................................................... 60

Figure 18: Beanplots for bigeye region 1 (western tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................................... 61

Figure 19: Beanplots for bigeye region 1 (western tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................... 62

Figure 20: Maps of the spatial distributions of clusters in region 1 (western tropical), for Japanese, Korean, and Taiwanese effort. ......................................................................................... 63

Figure 21: Beanplots for bigeye region 2 (eastern tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................................... 64

Figure 22: Beanplots for bigeye region 2 (eastern tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................... 65

Figure 23: Maps of the spatial distributions of clusters in region 2 (eastern tropical), for Japanese, Korean, and Taiwanese effort. ......................................................................................... 66

Figure 24: Beanplots for bigeye region 3 (western temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................... 67

Figure 25: Beanplots for bigeye region 3 (western temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................... 68

Figure 26: Maps of the spatial distributions of clusters in region 3 (western temperate), for Japanese, Korean, and Taiwanese effort. ......................................................................................... 69

Figure 27: Beanplots for bigeye region 4 (eastern temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ......................................................................................................................... 70

Figure 28: Beanplots for bigeye region 4 (eastern temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....71

Figure 29: Maps of the spatial distributions of clusters in region 4 (eastern temperate), for Japanese, Korean, and Taiwanese effort. ....72

Figure 30: Beanplots for yellowfin region 2 (western tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....73

Figure 31: Beanplots for yellowfin region 2 (western tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....74

Figure 32: Maps of the spatial distributions of clusters in region 2 (western tropical), for Japanese, Korean, and Taiwanese effort. ....75

Figure 33: Beanplots for yellowfin region 3 (western temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....76

Figure 34: Beanplots for yellowfin region 3 (western temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....77

Figure 35: Maps of the spatial distributions of clusters in region 3 (western temperate), for Japanese, Korean, and Taiwanese effort. ....78

Figure 36: Beanplots for yellowfin region 4 (eastern temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....79

Figure 37: Beanplots for yellowfin region 4 (eastern temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....80

Figure 38: Maps of the spatial distributions of clusters in region 4 (eastern temperate), for Japanese, Korean, and Taiwanese effort. ....81

Figure 39: Beanplots for yellowfin region 5 (eastern tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....82

Figure 40: Beanplots for yellowfin region 5 (eastern tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians. ....83

Figure 41: Maps of the spatial distributions of clusters in region 5 (eastern tropical), for Japanese, Korean, and Taiwanese effort. ....84

Figure 42: Estimated CPUE series for bigeye region 1 (western tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 85

Figure 43: Estimated CPUE series for bigeye region 2 (eastern tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 86

Figure 44: Estimated CPUE series for bigeye region 3 (western temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 87

Figure 45: Estimated CPUE series for bigeye region 4 (eastern temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 88

Figure 46: Estimated CPUE series for yellowfin region 2 (western tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 89

Figure 47: Estimated CPUE series for yellowfin region 3 (western temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 90

Figure 48: Estimated CPUE series for yellowfin region 4 (eastern temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 91

Figure 49: Estimated CPUE series for yellowfin region 5 (eastern tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects. ........ 92

Figure 50: Diagnostics plots for bigeye lognormal positive models in tropical regions 1 and 2, for 1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right). ........ 93

Figure 51: Diagnostics plots for bigeye lognormal positive models in temperate regions 3 and 4, for1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right). ........ 94

Figure 52: Diagnostics plots for yellowfin lognormal positive models in tropical regions 2 and 5, for 1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right). ........ 95

Figure 53: Diagnostics plots for yellowfin lognormal positive models in temperate regions 3 and 4, for1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right). ........ 96

Figure 54: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 1 (western tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right). ........ 97

Figure 55: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 2 (eastern tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right). ........ 98

Figure 56: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 3 (western temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right)............................................99

Figure 57: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 4 (eastern temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right)..........................................100

Figure 58: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 2 (western tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right)..........................................101

Figure 59: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 3 (western temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right). ...............................102

Figure 60: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 4 (eastern temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right). ...............................103

Figure 61: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 5 (eastern tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right)..........................................104

Figure 62: Bigeye residuals for tropical regions 1 (above) and 2 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right). .............................................................105

Figure 63: Bigeye residuals for temperate regions 3 (above) and 4 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right). ......................................................106

Figure 64: Yellowfin residuals for tropical regions 2 (above) and 5 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right). ......................................................107

Figure 65: Yellowfin residuals for temperate regions 3 (above) and 4 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right).......................................108

Figure 66: Comparison of the joint indices described in this paper (red) with the Japanese indices developed in 2015 and used in the 2015 yellowfin stock assessment. ............................................109

Figure 67: Ratio of the joint indices described in this paper divided by the Japanese indices developed in 2015 and used in the 2015 yellowfin stock assessment. ............................................110

Figure 68: Comparison of the joint indices described in this paper (red) with the Japanese indices developed in 2013 and used in the 2013 bigeye stock assessment (black). ....................................111

Figure 69: Ratio of the joint indices described in this paper divided by the Japanese indices developed in 2013 and used in the 2013 bigeye stock assessment. ................................................112

# Executive Summary

In March and April 2016 a collaborative study was conducted between national scientists with expertise in Japanese, Taiwanese, and Korean longline fleets, and an independent scientist. The meetings addressed Terms of Reference covering several important issues related to albacore, bigeye and yellowfin tuna CPUE indices in the Indian Ocean. A further meeting between the parties was held in July 2016 to update the tropical tuna indices. The study was funded by the International Seafood Sustainability Foundation (ISSF) and the Indian Ocean Tuna Commission (IOTC).

Terms of Reference:

1. To validate and improve methods for developing indices of abundance for tropical tunas.
2. To develop methods for providing indices of abundance for albacore tuna.
3. To provide indices of abundance for albacore tuna, and to draft a working paper to be presented at the 2016 WPTmT06 (18 – 21 July 2016).
4. To provide indices of abundance for bigeye and yellowfin tunas and to draft a working paper to be presented at the WPTT18 (5 – 10 November 2016).
5. To provide support and training to national scientists in their analyses of catch and effort data.

This document describes the development of indices of abundance for bigeye and yellowfin tunas.

Data were provided for the three fleets in similar formats, with varying combinations of species and variables, due to differences between the fisheries' data collection forms and processes and their changes through time. See Table 9 for a comparison of field availabilities among the three fleets. All datasets reported set date, number of hooks, hooks between floats for at least part of the time series, set location at some resolution, vessel identity for part or all of the dataset, and catch in number of albacore, bigeye, yellowfin, southern bluefin tuna, swordfish, blue marlin, striped marlin, and black marlin.

Japanese operational data were available from 1952-2015, with location reported to 1° of latitude and longitude, vessel call sign from 1979, hooks between floats for much of the time series, and date of trip start (Tables 2 and 3). The Taiwanese operational data were available 1979-2015, but data prior to 2005 were not used in tropical tuna analyses, due to concerns about data quality. Taiwanese vessel call sign was available for the whole time period along with information on vessel size; set location at 5° resolution until 1994, and 1° subsequently; number of hooks between floats from 1995; and catches in number for the species above plus other tuna, other billfish, skipjack, shark, and other species; equivalent values in weight for all species; SST; bait type fields ('Pacific saury', 'mackerel', 'squid', 'milkfish', and 'other'); depth of hooks (m); set type (type of target); remarks (indicating outliers); departure date from port; starting date of operations on a trip; stopping date of operations on a trip; and arrival date at port (Table 4). Korean data were available for 1971 to 2015 (Table 8), with the standard fields and vessel id, operation location to 1°, hooks between floats calculated for each set, and additional species 'other', sailfish, shark, and skipjack.

Data were cleaned by removing obvious errors and missing values (Figure 5). Unlikely but potentially plausible values (e.g. sets with very large catches of a species) were retained. Each set was allocated to bigeye and yellowfin regions according to several alternative regional definitions, and data outside these areas ignored. Standard datasets were produced for each fleet.

We applied cluster analysis methods to identify effort associated with different fishing strategies, using the approaches developed in the 2015 IOTC CPUE standardization workshop (Hoyle et al. 2015b). Data were aggregated by vessel-month and then clustered on species composition in the catch, using the Ward hclust method. Clustering was carried out by fleet and region, and a fleet/cluster group parameter was assigned to each set. The clustered data for all fleets were combined into a joint dataset, and then separated into regions for both bigeye and yellowfin. For each region and fleet, clusters were removed if the species of interest was a very small component of the catch.

Data for each region were standardized using regression techniques to estimate indices of abundance. The dependent variable was the presence/absence of the species of interest in the catch (binomial models), or the positive catch of the species of interest in numbers of fish (lognormal models). All models included the explanatory variables year-quarter and 5° cell as categorical variables, and a cubic spline on hooks as a covariate. Models for tropical regions included a cubic spline fitted to hooks between floats, while models for temperate areas included a categorical variable for cluster. Some models were run with vessel identity as a categorical variable. Models were run for the period 1952-1979 without vessel identity, for the later period 1979-2015 with vessel identity, and for the whole period 1952-2015 both with and without vessel identity. Indices were estimated using both a delta lognormal approach, and lognormal constant generalized linear models.

Figures and tables are provided for each set of indices, including both quarterly and annual indices. Diagnostic plots are also presented.

# Introduction

In March and April 2016 a collaborative study of longline data and CPUE standardization for bigeye, yellowfin, and albacore tuna was conducted between scientists with expertise in Japanese, Taiwanese, and Korean fleets, and an independent scientist. A further meeting was held in July 2016 to update the tropical tuna analyses with the most recent data. The study was funded by the International Seafood Sustainability Foundation (ISSF) and the Indian Ocean Tuna Commission (IOTC). The study addressed the Terms of Reference outlined below, which cover the most important issues that had previously been highlighted by different working parties. Work was carried out, for those factors relevant to them, for the following:

- Area: Indian Ocean
- Fleets: Japanese longline; Taiwanese longline, Korean longline
- Stocks: Bigeye tuna, yellowfin tuna, albacore tuna.

The current document addresses CPUE standardizations for the tropical tunas. The methods description includes approaches used for albacore in order to generalize the report, but to conserve space only bigeye and yellowfin tuna results are reported.

## Terms of Reference

- To organize a series of meetings between data holders and the consultant.
- To validate and improve methods for developing indices of abundance for tropical tunas.
- To develop methods for providing indices of abundance for albacore tuna.
- To provide indices of abundance for albacore tuna, and to draft a working paper to be presented at the 2016 IOTC WPTmT06 (18 – 21 July 2016).
- To provide indices of abundance for bigeye and yellowfin tunas and to draft a working paper to be presented at the IOTC WPTT18 (5 – 10 November 2016).
- To provide support and training to national scientists in their analyses of catch and effort data.
- The analyses will consider data to be provided by Japanese, Taiwanese, and Korean research agencies.
- Analyses will be carried out in a series of meetings in March and April, and in a final meeting focusing on tropical tunas following update of the data. After preliminary meetings between the consultant and each participating data provider to prepare each dataset and develop methods, there will be a first joint meeting between all participating parties and the consultant. This joint meeting will develop indices for albacore tuna and develop draft indices for bigeye and yellowfin tunas. A second joint meeting will occur in July or August to prepare final indices for bigeye and yellowfin tuna, and to provide training to national scientists in their analyses of catch and effort data.
- ***Data analysis tasks will include the following***:
- Load, prepare, and check each dataset, given that data formats and pre-processing often change between years and data extracts, and important changes to fleets and reporting sometimes occur in new data. The format of the Japanese data is expected to change before the second joint meeting which will require additional time during this meeting.

- Explore albacore catch and effort data from each CPC to check the reliability and coverage of reporting, as we did for tropical tunas
- Apply cluster analyses and BET + YFT CPUE standardization using reliable data from each CPC. Change regional structures from the generic 2015 approach to regions that are appropriate for each assessment, including alternate options.
- Address outstanding issues from 2015 tropical tuna analyses, including a) adjusting for the introduction of vessel effects in late-1970s Japanese data, and b) producing joint indices for temperate areas.
- Add functionality to provide estimates of relative observation error (CIs) by time period.
- Extend the approach to albacore standardization, i.e. cluster analyses and CPUE standardization with appropriate spatial structures.
- Thoroughly check all code and results in order to validate indices.
- All work is subject to the agreement of the respective fisheries agencies to make the data available.
- To document the analyses in accordance with the IOTC "*Guidelines for the presentation of CPUE standardisations and stock assessment models*", adopted by the IOTC Scientific Committee in 2014; and to provide draft reports to the IOTC Secretariat no later than 60 days prior to the meetings of the WPTmT06, i.e. **18 May 2016**, and WPTT18, i.e. **6 September 2016**, and the final report no later than 15 days prior to the meeting of the WPTT18, i.e. **21 October 2016**.
- To undertake any additional analyses deemed relevant by the WPTT18 or the IOTC Secretariat up to 60 days after the start date of the contract.

# Methods

## Data cleaning and preparation

The three datasets had many similarities but also significant differences. The variables differed somewhat among datasets, as did other aspects such as the sample sizes, the data coverage and the natures of the fleets.

Data preparation and analyses were carried out using R version 3.3.0 (R Core Team 2016).

The approaches used here are based on those applied by Chambers and Hoyle (2015), with modifications where required. For more detail about the Japanese, Korean, and Taiwanese fleets, see the descriptive figures in the following papers (Hoyle & Okamoto 2015; Hoyle et al. 2015a)

### Data

In this section we describe the datasets provided by Japanese, Taiwanese, and Korean data managers, and the methods that we used to prepare and clean the data for analysis. As the provided datasets were prepared for this collaborative study, the data do not include all information potentially included in logbook data. The cleaning described here differs from the standard cleaning procedures by national scientists when producing CPUE indices.

Japanese data were available from 1952-2015 (Figure 2), with fields year, month and day of operation, location to 1° of latitude and longitude, vessel call sign, no. of hooks between floats, number of hooks per set, date of the start of the fishing cruise, and catch in number of southern bluefin tuna, albacore, bigeye, yellowfin, swordfish, striped marlin, blue marlin, and black marlin (Table 1 and Table 2).

The Taiwanese operational data were available 1979-2015(Figure 3), but data prior to 2005 were not used in tropical tuna analyses, due to concerns about data quality applying to bigeye tuna in particular (see details in Chambers and Hoyle (2015)). Available fields were year, month and day of operation; vessel call sign; operational area (a code indicating fishing location at 5° resolution); operation location at 1° resolution (from 1994); number of hooks between floats (from 1995); number of hooks per set; catches in number for the species albacore, bigeye, yellowfin, bluefin (from 1993), southern bluefin (from 1994), other tuna, swordfish, striped marlin, blue marlin, black marlin, other billfish, skipjack, shark, and other species; equivalent values in weight for all species; SST; bait type fields for 'Pacific saury', 'mackerel', 'squid', 'milkfish', and 'other'; depth of hooks (m); set type (type of target, from 2006); remarks (indicating outliers); departure date from port; starting date of operations on a trip; stopping date of operations on a trip; arrival date at port (Table 3: Data format for Taiwanese longline dataset, and Table 6).

Korean operational data were available for 1971 to 2015 (Table 7, Figure 4), with fields vessel id, operation date, operation location to 1°, number of hooks, number of floats, and catch by species in number for albacore, bigeye, black marlin, blue marlin, striped marlin, other species, southern bluefin, sailfish, shark, skipjack, swordfish, and yellowfin.

The contents and preparation of logbook data is described below for each variable. See Table 8 for a comparison of field availability among the three fleets.

In the Japanese data international call sign was available 1979 - present, and was selected as the vessel identifier. Call sign is unique to the vessel and held throughout the vessel's working life. In the Taiwanese data, the international call sign was available for each set, and was also selected as the vessel identifier. The first digit of the Taiwanese callsign indicated the tonnage of the vessel (Table 4). In the Korean data the callsigns were understood to have changed through time to some extent, and so vessel ids were assigned based on a combination of vessel names and vessel callsigns. For all fleets, the vessel id was rendered anonymous by changing it to an arbitrary integer. Sets without a vessel call sign were allocated a vessel id of '1'. For joint analyses, a fleet code was added to differentiate vessels from different fleets.

In all Japanese and Korean data, and in most Taiwanese data from 1994, latitude and longitude were reported at 1° resolution, with a code to indicate north or south, west or east. Taiwanese fishing locations were otherwise reported at 5° square resolution using a logbook code. All data were adjusted to represent the south-western corner of the 1 x 1° square, and longitudes translated into 360° format. Each set was allocated to regions according to various alternative region definitions, including 2 definitions for yellowfin (Langley 2015), 3 for bigeye (Langley et al. 2013), and 6 for albacore. Data outside these areas were ignored. Location information was used to calculate the 5° square (latitude and longitude).

Hooks per set were reported in all datasets, and the few sets without hooks were deleted. For the purposes of further analyses, we cleaned the data by removing data likely to be in error. The criteria were selected after discussion with experts in the respective datasets. In the Japanese and Korean data, hooks per set above 5000 and less than 200 were removed. In the Taiwanese data hooks per set over 4500 and less than 200 were removed. The difference between fleets was unintentional, but there were very few sets with 4500-5000 sets, so there was little or no impact on results. A very high proportion of Taiwanese sets reported 3000 hooks per set, to an increasing degree through time. This difference from the other fleets and remarkable uniformity may be genuine, or may indicate a reporting problem, and warrants further investigation.

The three fleets all reported catch by species in numbers, but for slightly different species. The Japanese reported bigeye, yellowfin, albacore, southern bluefin tuna, swordfish, striped marlin, blue marlin, black marlin. The Taiwanese reported all these but included fields for skipjack, bluefin, sharks, other tunas, other billfish, and other species. The Taiwanese also reported catch by species in weight, but we used only the number information. Korea reported the same species as Japan and also skipjack, sailfish, sharks, and other species. The sailfish category may include shortbill spearfish (Uozumi 1999).

In the Taiwanese logbook, columns for bluefin and southern bluefin tuna were added in 1994. Prior to this bluefin were only recorded in the database when individuals changed the heading in the logbook. The number of reported bluefin increased substantially in 1994. We reassigned any fish reported as bluefin to the southern bluefin tuna category. The field labelled 'white marlin' represents striped marlin in the Indian Ocean. With the three fields for 'other' species, 'other tunas' are thought to be mostly neritic tunas, 'other billfish' may represent mostly sailfish and possibly shortbill spearfish, and 'other fish' particularly in recent years mostly oilfish.

In the logbooks of each fleet some very large catches were reported at times for individual species, but were not removed since there was anecdotal evidence that they may be genuine, and because they are unlikely to affect results substantially. Further investigation should consider the pros and cons of retaining these values.

In the Japanese logbook hooks between floats (HBF) were available for almost all sets 1971-2015 (Table 2), and for a high proportion of sets 1958-1966. Sets after 1975 with HBF missing or > 25 were removed. Sets before 1975 with missing HBF were allocated HBF of 5, according to standard practice with Japanese longline data (e.g. Langley et al. 2005; Hoyle et al. 2013; Ochi et al. 2014). In the Taiwanese logbook hooks between floats (HBF) were available from 1995. In the Korean logbook HBF was not available but the number of floats was reported, so we calculated HBF by dividing the number of hooks by the number of floats and rounding it to a whole number.

The remarks section of the Taiwanese dataset indicated outliers and other anomalies. Codes and criteria for outliers changed in 2012. Before 2012 an outlier was flagged if there was catch of more than 5 tons of a species per set, or outliers in the distribution of species catch number per set. From 2012 an outlier was flagged according to the 'IQR rule'. 1. Arrange average catch numbers per set (within a year) for all vessels in order. 2. Calculate first quartile (Q1), third quartile (Q3) and the interquartile range (IQR=Q3-Q1). 3. Compute Q1-1.5 x IQR and Compute Q3+1.5 x IQR. Anything outside this range is an outlier. This outlier information is used in the standard data cleaning procedures for Taiwanese standardisations. We did not use the outlier information in data cleaning for this paper.

After data cleaning, a standard dataset was produced for each fleet to be used in subsequent analyses (Figure 5).

Each set was allocated to bigeye, yellowfin, and albacore regions. These regions are based on the region definitions used in the stock assessments for each species. Several regional structures were explored for each species, but here we present one each for bigeye and yellowfin, and two for albacore (Figure 1). Data outside these regions were ignored. Subsequent analyses were performed separately for each region in each regional structure.

## Cluster analysis

Bigeye and yellowfin comprise a large proportion of the catch north of about 15° S, and a lower proportion further south (Figures 6 and 7). This pattern applied across all fleets, but there were also spatial and temporal differences in species composition patterns among fleets. The Taiwanese fishery included an oilfish fishery which developed from about 2005 in the southwest Indian Ocean (Figures 8 and 9).

We clustered the data using the approach applied by Hoyle et al. (2015b). We removed all sets with no catch of any of the species, and then aggregated by vessel-month. Set level data contains variability in species composition due to the randomness of chance encounters between fishing gear and schools of fish. This variability leads to some misallocation of sets using different fishing strategies. Aggregating the data tends to reduce the variability, and therefore reduce misallocation of sets. For these analyses we aggregated the data by vessel-month, assuming that individual vessels tend to follow a consistent fishing strategy through time. One trade-off with aggregation in this way is that vessels may change their fishing strategy within a month, which will result in misallocation of sets. For the purposes of this paper we refer to aggregation by vessel-month as trip-level aggregation, although the time scale is (for distant water vessels) in most cases shorter than a fishing trip. For Japanese data prior to 1979 vessel id was not available, but we were able to cluster them by vessel-month because the logbook id, available for the first time in the current data set, could be used to identify sets on the same vessel-trip.

We calculated proportional species composition by dividing the catch in numbers of each species by catch in numbers of all species in the vessel-month. Thus the species composition values of each vessel-month summed to 1, ensuring that large catches and small catches were given equivalent weight. The data were transformed by centring and scaling, so as to reduce the dominance of species with higher average catches. Centring was performed by subtracting the column (species) mean from each column, and scaling was performed by dividing the centred columns by their standard deviations.

We clustered the data using the hierarchical Ward hclust method, implemented with function hclust in R, option 'Ward.D', after generating a Euclidean dissimilarity structure with function 'dist'. This approach differs from the standard Ward D method which can be implemented by either taking the square of the dissimilarity matrix or using method 'ward.D2' (Murtagh & Legendre 2014). However in practice the method gives similar patterns of clusters to other methods, more reliably than ward.D2 (Hoyle et al 2015).

Data were also clustered using the kmeans method, which minimises the sum of squares from points to the cluster centres, using the algorithm of Hartigan and Wong (1979). It was implemented using function kmeans in the R stats package (R Core Team 2014).

### Selecting the number of groups

We used several subjective approaches to select the appropriate number of clusters. In most cases the approaches suggested the same or similar numbers of groups. First, we applied hclust to transformed trip-level data and examined the hierarchical trees, subjectively estimating the number of distinct branches. Second, we ran kmeans analyses on untransformed trip-level data with number of groups *k* ranging from 2 to 25, and plotted the deviance against *k*. The optimal group number was the lowest value of *k* after which the rate of decline of deviance became slower and smoother. Third, following Winker et al (2014) we applied the nScree() function from the R nFactors package (Raiche & Magis 2010), which uses various approaches (Scree test, Kaiser rule, parallel analysis, optimal coordinates, acceleration factor) to estimate the number of components to retain in an exploratory PCA. Where there was uncertainty about the number of clusters, we selected the option with more clusters.

### Plotting and data selection

We plotted the hclust clusters to explore the relationships between them and the species composition and other variables, such as HBF, number of hooks, year, and set location. Plots included boxplots of a) proportion of each species in the catch, by cluster; b) the distributions of variables by cluster; and c) maps of the spatial distribution of clusters, one map for each cluster.

In some analyses clusters that caught very few of the species of interest were omitted, because they provide little relevant information and may cause analysis problems due to large numbers of zeroes, and memory problems due to large sample sizes. Cluster selection was based on review and discussion of the plots of covariates and species compositions by cluster. Analyses were run both with and without these clusters – see the 'Models and datasets' section.

We pooled data from multiple fleets into a single analysis for years 1952-2015. The pooled dataset included all data from the Japanese (1952-2015) and Korean (1971-2015) fleets. For the Taiwanese fleet 1979-2015 were included for albacore, and 2005-2015 for tropical tunas.

For standardization of each region, data were selected for vessels that had fished for at least N1 quarters in that region. The standard level of N1 was 8 quarters in the equatorial regions and 2 quarters in the southern regions. Subsequently, vessels, 5° cells, and year-quarters were included if they had at least 100 sets. For analyses of the 1952-1979 period this criterion was reduced to 50 sets, to increase the size of the dataset. For datasets with more than 60,000 sets the number of sets in each stratum (5° square * year-quarter) was limited by randomly selecting 60 sets without replacement from strata with more than this number of sets. Testing suggested that this approach did not cause bias, and the effects on trends of random variation were reduced to very low levels at 30 sets per stratum (Hoyle & Okamoto 2011), suggesting that 60 sets was more than adequate.

## CPUE standardization, and fleet efficiency analyses

CPUE standardization methods generally followed the approaches used by Hoyle and Okamoto (2011) with some modifications. The operational data were standardized using generalized linear models in R. A large number of analyses were carried out.

1. Analyses were carried out for each species.
2. Initially analyses were carried out for multiple regional structures, though this was later reduced to one each for bigeye and yellowfin, and two for albacore (Figure 1).
3. Analyses for bigeye and yellowfin were conducted using five alternative models and datasets, described below.
4. Separate analyses were run for each region, ranging from one to four regions per structure.
5. Up to three modelling distributions were used: lognormal constant, delta lognormal, and negative binomial. Lognormal constant was used for all species, delta lognormal for bigeye and yellowfin, and negative binomial for albacore.
6. Analyses were run for four alternative data groups, as described below.

### Distributions

Lognormal constant analyses were carried out using generalized linear models that assumed a lognormal distribution. In this approach the response variable $\log(CPUE + k)$ was used, and a Normal distribution assumed. The constant *k*, added to allow for modelling sets with zero catches of the species of interest, was 10% of the mean CPUE for all sets.

Delta lognormal analyses (Lo et al. 1992; Maunder & Punt 2004) used a binomial distribution for the probability *w* of catch rate being zero and a probability distribution *f(y)* , where y was log(catch/hooks set), for non-zero (positive) catch rates. The index estimated for each year-quarter was the product of the year effects for the two model components, $(1 - w).E(y|y \neq 0)$.

$$\Pr(Y = y) = \begin{cases} w, & y = 0 \\ (1 - w)f(y) & otherwise \end{cases}$$

$\mathrm{g}(w) = (CPUE = 0) \sim covariates + \epsilon$, where *g* is the logistic function.

$\mathrm{f}(y) = CPUE \sim covariates + \epsilon$

Negative binomial analyses used the function glm.nb from the MASS package (Venables & Ripley 2002) in R, using the default options. The response variable was catch in numbers.

In each case the covariates included year-quarter (*yrqtr*), 5° cell (*latlong5*), and cluster (*cl*) fitted as categorical variables, and a cubic spline function *h* with 10 degrees of freedom applied to the continuous variable *hooks*. Some analyses included the vessel identifier *vessid* as a categorical variable. Some analyses included a cubic spline $\varphi$ applied to the continuous variable hooks between floats (*hbf*).

Data in all models except the binomial model were 'area-weighted', with the weights of the sets adjusted so that the total weight per year-quarter in each 5° square would sum to 1. This method was based on the approach identified using simulation by Punsly (1987) and Campbell (2004), that for set *j* in area *i* and year-quarter *t*, the weighting function that gave the least average bias was:

$w_{ijt} = \frac{log(h_{ijt}+1)}{\sum_{j=1}^{n} \log(h_{ijt}+1)}$. Given the relatively low variation in number of hooks between sets in a stratum, we simplified this to $w_{ijt} = \frac{h_{ijt}}{\sum_{j=1}^{n} h_{ijt}}$.

For the lognormal constant and positive lognormal GLMs, model fits were examined by plotting the residual densities and using Q-Q plots.

### Models and datasets

In order to explore alternative approaches to the analysis, the four approaches below were applied for each of the tropical tuna species. Albacore was modelled with the second approach only.

1. Data omitted low-target clusters. Model included HBF but not cluster.
2. Data omitted low-target clusters. Model included cluster but not HBF.
3. Data omitted low-target clusters. Model included neither HBF nor cluster.
4. All data included. Model included HBF but not cluster.

### Data periods

Vessel identity information was only available from 1979, so could not be applied uniformly across all years. The discontinuity in 1979 could be addressed in several different ways. We therefore analysed the data in several ways so as to provide the assessment scientists with appropriate data. For each of the approaches above, four analyses were carried out (Table 1).

*Table 1: Analysis approaches for addressing the discontinuity in availability of vessel identity.*

| Analysis | Years | Vessel effects |
|---|---|---|
| 1 | 1952-1979 | No |
| 2 | 1979-2015 | Yes |
| 3 | 1952-2015 | No |
| 4 | 1952-2015 | Yes |

It is possible to standardize the time series with vessel effects by assigning an identical dummy value to all vessels without vessel identity information. This was done for analysis 3). However using a dummy value introduces several problems. First, not all vessels begin to report their callsign at once in 1979, and those that do are self-selected and not randomly selected from the vessel population. Therefore it cannot be assumed that fishing power remains constant after 1979 for the dummy vessel id, so the transition in 1979 may introduce a discontinuity into the time series. The discontinuity can be limited in scope by restricting the overlap between dummy and real vessel IDs to one year – 1979 – and removing sets with missing vessel IDs after this time. Secondly, residuals may be more variable before 1979, without a true vessel ID in the model, which can introduce bias into the standardization.

One approach for addressing the discontinuity in analysis 3) is to adjust the time period 1952-1978 so that the relative averages in 1978 and 1979 are the same as they are in analysis 4), without vessel effects. However we considered that a better approach may be to estimate two time series 1952-1979 without vessel effects, and a second time series 1979-2015 with vessel effects (omitting all sets without vessel IDs). These are analyses 1) and 2) above. Subsequently the analyst can use them as desired, for example concatenating them after adjusting the averages so that the estimates for 1979 are the same.

### Covariate effects

The effects of covariates were examined by plotting the predicted effects, with 95% confidence limits, of each parameter at observed values of the explanatory variables. Spatial effects with 95% confidence intervals were plotted by latitude. The cumulative vessel effects through time were examined by plotting each vessel's effect at every time that vessel made a set. An average vessel effect over time was examined by calculating the mean of the vessel effects for all sets made by the fleet during each time period, and this was also plotted. There was insufficient space to include all plots in the report, but these are available on request.

Changes in catchability through time were investigated by fitting to the operational data both with and without a term for individual vessel. The two models were designated respectively the 'base model' and the 'vessel-effects model'. Abundance indices were calculated for each model, and normalized to average 1.

For all model comparisons, the indices estimated for each year-quarter were compared by dividing the base model by the vessel effects model, plotting the time series of ratios, and fitting a log-linear regression. The slope of the regression represented the average annual compounding rate of change in fishing power attributable to changes in the vessel identities; i.e. the introduction of new vessels and retirement of old vessels. Gradients are shown on the figures, together with confidence intervals.

### Indices of abundance

Indices of abundance were obtained by applying the R function predict.glm to model objects. Binomial time effects were obtained by generating time effects from the glm and adjusting them so that their mean was the proportion of positive sets across the whole dataset. The main aim with this approach is to obtain a CPUE that varies appropriately, since variability for a binomial is greater when the mean is at 0.5 than at 0.02 or 0.98, and the multiplicative effect of the variability is greater when the mean is lower. The outcomes were normalised and reported as relative CPUE with mean of 1.

Uncertainty estimates were provided by applying the R function predict.glm with type = ”terms” and se.fit=TRUE, and taking the standard error of the year-quarter effect. For the delta lognormal models we used only the uncertainty in the positive component. Uncertainty estimates from standardizing commercial logbook data are in general biased low and often ignored by assessment scientists, since they assume independence and ignore autocorrelation associated with (for example) consecutive sets by the same vessels in the same areas. There may be a very large mismatch between the observation error in CPUE indices and the process error in the indices that is estimated in the assessment. This is particularly true for distant water longline CPUE, where very large sample sizes generate small observation errors.

Residual distributions and Q-Q plots were produced for all but the binomial analyses. For the lognormal positive analyses that included cluster in the model, median residuals were plotted by cluster. For all lognormal positive analyses, residuals by year-quarter were plotted by flag; median residuals by year-quarter were plotted by flag; and median residuals by 5° cell were mapped onto a contour plot for each flag.

We compared the indices with the area-specific Japanese bigeye indices from 2013 (Matsumoto et al. 2013) and yellowfin indices from 2015 (Ochi et al. 2015). The 2013 bigeye indices provided only a whole-of-area index in the southern temperate area, so this was compared with both the east and west joint indices. For each comparison, each dataset was first normalised by dividing through by its mean for 1980-2000, and the datasets plotted on the same axes. Secondly, the joint indices were divided by the matching year-quarter values from the Japanese indices, and these ratios were plotted to show the relative trends of the two time series.

# Results

## Cluster analysis

The aim of the cluster analysis was first to identify separate fishing strategies in the data for each species, regional structure, fleet, and region, and so to better understand the fishing practices; and second to assign each unit of fishing effort to a particular fishing strategy, so that the clusters could be used in standardization.

We clustered the data using hclust and kmeans methods for each region and fleet. The hclust trip and untransformed kmeans set methods separated Japanese, Korean and Taiwanese effort into 3-5 fishing strategies in each region (Table 10, Figures 10 - 17). Please note that the order of the clusters in the dendrograms does not match the cluster numbers.

Species compositions were plotted by cluster for each region and fleet, as were the relative distributions of covariates (Figures 18 - 40).

Clusters with low levels of the target species were excluded from standardization datasets 1-3. Clusters retained in these datasets are reported in Table 11.

For bigeye in region 1 we excluded 2 Japanese clusters and 1 Taiwanese cluster that were characterized by a low proportion of bigeye, and either high albacore or high yellowfin (Figure 18). The excluded clusters derived mostly from southern areas (Figures 19 and 20).

For bigeye in region 2 only one Japanese and one Taiwanese cluster were excluded, and no clusters for Korea. The excluded Japanese cluster had high catch of southern bluefin tuna (Figure 21) mostly from the period before 1970 (Figure 22). The excluded Taiwanese cluster included mostly albacore (Figure 21), and was from the far south of the region (Figure 23).

For bigeye in region 3 only 2 Japanese, 1 Korean, and 2 Taiwanese clusters were included in the dataset, and even these had relatively low proportions of bigeye (Figure 24). Data covered the entire period though with low sample sizes in the early period (Figure 25), and the effort covering most of the area except the Mozambique channel (Figure 26).

For bigeye in region 4, there were 2 Japanese clusters selected but all Korean and Taiwanese clusters were included. The excluded Japanese clusters caught mostly yellowfin and southern bluefin tuna

(Figure 27). Sample sizes were relatively low before 1990 (Figure 28). Effort was distributed broadly across the region (Figure 29).

For yellowfin in region 2 , all Korean and Taiwanese clusters were included. One Japanese cluster was excluded that caught mostly bigeye (Figure 30) in the period after 1980 (Figure 31). The excluded cluster was found more in the northeast of the region (Figure 32).

For yellowfin in region 3, two Japanese clusters targeting albacore and southern bluefin tuna, one Korean cluster targeting albacore and southern bluefin tuna, and one Taiwanese cluster targeting other species (oilfish and escolar) were excluded. The remaining data had relatively few sets before 1980 (Figure 34). The covered most of the region, except for the far south (Figure 35).

For yellowfin in region 4, one Japanese, two Korean and 1 Taiwanese cluster were included. The Japanese data covered the whole time series (Figure 37) but was mostly from the northwest and around the northern Australian coast, and the Taiwanese data was mostly from the north (Figure 38).

For yellowfin in region 5, all Japanese, one of the three Korean, and all Taiwanese clusters were included.

## CPUE indices

We estimated indices for all regions of regional structures B2 (Figures 42 - 45), Y (Figures 46 - 49), and for the two albacore indices. Diagnostics for the lognormal positive distribution indicated some negative skewness in the distributions of residuals (Figures 50 - 53). Figures are presented here for the bigeye and yellowfin only.

In tropical areas (bigeye regions 1 and 2, and yellowfin regions 2 and 5) we have selected figures from analysis 1, which omits low-target clusters from the dataset, and includes HBF but not cluster in the model. In temperate areas (bigeye and yellowfin regions 3 and 4) we have selected figures from analysis 4, which omits low-target clusters from the dataset, and conversely includes cluster but not HBF in the model. In southern regions there are well known differences in fishing behaviour between vessels targeting different species, and these differences are reflected in the species composition, making it appropriate to use cluster in the standardization model. For example, the Japanese southern bluefin tuna fishery takes largely SBT, with some catch of albacore. The Taiwanese oilfish fishery is also a clear example, with a very high representation of species ‘other’.

In tropical areas however, although there have been changes in targeting through time, vessels target bigeye and yellowfin at the same time and using similar methods, but to different extents by area and season, and with changes through time. In this complex situation clustering may be useful to remove data from clearly separate fisheries (such as the southern bluefin tuna fishery that occurred in eastern areas near Indonesia in the 1960s and 70s). However including cluster in the model may be problematic due to confounding with abundance change. We have therefore used hooks between floats in the models for tropical areas, as was done in previous years’ analyses. Unlike previous analyses, we excluded clusters with minimal catch of the species of interest.

In reporting results we focus mainly on the long term indices that include vessel effects, in the top-right quadrant of each set of figures.

For bigeye tuna the long term tropical indices in regions 1 (Figure 42) and 2 (Figure 43) show a decline followed by a sharp increase in the late 1970s (see below). The declining trends subsequently resume and continue until about 2010. The CPUE is estimated to be currently at its lowest observed level in the western tropical region 1, and slightly above its lowest level in the eastern tropical region 2.

For bigeye in western temperate region 3, data are relatively sparse in the period before 1990, due to low sample sizes in the original data, and omission of sets with very low bigeye species composition. For the period since 1990 the indices show a similar pattern to the northern indices, with declining CPUE overall, but a suggestion of some increase since 2010. In eastern temperate region 4 the pattern is very different and highly variable. Standardized CPUE increases during the 1960s to a peak in the early 1970s, after which it drops until 1990, increases until 2000, declines again until about 2010, and then increases again.

For yellowfin tuna, indices in the northern areas were characterized by very steep declines in standardized CPUE prior to 1975 (Figure 46). From 1980-1989 the north-western region 2 CPUE increased, then declined until 1995, increased again until 2005, and then decreased again. Since that time it has remained close to the lowest level observed. The north-eastern region 5 shows a similar pattern until 1990 but has declined steadily since that time, and is currently also close to the lowest level in the time series (Figure 49).

Yellowfin in western temperate region 3 show a similar pattern to the western tropical indices, with a decline until the mid-1970s followed by an increase until the late 1980s, decline until 2000, and more stable CPUE since then. In eastern temperate region 4 the pattern is, as for bigeye, very different and highly variable. Indices are very sparse due to low sample sizes and it is difficult to draw conclusions.

For the analyses that included cluster, median residuals by cluster from the lognormal constant models showed some trends, both in the recent past and over the long term (Figures 54 - 61). Recent residuals were less problematic in the time series that included vessel effects, suggesting that some of the patterns are due to changes in fleet composition. However some patterns and long-term trends remained, with stronger trends in the tropical areas. Median residuals by fleet were both mapped and plotted by year-quarter (Figures 62 - 65). There were clear spatial patterns with differences between fleets. Similarly, temporal trends were observed within the fleets.

Comparisons of the joint indices with indices developed in 2013 (bigeye) and 2015 (yellowfin) for the Japanese fleet showed differing trends, with less decline in both tropical yellowfin indices and in the western tropical bigeye indices, but the converse in the eastern tropical bigeye tuna (Figures 66 - 69).

## Discussion

An important feature of these indices is the discontinuity in CPUE in about 1978 for bigeye in tropical areas 1 and 2. It also appears in the indices without vessel effects for yellowfin in the north-eastern region 5 and there is a sharp and short-lived peak in north-western area 2. Similar patterns are observed in the 2013 and 2015 Japanese indices. This discontinuity occurred during a period with little effort in northern areas (Figure 2) and an almost complete change in hooks between floats in the Japanese fishery, from HBF less than 8 to greater than 8 (Hoyle & Okamoto 2015, Figure 7). Indices for the northern areas use HBF in the standardization model, and the limited overlap suggests that there may be some confounding between the year-quarter effects and the HBF effects. However, results for model 3 (with neither HBF nor cluster in the model) show almost the same level of discontinuity, suggesting that there is a genuine change in catch rate at this time. The change in HBF strongly suggests a change in fishing behaviour, and given the low effort at this time there may have been a compete change in the fleet. However, vessel identities are not currently available for this period. Provision of vessel identity information is therefore considered to be a high priority.

A relatively high proportion of effort was removed by data cleaning during the period of the discontinuity (Figure 5). The reasons for the removal are currently unclear and should be further investigated by those with access to the data.

The CPUE trends estimated here address a number of concerns about indices used in previous assessments. Models are run separately for different areas, which addresses concerns about differing parameter estimates and uncertainty distributions in different areas (Chang et al. 2011). The models use 5° cell area effects, as recommended by the 2013 IOTC CPUE workshop (Anon 2013) to account for changes in effort distribution, and adjusts statistical weights to allow for changing effort concentration (Punsly 1987; Campbell 2004). The models include vessel effects, which accounts for some effects of changing fishing power and targeting within the fleet (Hoyle & Okamoto 2011). It also uses cluster analysis based on species composition in order to identify target change, and to separate out effort using different fishing strategies (He et al. 1997).

CPUE trends are affected by the abundance of the species of interest, but also by catchability factors such as targeting, the fishing power of the fleet, and environmental changes, and by changes in species distribution. The spatial distribution of effort also affects CPUE because there is no information about areas without data.

We assume uniform population trends within regions, but permit varying population trends between regions. Regional models are completely independent, which is appropriate given the potential for differing error distributions and covariate relationships. Within each region the models allow for spatial variation in catch rates at the 5° cell level, but assume that the spatial variation is consistent through time, without time-area interactions. Thus areas without data are assumed to follow the same trends as areas with data. This approach would be reasonable for a well-mixed population, or for a population in which areas without data were fished by fleets not included in the dataset. These factors may be true to some extent for Indian Ocean tuna, but mixing at a large scale is probably limited (Kolody & Hoyle 2014), and reduced fishing pressure in defined areas (e.g. due to piracy in the area near Somalia) may lead to spatial variation in population density.

The spatial and temporal patterns observed in the residuals by fleet may reflect time-area, time-fleet, and/or time-vessel group interactions that are not accommodated by the models. There may be changes in targeting and fishing practices that are not adequately accounted for by the models, or biases in fleets' reporting of effort. It may be useful to include time area interactions in the models, but to be estimable these subareas would need to be at a larger spatial scale than 5° cells. Further exploring residual patterns would help to determine the appropriate scale and locations for subareas. It would also be useful to explore spatial and temporal residual patterns among fleets and vessels.

Concerns remain about the indices estimated in this study. The declines in the indices before 1970 are too steep to represent abundance change, given the relatively low catches taken during this period. Similar declines are seen in other oceans (e.g. Hampton et al. 2005), even after clustering (Bigelow & Hoyle 2012). Factors causing the declines are unclear, but in addition to unresolved effects of target change may include changing catchability due to removal of the most vulnerable individuals (Gulland 1974; Maunder et al. 2006). The discontinuities discussed above also require resolution.

Some of the southern indices show different CPUE trends from the northern areas, which is likely to reflect the fact that vessels are targeting other species and bigeye and yellowfin tuna are bycatch species. Cluster analysis has not fully accounted for target change, and these indices may be biased. Further investigation is needed to explore this issue, which should include investigating residuals by fleet, the effects of piracy on fleet distribution, exploring the timing of the changes seasonally, and possible relationships with target switching by the southern bluefin tuna fleet after quotas have been met.

CPUE indices are very influential components of stock assessment models, and further work to improve and validate indices is a high priority. We suggest the following priorities for further work:

1) Develop simulator to test methods for standardizing CPUE, and to allow the development and testing of new code during periods when the joint data are unavailable.
2) Investigate the 1976-80 discontinuity in the tropical CPUE of bigeye and (to a lesser extent) yellowfin.
3) Explore options for extending the Japanese time series of vessel effects into the pre-1979 period.
4) Increase understanding of the fisheries that provide the CPUE by a) exploring the size data associated with each fleet, if possible with size data at the vessel set level; and b) exploring vessel movement patterns through time.
5) Develop standard methods for estimating relative regional weights so as to apportion relative abundance among regions.
6) Explore alternative modelling and data transformation methods in order to normalise residuals and to accommodate strata with no zero catches.
7) Develop separate indices for each fleet.
8) Add subarea-time interactions to the standardization models, to address differences in trends among areas. Explore residual patterns spatially and among clusters, fleets and vessels through time, and change models where necessary to address any problems identified. Develop additional residual and exploratory plots to explore possible confounding effects, such as maps of residuals by season to explore seasonal catchability changes.
9) Test alternative methods for identifying and accounting for targeting.

## Acknowledgments

Thanks to the International Seafood Sustainability Foundation (ISSF) and the Indian Ocean Tuna Commission (IOTC) for funding this work. We are grateful to the IOTC for facilitating, and particularly David Wilson and James Geehan. Thanks to the Taiwanese Fisheries Agency, Taiwanese Overseas Fisheries Development Council, and the Japanese Institute for Far Seas Fisheries at Shimizu for providing their facilities and support. Thanks to Ren-Fen Wu and Lisa Chang for their thoughtful contributions and organizational support. Thanks also to Shanghai Ocean University for their facilities and support.

## References


Anon 2013. Report of the IOTC CPUE Workshop, San Sebastian, Spain, 21–22 October, 2013 IOTC–2013–SC16–12[E]. 32 pp p.

Bigelow KA, Hoyle SD 2012. Standardized CPUE for South Pacific albacore WCPFC-SC8-SA-IP-14.

Campbell RA 2004. CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models. Fisheries Research 70(2-3): 209-227.

Chambers M, Hoyle SD 2015. Estimates of non-member catch of SBT in the Indian and Pacific Oceans, CCSBT-ESC/1509/10. 20th Extended Scientific Committee of the CCSBT. Incheon, Republic of Korea.

Chang S-K, Hoyle S, Liu H-I 2011. Catch rate standardization for yellowfin tuna (Thunnus albacares) in Taiwan's distant-water longline fishery in the Western and Central Pacific Ocean, with consideration of target change. Fisheries Research 107(1-3): 210-220.

Gulland JA 1974. Catch per unit effort as a measure of abundance. ICCAT Collective Volume of Scientific 3: 1-11.

Hampton J, Sibert JR, Kleiber P, Maunder MN, Harley SJ 2005. Fisheries: decline of Pacific tuna populations exaggerated? Nature 434(7037): E1-2; discussion E2.

Hartigan JA, Wong MA 1979. Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society.Series C (Applied Statistics) 28(1): 100-108.

He X, Bigelow KA, Boggs CH 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fisheries Research 31(1-2): 147-158.

Hoyle S, Davies N, Chang S-K 2013. Analysis of swordfish catch per unit effort data for Japanese and Chinese Taipei longline fleets in the southwest Pacific Ocean, WCPFC-SC9-2013/SA-IP-03. WCPFC Scientific Committee, Ninth Regular Session, 7-15 August 2012, Busan, Republic of Korea.

Hoyle SD, Okamoto H 2011. Analyses of Japanese longline operational catch and effort for bigeye and yellowfin tuna in the WCPO, WCPFC-SC7-SA-IP-01. Western and Central Pacific Fisheries Commission, 9th Scientific Committee. Pohnpei, Federated States of Micronesia.

Hoyle SD, Okamoto H 2015. Descriptive analyses of the Japanese Indian Ocean longline fishery, focusing on tropical areas. 54 pp. p.

Hoyle SD, Yeh Y-M, Chang S-T, Wu R-F 2015a. Descriptive analyses of the Taiwanese Indian Ocean longline fishery, focusing on tropical areas. 84 pp. p.

Hoyle SD, Okamoto H, Yeh Y-m, Kim ZG, Lee SI, Sharma R 2015b. IOTC–CPUEWS02 2015: Report of the 2nd CPUE Workshop on Longline Fisheries, 30 April – 2 May 2015. 126 p.

Kolody D, Hoyle S 2014. Evaluation of tag mixing assumptions in western Pacific Ocean skipjack tuna stock assessment models. Fisheries Research.

Langley A 2015. Stock assessment of yellowfin tuna in the Indian Ocean using Stock Synthesis, IOTC–2015–WPTT17–30. IOTC Working Party on Tropical Tunas.

Langley A, Herrera M, Sharma R 2013. Stock assessment of bigeye tuna in the Indian Ocean for 2012. IOTC Working Party Document.

Langley A, Bigelow K, Maunder M, Miyabe N 2005. Longline CPUE indices for bigeye and yellowfin in the Pacific Ocean using GLM and statistical habitat standardisation methods. WCPFC-SC1, Noumea, New Caledonia SA WP-8. 8-19 p.

Lo NCH, Jacobson LD, Squire JL 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. Canadian Journal of Fisheries and Aquatic Sciences 49(12): 2515-2526.

Matsumoto T, Satoh K, Okamoto H 2013. Japanese longline CPUE for bigeye tuna in the Indian Ocean standardized by GLM IOTC-2013-WPTT15-25. 28 p.

Maunder MN, Punt AE 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70(2-3): 141-159.

Maunder MN, Sibert JR, Fonteneau A, Hampton J, Kleiber P, Harley SJ 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities. Ices Journal of Marine Science 63: 1373-1385.

Murtagh F, Legendre P 2014. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification 31(3): 274-295.

Ochi D, Matsumoto T, Satoh K, Okamoto H 2014. Japanese longline CPUE for bigeye tuna in the Indian Ocean standardized by GLM. IOTC–2014–WPTT16–29 Rev_1. 28 p.

Ochi D, Matsumoto T, Nishida T, Kitakado T 2015. Update of standardized Japanese longline CPUE for yellowfin tuna in the Indian Ocean and consideration of standardization methods IOTC–2015–WPTT17–26. 53 p.

Punsly R 1987. Estimation of the relative annual abundance of yellowfin tuna, Thunnus albacares , in the eastern Pacific Ocean during 1970-1985. LA JOLLA, CA ( ), I-ATTC.

R Core Team 2014. R: A Language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

R Core Team 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Raiche G, Magis D 2010. nFactors: Parallel analysis and non graphical solutions to the Cattell Scree Test. R package version 2(3).

Uozumi Y 1999. BBRG-6. Review of Problems on Stock Assessment of Marlins Laying Stress on the Coverage of landing and Catch and Effort Information in the Pacific Ocean. 12th Standing Committee on Tuna and Billfish (SCTB). Tahiti, French Polynesia. Pp. 9 pages.

Venables WN, Ripley BD 2002. Modern Applied Statistics with S, Springer.

Winker H, Kerwath SE, Attwood CG 2014. Proof of concept for a novel procedure to standardize multispecies catch and effort data. Fisheries Research 155: 149-159.

# Tables

*Table 2: Data format for Japanese longline dataset.*

| Items | Type | 1952-1957 | 1959-1966 | 1967-1975 | 1976-1993 | 1994-2015 |
|---|---|---|---|---|---|---|
| operation year | integer | YES | YES | YES | YES | YES |
| operation month | integer | YES | YES | YES | YES | YES |
| operation day | integer | YES | YES | YES | YES | YES |
| operation latitude | integer | YES | YES | YES | YES | YES |
| operation latitude code | integer | YES | YES | YES | YES | YES |
| operation longitude | integer | YES | YES | YES | YES | YES |
| operation longitude code | integer | YES | YES | YES | YES | YES |
| call sign | character | NO | NO | NO | YES | YES |
| no. of hooks between float | integer | NO | YES | NO | YES | YES |
| total no. of hooks per set | integer | YES | YES | YES | YES | YES |
| SBT catch in number | integer | YES | YES | YES | YES | YES |
| albacore catch in number | integer | YES | YES | YES | YES | YES |
| bigeye catch in number | integer | YES | YES | YES | YES | YES |
| yellowfin catch in number | integer | YES | YES | YES | YES | YES |
| swordfish catch in number | integer | YES | YES | YES | YES | YES |
| striped marlin catch in number | integer | YES | YES | YES | YES | YES |
| blue marlin catch in number | integer | YES | YES | YES | YES | YES |
| black marlin catch in number | integer | YES | YES | YES | YES | YES |
| shark catch in number | Integer | YES | YES | YES | YES | YES |
| prefecture code | character | YES | YES | YES | YES | YES |
| logbook ID | integer | YES | YES | YES | YES | YES |
| day of cruise start | integer | NO | YES | NO | YES (79-93) | YES |

*Table 3: Number of available data by variable in the Japanese longline dataset.*

| YEAR | No. of operation | Operation Date | Latitude | Longitude | Call sign | HBF | Total number of hooks per set | SBT catch in number | ALB catch in number | BET catch in number | YFT catch in number | SWO catch in number | MLS catch in number | BUM catch in number | BLA catch in number | day of cruise start |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1952** | 136 | 136 | 136 | 136 | 0 | 0 | 136 | 136 | 136 | 136 | 136 | 136 | 136 | 136 | 136 | 0 |
| **1953** | 1065 | 1065 | 1065 | 1065 | 0 | 0 | 1065 | 1065 | 1065 | 1065 | 1065 | 1065 | 1065 | 1065 | 1065 | 0 |
| **1954** | 4289 | 4289 | 4289 | 4289 | 0 | 0 | 4289 | 4289 | 4289 | 4289 | 4289 | 4289 | 4289 | 4289 | 4289 | 0 |
| **1955** | 6411 | 6411 | 6411 | 6411 | 0 | 0 | 6411 | 6411 | 6411 | 6411 | 6411 | 6411 | 6411 | 6411 | 6411 | 0 |
| **1956** | 11293 | 11293 | 11293 | 11293 | 0 | 0 | 11293 | 11293 | 11293 | 11293 | 11293 | 11293 | 11293 | 11293 | 11293 | 0 |
| **1957** | 7833 | 7833 | 7833 | 7833 | 0 | 99 | 7833 | 7833 | 7833 | 7833 | 7833 | 7833 | 7833 | 7833 | 7833 | 103 |
| **1958** | 8149 | 8149 | 8149 | 8149 | 0 | 6055 | 8149 | 8149 | 8149 | 8149 | 8149 | 8149 | 8149 | 8149 | 8149 | 7086 |
| **1959** | 9983 | 9983 | 9983 | 9983 | 0 | 7048 | 9983 | 9983 | 9983 | 9983 | 9983 | 9983 | 9983 | 9983 | 9983 | 9111 |
| **1960** | 13701 | 13701 | 13701 | 13701 | 0 | 10139 | 13701 | 13701 | 13701 | 13701 | 13701 | 13701 | 13701 | 13701 | 13701 | 12546 |
| **1961** | 12553 | 12553 | 12553 | 12553 | 0 | 10103 | 12553 | 12553 | 12553 | 12553 | 12553 | 12553 | 12553 | 12553 | 12553 | 11655 |
| **1962** | 22365 | 22365 | 22365 | 22365 | 0 | 11759 | 22365 | 22365 | 22365 | 22365 | 22365 | 22365 | 22365 | 22365 | 22365 | 21195 |
| **1963** | 23315 | 23315 | 23315 | 23315 | 0 | 11397 | 23315 | 23315 | 23315 | 23315 | 23315 | 23315 | 23315 | 23315 | 23315 | 23278 |
| **1964** | 28868 | 28868 | 28868 | 28868 | 0 | 13686 | 28865 | 28868 | 28868 | 28868 | 28868 | 28868 | 28868 | 28868 | 28868 | 28868 |
| **1965** | 28631 | 28631 | 28631 | 28631 | 0 | 25152 | 28631 | 28631 | 28631 | 28631 | 28631 | 28631 | 28631 | 28631 | 28631 | 28631 |
| **1966** | 32773 | 32773 | 32272 | 32773 | 0 | 31574 | 32773 | 11057 | 32773 | 32773 | 32773 | 32773 | 19904 | 17978 | 13959 | 32773 |
| **1967** | 58000 | 58000 | 57853 | 58000 | 0 | 9215 | 58000 | 51436 | 58000 | 58000 | 58000 | 58000 | 53732 | 53166 | 51628 | 9343 |
| **1968** | 40033 | 40033 | 40033 | 40033 | 0 | 0 | 40033 | 40033 | 40033 | 40033 | 40033 | 40033 | 40033 | 40033 | 40033 | 0 |
| **1969** | 36172 | 36172 | 36172 | 36172 | 0 | 0 | 36172 | 36172 | 36172 | 36172 | 36172 | 36172 | 36172 | 36172 | 36172 | 0 |
| **1970** | 29393 | 29393 | 29393 | 29393 | 0 | 0 | 29393 | 29393 | 29393 | 29393 | 29393 | 29393 | 29393 | 29393 | 29393 | 0 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1971** | 27402 | 27402 | 27402 | 27402 | 0 | 26248 | 27402 | 27402 | 27402 | 27402 | 27402 | 27402 | 27402 | 27402 | 27402 | 0 |
| **1972** | 21220 | 21220 | 21220 | 21220 | 0 | 20571 | 21220 | 21220 | 21220 | 21220 | 21220 | 21220 | 21220 | 21220 | 21220 | 0 |
| **1973** | 24968 | 24968 | 24968 | 24968 | 0 | 24036 | 24968 | 24968 | 24968 | 24968 | 24968 | 24968 | 24968 | 24968 | 24968 | 0 |
| **1974** | 28492 | 28492 | 28492 | 28492 | 0 | 27700 | 28492 | 28492 | 28492 | 28492 | 28492 | 28492 | 28492 | 28492 | 28492 | 0 |
| **1975** | 30287 | 30287 | 30287 | 30287 | 0 | 29062 | 30287 | 30287 | 30287 | 30287 | 30287 | 30287 | 30287 | 30287 | 30287 | 0 |
| **1976** | 26590 | 26590 | 26590 | 26590 | 0 | 26039 | 26590 | 26590 | 26590 | 26590 | 26590 | 26590 | 26590 | 26590 | 26590 | 0 |
| **1977** | 22150 | 22150 | 22150 | 22150 | 0 | 21780 | 22150 | 22150 | 22150 | 22150 | 22150 | 22150 | 22150 | 22150 | 22150 | 0 |
| **1978** | 22530 | 22530 | 22530 | 22530 | 0 | 22080 | 22530 | 22530 | 22530 | 22530 | 22530 | 22530 | 22530 | 22530 | 22530 | 0 |
| **1979** | 28551 | 28551 | 28551 | 28551 | 27857 | 23552 | 28551 | 28551 | 28551 | 28551 | 28551 | 28551 | 28551 | 28551 | 28551 | 28551 |
| **1980** | 31506 | 31506 | 31506 | 31506 | 30464 | 30454 | 31506 | 31506 | 31506 | 31506 | 31506 | 31506 | 31506 | 31506 | 31506 | 31506 |
| **1981** | 31368 | 31368 | 31368 | 31368 | 30288 | 30929 | 31368 | 31368 | 31368 | 31368 | 31368 | 31368 | 31368 | 31368 | 31368 | 31368 |
| **1982** | 32732 | 32732 | 32732 | 32732 | 31638 | 31994 | 32732 | 32732 | 32732 | 32732 | 32732 | 32732 | 32732 | 32732 | 32732 | 32732 |
| **1983** | 40153 | 40153 | 40153 | 40153 | 39541 | 38643 | 40153 | 40153 | 40153 | 40153 | 40153 | 40153 | 40153 | 40153 | 40153 | 40153 |
| **1984** | 42800 | 42800 | 42800 | 42800 | 41992 | 41438 | 42800 | 42800 | 42800 | 42800 | 42800 | 42800 | 42800 | 42800 | 42800 | 42800 |
| **1985** | 46245 | 46245 | 46245 | 46245 | 45431 | 45332 | 46245 | 46245 | 46245 | 46245 | 46245 | 46245 | 46245 | 46245 | 46245 | 46245 |
| **1986** | 42564 | 42564 | 42564 | 42564 | 41657 | 41762 | 42564 | 42564 | 42564 | 42564 | 42564 | 42564 | 42564 | 42564 | 42564 | 42564 |
| **1987** | 35539 | 35539 | 35539 | 35539 | 34475 | 35150 | 35539 | 35539 | 35539 | 35539 | 35539 | 35539 | 35539 | 35539 | 35539 | 35539 |
| **1988** | 28739 | 28739 | 28739 | 28739 | 28302 | 28638 | 28739 | 28739 | 28739 | 28739 | 28739 | 28739 | 28739 | 28739 | 28739 | 28739 |
| **1989** | 25988 | 25988 | 25988 | 25988 | 25818 | 25317 | 25988 | 25988 | 25988 | 25988 | 25988 | 25988 | 25988 | 25988 | 25988 | 25988 |
| **1990** | 17475 | 17475 | 17475 | 17475 | 17450 | 17218 | 17475 | 17475 | 17475 | 17475 | 17475 | 17475 | 17475 | 17475 | 17475 | 17475 |
| **1991** | 20227 | 20227 | 20227 | 20227 | 20227 | 19354 | 20227 | 20227 | 20227 | 20227 | 20227 | 20227 | 20227 | 20227 | 20227 | 20227 |
| **1992** | 19672 | 19672 | 19672 | 19672 | 19672 | 19338 | 19672 | 19672 | 19672 | 19672 | 19672 | 19672 | 19672 | 19672 | 19672 | 19672 |
| **1993** | 17153 | 17153 | 17153 | 17153 | 17153 | 16990 | 17153 | 17153 | 17153 | 17153 | 17153 | 17153 | 17153 | 17153 | 17153 | 17153 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1994** | 25637 | 25637 | 25637 | 25637 | 25637 | 25471 | 25637 | 25637 | 25637 | 25637 | 25637 | 25637 | 25637 | 25637 | 25637 | 25637 |
| **1995** | 30588 | 30588 | 30588 | 30588 | 30588 | 30437 | 30588 | 30588 | 30588 | 30588 | 30588 | 30588 | 30588 | 30588 | 30588 | 30588 |
| **1996** | 35991 | 35991 | 35991 | 35991 | 35991 | 35713 | 35991 | 35991 | 35991 | 35991 | 35991 | 35991 | 35991 | 35991 | 35991 | 35991 |
| **1997** | 40691 | 40691 | 40691 | 40691 | 40691 | 40459 | 40691 | 40691 | 40691 | 40691 | 40691 | 40691 | 40691 | 40691 | 40691 | 40691 |
| **1998** | 37609 | 37609 | 37609 | 37609 | 37609 | 37262 | 37609 | 37609 | 37609 | 37609 | 37609 | 37609 | 37609 | 37609 | 37609 | 37609 |
| **1999** | 33249 | 33249 | 33249 | 33249 | 33249 | 32875 | 33249 | 33249 | 33249 | 33249 | 33249 | 33249 | 33249 | 33249 | 33249 | 33249 |
| **2000** | 32199 | 32199 | 32199 | 32199 | 32199 | 31767 | 32199 | 32199 | 32199 | 32199 | 32199 | 32199 | 32199 | 32199 | 32199 | 32199 |
| **2001** | 34827 | 34827 | 34827 | 34827 | 34827 | 34204 | 34827 | 34827 | 34827 | 34827 | 34827 | 34827 | 34827 | 34827 | 34827 | 34827 |
| **2002** | 31471 | 31471 | 31471 | 31471 | 31471 | 30926 | 31471 | 31471 | 31471 | 31471 | 31471 | 31471 | 31471 | 31471 | 31471 | 31471 |
| **2003** | 23827 | 23827 | 23827 | 23827 | 23827 | 23021 | 23827 | 23827 | 23827 | 23827 | 23827 | 23827 | 23827 | 23827 | 23827 | 23827 |
| **2004** | 30271 | 30271 | 30271 | 30271 | 30271 | 29330 | 30271 | 30271 | 30271 | 30271 | 30271 | 30271 | 30271 | 30271 | 30271 | 30271 |
| **2005** | 34389 | 34389 | 34389 | 34389 | 34389 | 33294 | 34389 | 34389 | 34389 | 34389 | 34389 | 34389 | 34389 | 34389 | 34389 | 34389 |
| **2006** | 34021 | 34021 | 34021 | 34021 | 34021 | 33634 | 34021 | 34021 | 34021 | 34021 | 34021 | 34021 | 34021 | 34021 | 34021 | 34021 |
| **2007** | 30708 | 30708 | 30708 | 30708 | 30708 | 30675 | 30708 | 30708 | 30708 | 30708 | 30708 | 30708 | 30708 | 30708 | 30708 | 30708 |
| **2008** | 25552 | 25552 | 25552 | 25552 | 25552 | 25519 | 25552 | 25552 | 25552 | 25552 | 25552 | 25552 | 25552 | 25552 | 25552 | 25552 |
| **2009** | 20454 | 20454 | 20454 | 20454 | 20454 | 20421 | 20454 | 20454 | 20454 | 20454 | 20454 | 20454 | 20454 | 20454 | 20454 | 20454 |
| **2010** | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 | 12286 |
| **2011** | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 | 10131 |
| **2012** | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 | 10607 |
| **2013** | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 | 9974 |
| **2014** | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 | 10236 |
| **2015** | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 | 8137 |

*Table 4: Data format for Taiwanese longline dataset.*

| Items | Type | Column | 1979-1994 | 1995-2005 | 2006-2013 | Remarks |
|---|---|---|---|---|---|---|
| call sign | character | 1-5 | YES | YES | YES | See below re first digit |
| operation year | integer | 6-9 | YES | YES | YES | |
| operation month | integer | 10-11 | YES | YES | YES | |
| operation day | integer | 12-13 | YES | YES | YES | |
| operational area | integer | 14-17 | YES | YES | YES | Reference to map |
| no. of hooks between floats | integer | 18-20 | NO | YES | YES | |
| total no. of hooks per set | integer | 21-25 | YES | YES | YES | |
| albacore catch in number | integer | 26-29 | YES | YES | YES | |
| bigeye catch in number | integer | 30-33 | YES | YES | YES | |
| yellowfin catch in number | integer | 34-37 | YES | YES | YES | |
| bluefin catch in number | integer | 38-41 | YES | YES | YES | |
| southern bluefin catch in number | integer | 42-45 | YES | YES | YES | |
| other tuna catch in number | integer | 46-49 | YES | YES | YES | |
| swordfish catch in number | integer | 50-53 | YES | YES | YES | |
| white marlin catch in number | integer | 54-57 | YES | YES | YES | |
| blue marlin catch in number | integer | 58-61 | YES | YES | YES | |
| black marlin catch in number | integer | 62-65 | YES | YES | YES | |
| other billfish catch in number | integer | 66-69 | YES | YES | YES | |
| skipjack catch in number | integer | 70-73 | YES | YES | YES | |
| shark catch in number | integer | 74-77 | YES | YES | YES | |
| other species catch in number | integer | 78-81 | YES | YES | YES | |
| albacore catch in weight | integer | 82-86 | YES | YES | YES | |
| bigeye catch in weight | integer | 87-91 | YES | YES | YES | |
| yellowfin catch in weight | integer | 92-96 | YES | YES | YES | |
| bluefin catch in weight | integer | 97-101 | YES | YES | YES | |
| southern bluefin catch in wt | integer | 102-106 | YES | YES | YES | |
| other tuna catch in wt | integer | 107-111 | YES | YES | YES | |
| swordfish catch in wt | integer | 112-116 | YES | YES | YES | |
| white marlin catch in wt | integer | 117-121 | YES | YES | YES | |
| blue marlin catch in wt | integer | 122-126 | YES | YES | YES | |
| black marlin catch in wt | integer | 127-131 | YES | YES | YES | |
| other billfish catch in wt | integer | 132-136 | YES | YES | YES | |
| skipjack catch in number | integer | 137-141 | YES | YES | YES | |
| shark catch in number | integer | 142-146 | YES | YES | YES | |
| other spp catch in number | integer | 147-151 | YES | YES | YES | |
| SST | Integer | 152-153 | YES | YES | YES | |
| bait type: pacific saury | integer | 154 | YES | YES | YES | |
| bait type: mackerel | integer | 155 | YES | YES | YES | |
| bait type: squid | integer | 156 | YES | YES | YES | |
| bait type: milkfish | integer | 157 | YES | YES | YES | |
| bait type: others | integer | 158 | YES | YES | YES | |
| Depth of hooks (m) | Integer | 159-161 | NO | YES | YES | |
| set type (type of target) | character | 162-163 | NO | NO | YES | 1.BET, 2. ALB, 3.both |
| Remark | integer | 164-165 | NO | NO | YES | See below |
| operation latitude code | character | 166-166 | NO | YES | YES | N: 4, S: 3 |
| operation latitude | Integer | 167-168 | NO | YES | YES | |
| operation longitude code | Character | 169-169 | NO | YES | YES | E: 1, W: 2 |
| operation longitude | Integer | 170-172 | NO | YES | YES | |
| departure date from port | Integer | 176-183 | YES | YES | YES | |
| starting date to operation | Integer | 185-192 | NO | YES | YES | |
| stop date to operation | Integer | 194-201 | NO | YES | YES | |
| arrival date at port | Integer | 203-210 | YES | YES | YES | |

*Table 5: Tonnage as indicated by first digit of TW callsign.*

| First digit | Tonnage |
|---|---|
| 1 | >= 5 and < 10 tonnes |
| 2 | >= 10 and < 20 tonnes |
| 3 | >= 20 and < 50 tonnes |
| 4 | >= 50 and < 100 tonnes |
| 5 | >= 100 and < 200 tonnes |
| 6 | >= 200 and < 500 tonnes |
| 7 | >= 500 and < 1,000 tonnes |
| 8 | >= 1,000 tonnes |

*Table 6: Codes in the Remarks field of the TW dataset, indicating outliers.*

| Dates | Code | Outliers |
|---|---|---|
| 2007-2011 | G1 | extremely high BET catch |
| | G4 | extremely high ALB |
| | G6 | extremely high YFT catch |
| | G8 | extremely high SWO; |
| | SF | for a given year and vessel, record only single species catch for 3 successive months |
| 2012-2015 | G1 | extremely high ALB catch |
| | G2 | extremely high BET |
| | G3 | extremely high YFT catch |
| | G7 | extremely high SWO |
| | GH | abnormal total no. of hooks per set |
| | GL | more than one anomaly |
| | SF | for a given year and vessel, only record single species catch for 3 successive months |

**2007-2011:**
1.G1:extremely high BET catch ( > 5 tons per set or outliers in the distribution of bet catch number per set) ; G4: extremely high ALB;
G6: extremely high YFT catch; G8: extremely high SWO;
SF: for a given year and a given vessel, record only single species catch for three successive months.

**2012-2015:**
G1: extremely high ALB catch (Based on definition of IOTC BET regions, for a given year and a given region, average catch numbers per set for a given vessel. Then use the IQR Rule*. Remark all sets by the vessel which reported the outlier for the given year and region);
G2: extremely high BET;
G3: extremely high YFT catch; G7: extremely high SWO;
GH: abnormal total no. of hooks per set;
GL: if there are more than one anomaly.
SF: for a given year and a given vessel, only record single species catch for three successive months.

Criteria for outliers
( > 5 tons per set or outliers in the distribution of bet catch number per set)

*IQR Rule for Outliers
1. Arrange average catch numbers per set for all vessels in order.
2. Calculate first quartile (Q1), third quartile (Q3) and the interquartile range (IQR=Q3-Q1).
3. Compute Q1-1.5 x IQR and Compute Q3+1.5 x IQR. Anything outside this range is an outlier.

*Table 7a: Taiwanese data sample sizes by variable.*

| Year | No. of ops | Cruise start date | Cruise end date | Op start date | Op end date |
|---|---|---|---|---|---|
| 1979 | 16,056 | 15,996 | 16,056 | 0 | 0 |
| 1980 | 21,021 | 20,682 | 21,021 | 0 | 0 |
| 1981 | 16,969 | 16,835 | 16,969 | 0 | 0 |
| 1982 | 23,110 | 23,110 | 23,110 | 0 | 0 |
| 1983 | 22,048 | 22,048 | 22,048 | 0 | 0 |
| 1984 | 17,551 | 17,551 | 17,551 | 0 | 0 |
| 1985 | 13,531 | 13,531 | 13,531 | 0 | 0 |
| 1986 | 13,257 | 13,257 | 13,257 | 0 | 0 |
| 1987 | 14,431 | 14,431 | 14,431 | 0 | 0 |
| 1988 | 12,497 | 12,497 | 12,497 | 0 | 0 |
| 1989 | 9,045 | 9,045 | 9,045 | 0 | 0 |
| 1990 | 7,181 | 7,181 | 7,181 | 0 | 0 |
| 1991 | 5,738 | 5,738 | 5,738 | 0 | 0 |
| 1992 | 3,499 | 3,499 | 3,499 | 0 | 0 |
| 1993 | 17,869 | 17,869 | 17,869 | 0 | 0 |
| 1994 | 20,315 | 7,726 | 7,726 | 1,359 | 2,021 |
| 1995 | 19,341 | 19,341 | 19,196 | 19,077 | 19,341 |
| 1996 | 24,492 | 24,402 | 24,492 | 24,492 | 24,492 |
| 1997 | 25,503 | 23,137 | 25,503 | 25,503 | 25,503 |
| 1998 | 24,041 | 23,653 | 24,041 | 24,041 | 24,041 |
| 1999 | 29,608 | 29,037 | 29,608 | 29,563 | 29,608 |
| 2000 | 31,664 | 30,489 | 31,569 | 31,593 | 31,569 |
| 2001 | 40,636 | 39,073 | 40,486 | 40,486 | 40,486 |
| 2002 | 42,017 | 41,522 | 42,017 | 42,017 | 42,017 |
| 2003 | 69,329 | 68,205 | 65,718 | 69,329 | 69,329 |
| 2004 | 80,508 | 77,186 | 76,430 | 80,508 | 80,508 |
| 2005 | 72,204 | 68,983 | 63,761 | 72,204 | 72,204 |
| 2006 | 51,798 | 47,281 | 47,784 | 51,798 | 51,798 |
| 2007 | 44,016 | 36,749 | 37,705 | 44,016 | 44,016 |
| 2008 | 31,809 | 24,716 | 25,335 | 31,809 | 31,809 |
| 2009 | 40,097 | 31,527 | 31,265 | 40,097 | 40,097 |
| 2010 | 29,856 | 26,057 | 23,609 | 29,801 | 29,801 |
| 2011 | 22,544 | 19,182 | 17,000 | 22,544 | 22,544 |
| 2012 | 25,284 | 19,717 | 20,000 | 25,284 | 25,284 |
| 2013 | 23,723 | 20,901 | 20.551 | 23,723 | 23,723 |
| 2014 | 16,742 | 15,954 | 13,632 | 16,742 | 16,742 |
| 2015 | 20,442 | 0 | 0 | 0 | 0 |

*Table 7b: Taiwanese data sample sizes by variable.*

| Year | No. of ops | Set type | Lat & long in 1° | NHBF | No of ops after cleaning |
|---|---|---|---|---|---|
| 1979 | 16,056 | 0 | 0 | 0 | 13,461 |
| 1980 | 21,021 | 0 | 0 | 0 | 18,017 |
| 1981 | 16,969 | 0 | 0 | 0 | 14,822 |
| 1982 | 23,110 | 0 | 0 | 0 | 18,678 |
| 1983 | 22,048 | 0 | 0 | 0 | 18,253 |
| 1984 | 17,551 | 0 | 0 | 0 | 15,053 |
| 1985 | 13,531 | 0 | 0 | 0 | 12,280 |
| 1986 | 13,257 | 0 | 0 | 0 | 11,342 |
| 1987 | 14,431 | 0 | 0 | 0 | 12,110 |
| 1988 | 12,497 | 0 | 0 | 0 | 10,875 |
| 1989 | 9,045 | 0 | 0 | 0 | 7,725 |
| 1990 | 7,181 | 0 | 0 | 0 | 6,270 |
| 1991 | 5,738 | 0 | 0 | 0 | 5,273 |
| 1992 | 3,499 | 0 | 0 | 0 | 3,113 |
| 1993 | 17,869 | 0 | 0 | 0 | 13,806 |
| 1994 | 20,315 | 0 | 20,315 | 0 | 17,594 |
| 1995 | 19,341 | 0 | 12,051 | 7,116 | 10,904 |
| 1996 | 24,492 | 0 | 18,408 | 10,884 | 15,886 |
| 1997 | 25,503 | 0 | 20,565 | 9,495 | 20,537 |
| 1998 | 24,041 | 0 | 19,785 | 10,022 | 17,663 |
| 1999 | 29,608 | 0 | 24,603 | 14,198 | 23,193 |
| 2000 | 31,664 | 0 | 26,723 | 16,022 | 24,734 |
| 2001 | 40,636 | 0 | 37,853 | 32,575 | 35,534 |
| 2002 | 42,017 | 0 | 38,204 | 40,768 | 39,782 |
| 2003 | 69,329 | 0 | 53,455 | 69,183 | 40,296 |
| 2004 | 80,508 | 0 | 76,388 | 80,402 | 75,399 |
| 2005 | 72,204 | 0 | 70,135 | 72,204 | 65,655 |
| 2006 | 51,798 | 51,798 | 50,987 | 51,798 | 46,073 |
| 2007 | 44,016 | 44,016 | 43,506 | 44,016 | 39,005 |
| 2008 | 31,809 | 31,809 | 31,176 | 31,809 | 28,632 |
| 2009 | 40,097 | 40,097 | 39,355 | 40,097 | 36,732 |
| 2010 | 29,856 | 29,856 | 29,756 | 29,856 | 28,244 |
| 2011 | 22,544 | 22,544 | 22,544 | 22,544 | 21,161 |
| 2012 | 25,284 | 25,284 | 25,284 | 25,284 | 24,350 |
| 2013 | 23,723 | 23,723 | 23,723 | 23,723 | 21,764 |
| 2014 | 16,742 | 16,742 | 16,742 | 16,742 | 15,768 |
| 2015 | 20,442 | 20,442 | 20,442 | 20,442 | 19,212 |

*Table 8: Korean data description.*

| Year | No. of ops | VESSEL NAME_rev | Vessel id coverage (%) | Hooks | Floats | Op date |
|---|---|---|---|---|---|---|
| 1971 | 34 | 34 | 100.0 | 34 | 34 | 34 |
| 1972 | 3265 | 53 | 1.6 | 3265 | 3265 | 3265 |
| 1973 | 508 | 508 | 100.0 | 508 | 241 | 508 |
| 1974 | 1255 | 1255 | 100.0 | 1255 | 93 | 1255 |
| 1975 | 5313 | 5051 | 95.1 | 5021 | 334 | 5313 |
| 1976 | 119 | 119 | 100.0 | 119 | 119 | 119 |
| 1977 | 3714 | 3714 | 100.0 | 3714 | 3714 | 3736 |
| 1978 | 23191 | 22882 | 98.7 | 23191 | 23191 | 23191 |
| 1979 | 10509 | 10433 | 99.3 | 10509 | 10509 | 10651 |
| 1980 | 20446 | 19874 | 97.2 | 20446 | 20446 | 20408 |
| 1981 | 15566 | 15527 | 99.7 | 15566 | 15566 | 15585 |
| 1982 | 17119 | 16593 | 96.9 | 17119 | 17119 | 17176 |
| 1983 | 19255 | 18216 | 94.6 | 19255 | 19255 | 19255 |
| 1984 | 7912 | 7684 | 97.1 | 7912 | 7912 | 8080 |
| 1985 | 11386 | 10887 | 95.6 | 11386 | 11386 | 11530 |
| 1986 | 14374 | 14157 | 98.5 | 14374 | 14374 | 14462 |
| 1987 | 14810 | 14660 | 99.0 | 14810 | 14810 | 14810 |
| 1988 | 17568 | 17409 | 99.1 | 17568 | 17568 | 17568 |
| 1989 | 18771 | 18127 | 96.6 | 18771 | 18771 | 18771 |
| 1990 | 14162 | 14073 | 99.4 | 14162 | 14162 | 14162 |
| 1991 | 4533 | 4533 | 100.0 | 4533 | 4533 | 4533 |
| 1992 | 7005 | 7005 | 100.0 | 7005 | 7005 | 7005 |
| 1993 | 9569 | 9569 | 100.0 | 9569 | 9569 | 9569 |
| 1994 | 10141 | 9065 | 89.4 | 10141 | 10141 | 10141 |
| 1995 | 7577 | 5332 | 70.4 | 7577 | 7577 | 7577 |
| 1996 | 12218 | 7501 | 61.4 | 12218 | 12218 | 12218 |
| 1997 | 13740 | 8031 | 58.4 | 13740 | 13740 | 13740 |
| 1998 | 5165 | 2239 | 43.3 | 5165 | 5165 | 5165 |
| 1999 | 2833 | 1783 | 62.9 | 2833 | 2833 | 2833 |
| 2000 | 4236 | 2394 | 56.5 | 4236 | 4236 | 4236 |
| 2001 | 3162 | 1929 | 61.0 | 3162 | 3162 | 3162 |
| 2002 | 1479 | 1341 | 90.7 | 1479 | 1479 | 1638 |
| 2003 | 2627 | 1474 | 56.1 | 2627 | 2627 | 2627 |
| 2004 | 4345 | 3004 | 69.1 | 4345 | 4345 | 4345 |
| 2005 | 2443 | 2443 | 100.0 | 2443 | 2443 | 2444 |
| 2006 | 3597 | 3508 | 97.5 | 3597 | 3597 | 3597 |
| 2007 | 3371 | 3197 | 94.8 | 3371 | 3371 | 3371 |
| 2008 | 2330 | 2330 | 100.0 | 2330 | 2330 | 2330 |
| 2009 | 3273 | 3273 | 100.0 | 3273 | 3273 | 3273 |
| 2010 | 1851 | 1851 | 100.0 | 1851 | 1851 | 1851 |
| 2011 | 1658 | 1658 | 100.0 | 1658 | 1658 | 1658 |
| 2012 | 1295 | 1295 | 100.0 | 1295 | 1295 | 1295 |
| 2013 | 1659 | 1659 | 100.0 | 1659 | 1659 | 1659 |
| 2014 | 1802 | 1802 | 100.0 | 1802 | 1802 | 1802 |
| 2015 | 2323 | 2323 | 100.0 | 2323 | 2323 | 2323 |

*Table 9: Comparison of field availability among the three fleets.*

| Items | JP | TW | KR |
|---|---|---|---|
| call sign | 1979- | Y | Y |
| operation date | Y | Y | Y |
| Location – 5x5 | Y | Y | Y |
| Location – 1x1 | Y | 1994- | Y |
| no. of hooks between float | * | # | & |
| total no. of hooks per set | Y | Y | Y |
| albacore catch in number | Y | Y | Y |
| bigeye catch in number | Y | Y | Y |
| yellowfin catch in number | Y | Y | Y |
| southern bluefin catch in number | Y | 1994- | Y |
| other tuna catch in number | N | Y | N |
| swordfish catch in number | Y | Y | Y |
| striped marlin catch in number | Y | Y | Y |
| blue marlin catch in number | Y | Y | Y |
| black marlin catch in number | Y | Y | Y |
| sailfish catch in numbers | N | ^ | Y |
| skipjack catch in number | N | Y | Y |
| shark catch in number | N | Y | Y |
| other species catch in number | N | Y[1] | Y[1] |
| Bait type: Pacific saury | Y | N | N |
| Bait type: mackerel | Y | N | N |
| Bait type: squid | Y | N | N |
| Bait type: milkfish | Y | N | N |
| Bait type: others | Y | N | N |

* High coverage since 1971, variable earlier

# Coverage increasing from 1994 to reach 100% by 2003

& number of floats reported for full dataset, and HBF estimated as HBF= hooks/floats

$ No field for SBT before 1994, only reported when skipper changed the field code

^ Reported in ‘other billfish catch’

[1] Different species mix between TW and KR.

*Table 10: Numbers of clusters included in clustering of each region and fishing fleet.*

| Species/ design | Region | JP | TW | KR |
|---|---|---|---|---|
| B2 | 1 | 5 | 5 | 4 |
| | 2 | 5 | 5 | 4 |
| | 3 | 4 | 4 | 4 |
| | 4 | 4 | 4 | 4 |
| Y | 2 | 4 | 4 | 4 |
| | 3 | 4 | 4 | 4 |
| | 4 | 5 | 5 | 5 |
| | 5 | 4 | 4 | 4 |
| A2 | 1 | 4 | 4 | 4 |
| | 2 | 4 | 4 | 4 |
| | 3 | 4 | 4 | 4 |
| | 4 | 4 | 4 | 4 |
| A3 | 1 | 4 | 4 | 4 |
| | 2 | 4 | 3 | 4 |
| | 3 | 4 | 3 | 4 |
| | 4 | 4 | 3 | 4 |
| A5 | 1 | 5 | 5 | 5 |

*Table 11: Clusters included in indices for each fleet and region.*

| Species/design | Region | JP | KR | TW |
|---|---|---|---|---|
| B2 | 1 | 2,3,4 | 1,2,3,4 | 2,3,4,5 |
| | 2 | 1,2,3,4 | 1,2,3,4 | 1,2,3,5 |
| | 3 | 2,4 | 2 | 2,3 |
| | 4 | 1,2 | 1,2,3,4 | 1,2,3,4 |
| Y0 | 2 | 1,2,4 | 1,2,3,4 | 1,2,3,4 |
| | 3 | 1,2 | 1,2,3 | 1,2,3 |
| | 4 | 2 | 2,3 | 3 |
| | 5 | 1,2,3,4 | 2 | 1,2,3,4 |
| A2 | 1 | 2,4 | 3,4 | 1 |
| | 2 | 3 | 3 | 1 |
| | 3 | 3,4 | 3,4 | 1,2 |
| | 4 | 1,3 | 4 | 1,4 |
| A3 | 1 | 2,3 | 4 | 1 |
| | 2 | 3 | 3 | 1 |
| | 3 | 3 | 4 | 1,2 |
| | 4 | 2 | 2,4 | 1,2 |
| A5 | 1 | 2,4 | 5 | 1,2,4 |

# Figures

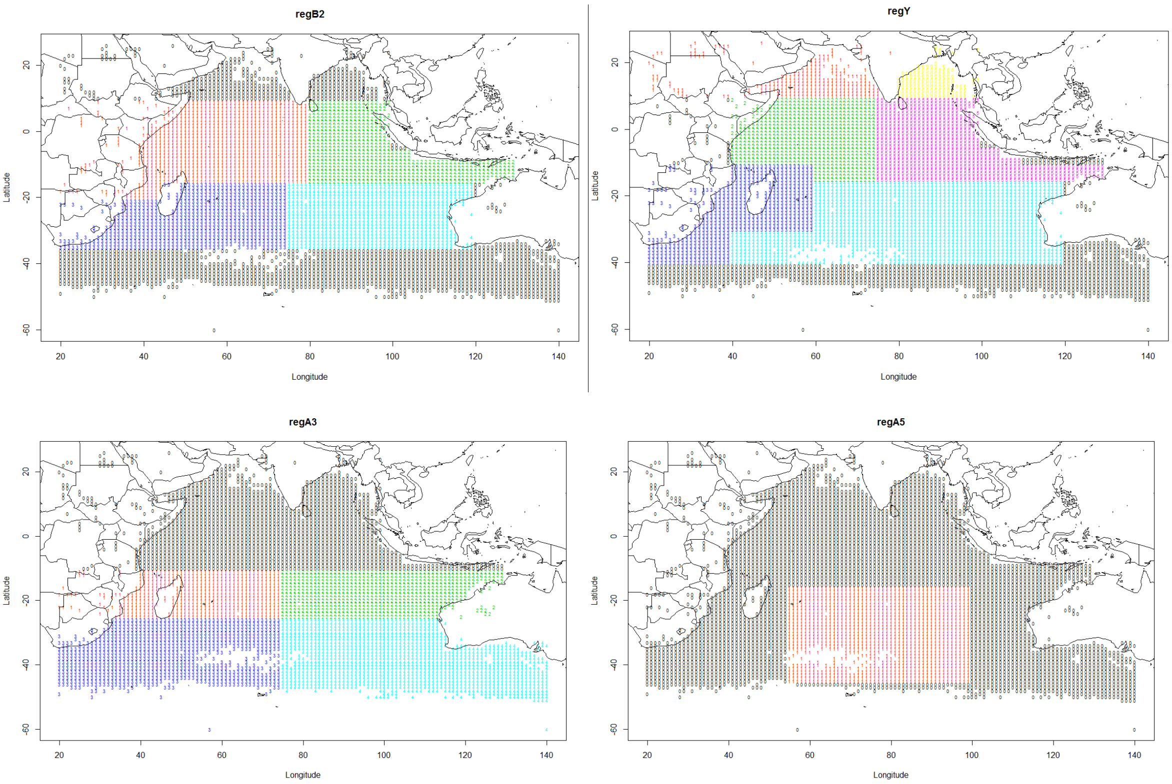


*Figure 1: Maps of the regional structures used to estimate bigeye (top left), yellowfin (top right) and albacore (lower) CPUE indices.*

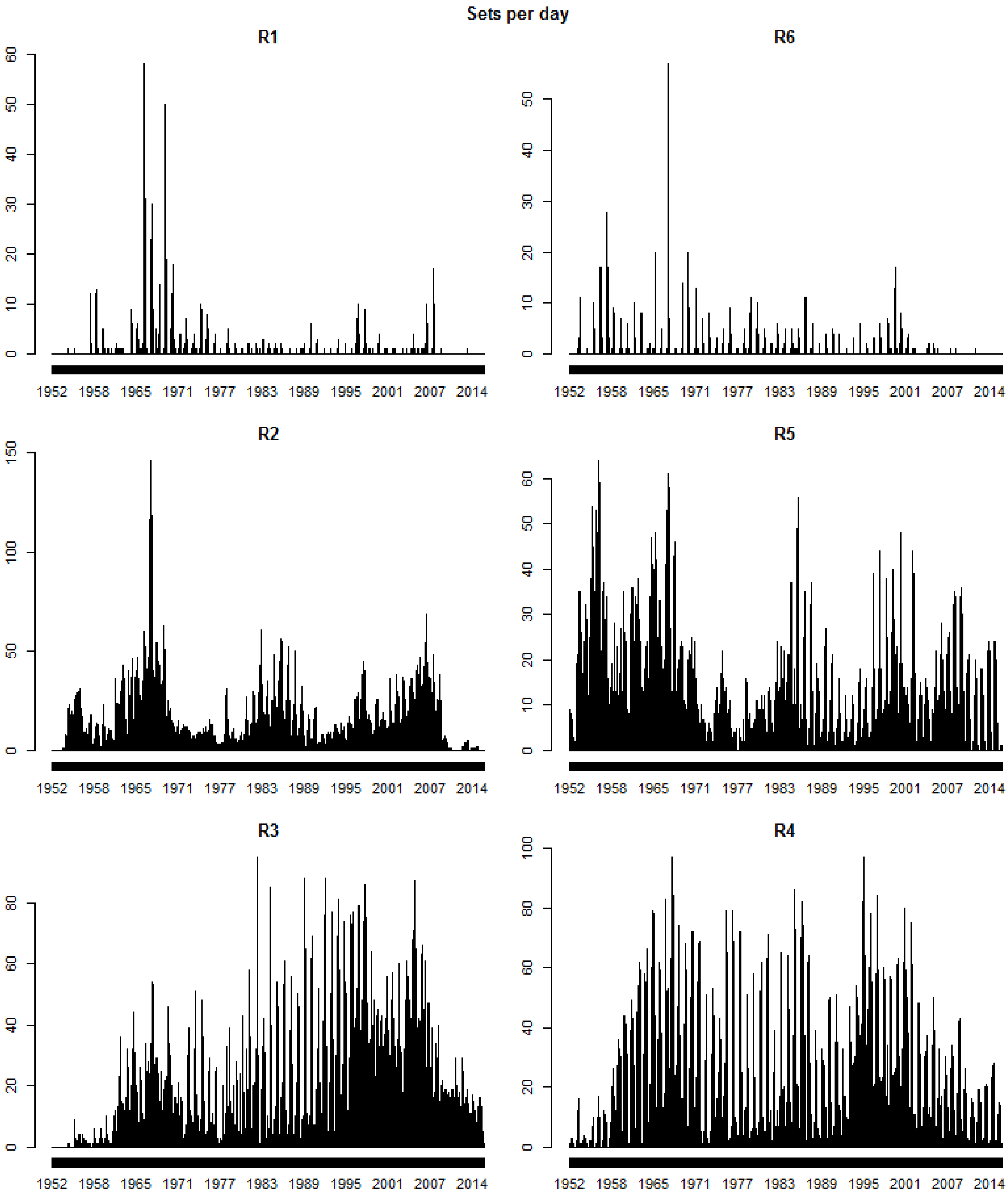


*Figure 2: Sets per day by region for the Japanese fleet in yellowfin regional structure.*

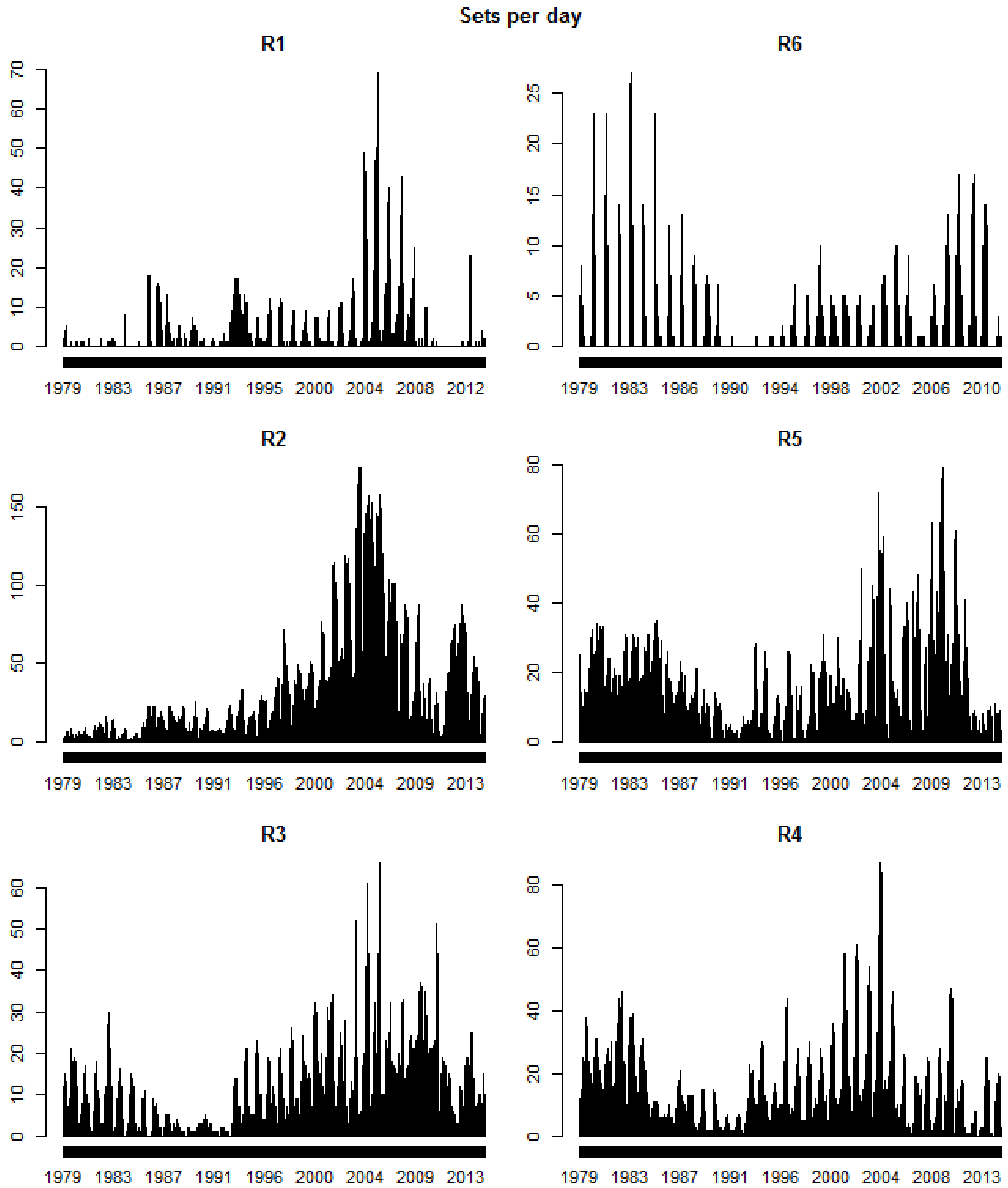


*Figure 3: Sets per day by region for the Taiwanese fleet in the yellowfin regional structure.*

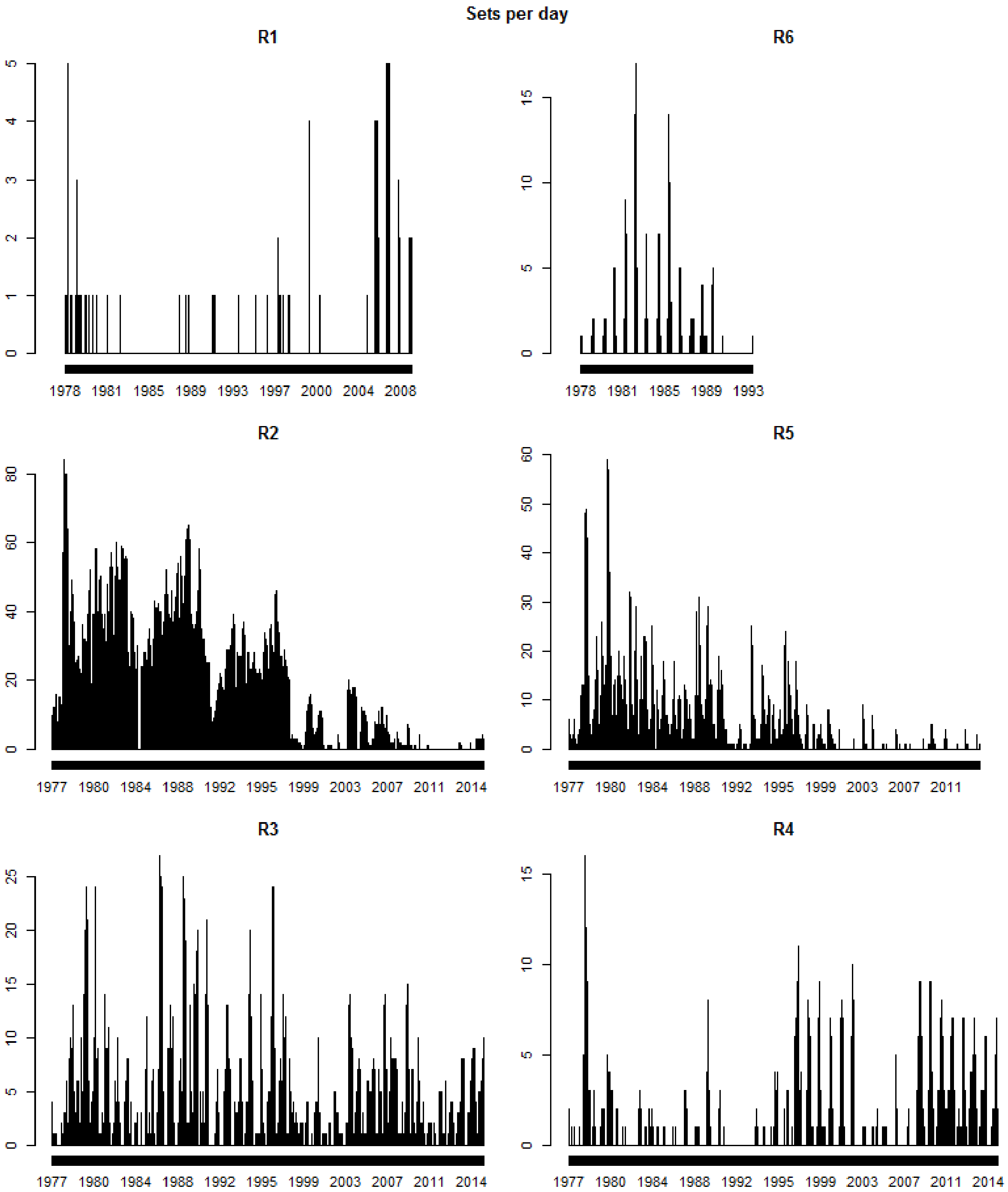


*Figure 4: Sets per day by region for the Korean fleet in the yellowfin regional structure.*

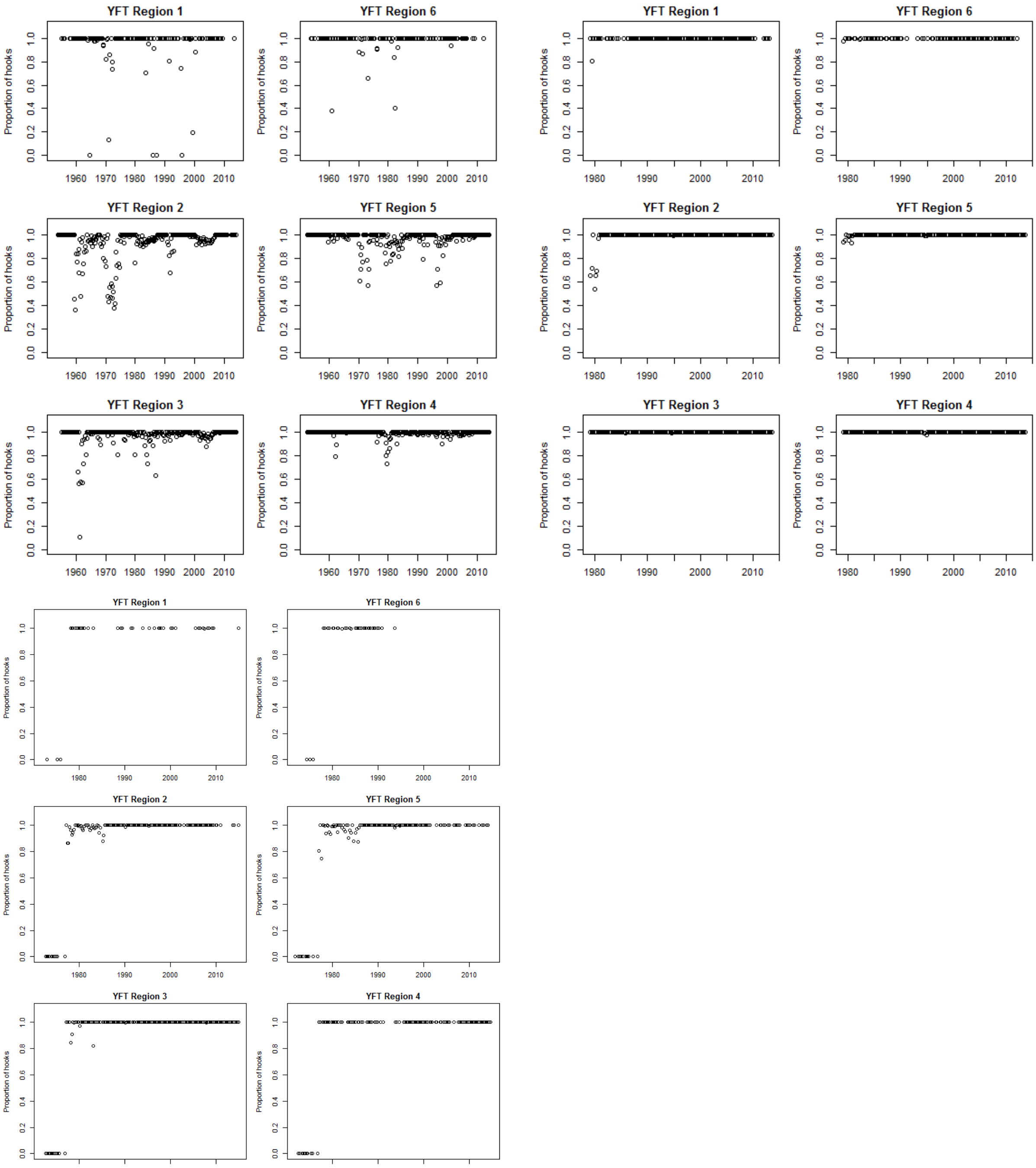


*Figure 5: Proportions of sets retained after data cleaning for analyses in this paper, by region and year-quarter, for Japanese (top left), Taiwanese (top right), and Korean (bottom left) data.*

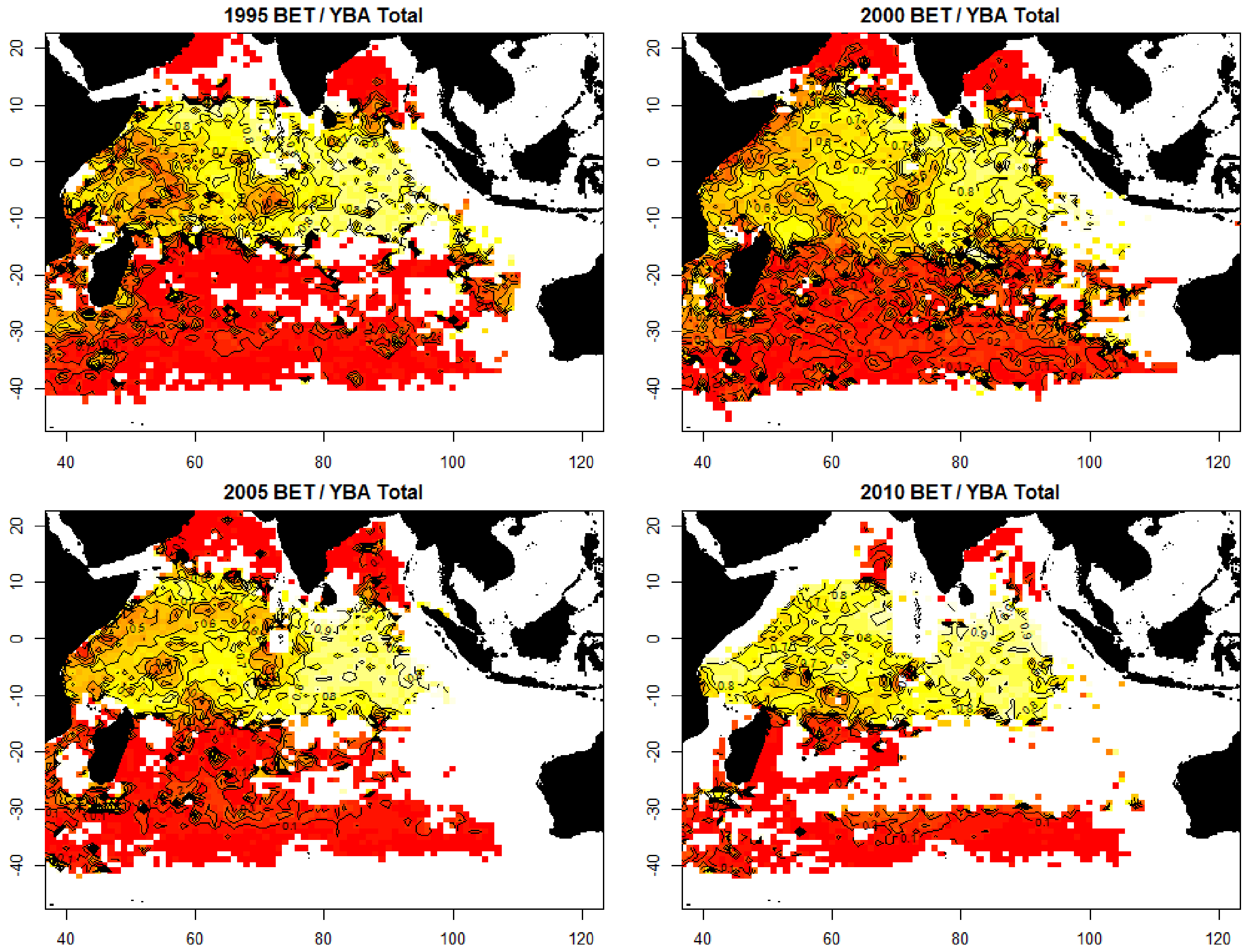
1995 BET / YBA Total
2000 BET / YBA Total
2005 BET / YBA Total
2010 BET / YBA Total
20
10
0
-10
-20
-30
-40
40
60
80
100
120

*Figure 6: Proportions of Taiwanese catch in number reported as bigeye, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of bigeye. Contour lines occur at 5% intervals.*

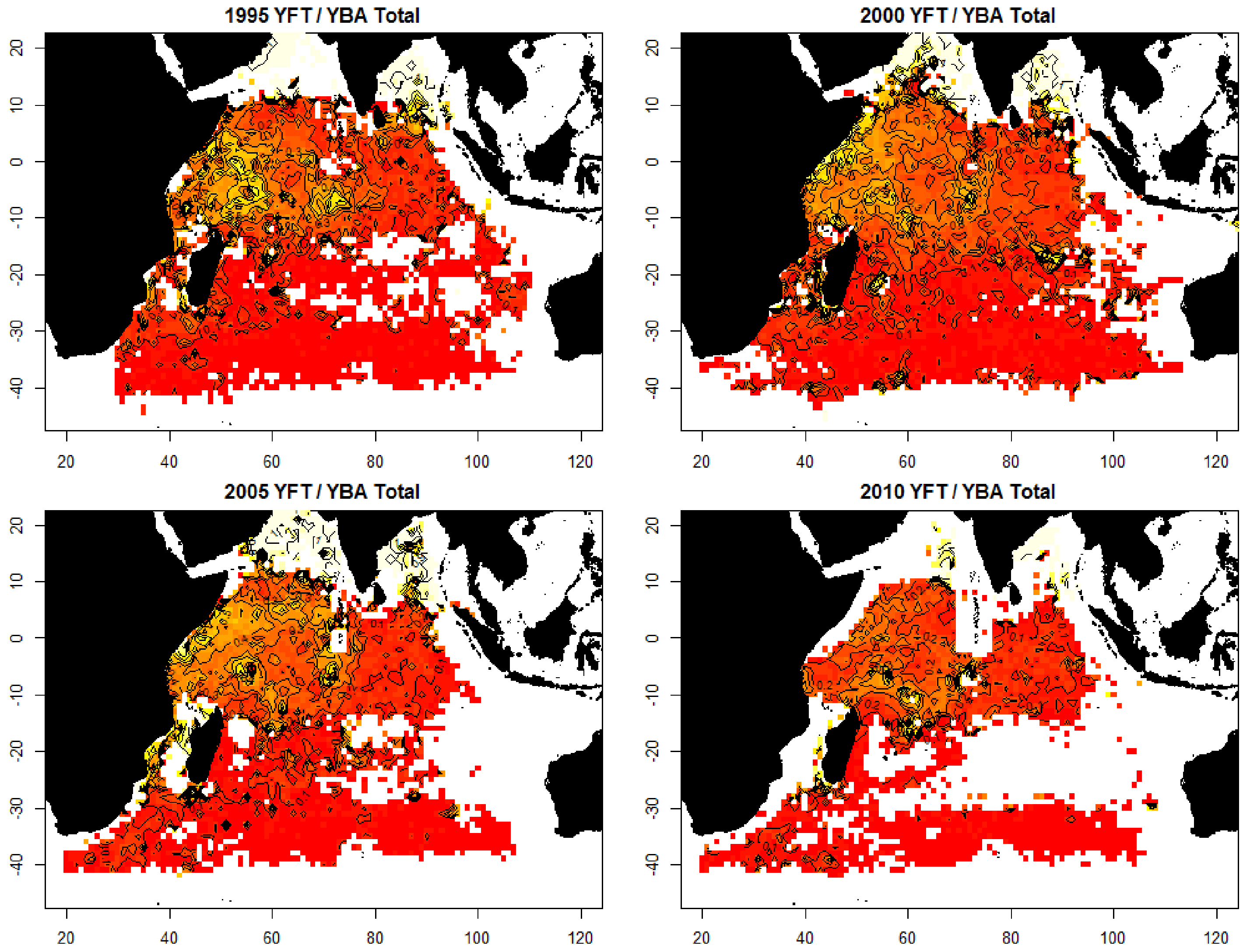
1995 YFT / YBA Total
2000 YFT / YBA Total
2005 YFT / YBA Total
2010 YFT / YBA Total
20
10
0
-10
-20
-30
-40
20
40
60
80
100
120

*Figure 7: Proportions of Taiwanese catch in number reported as yellowfin tuna, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of yellowfin. Contour lines occur at 5% intervals.*

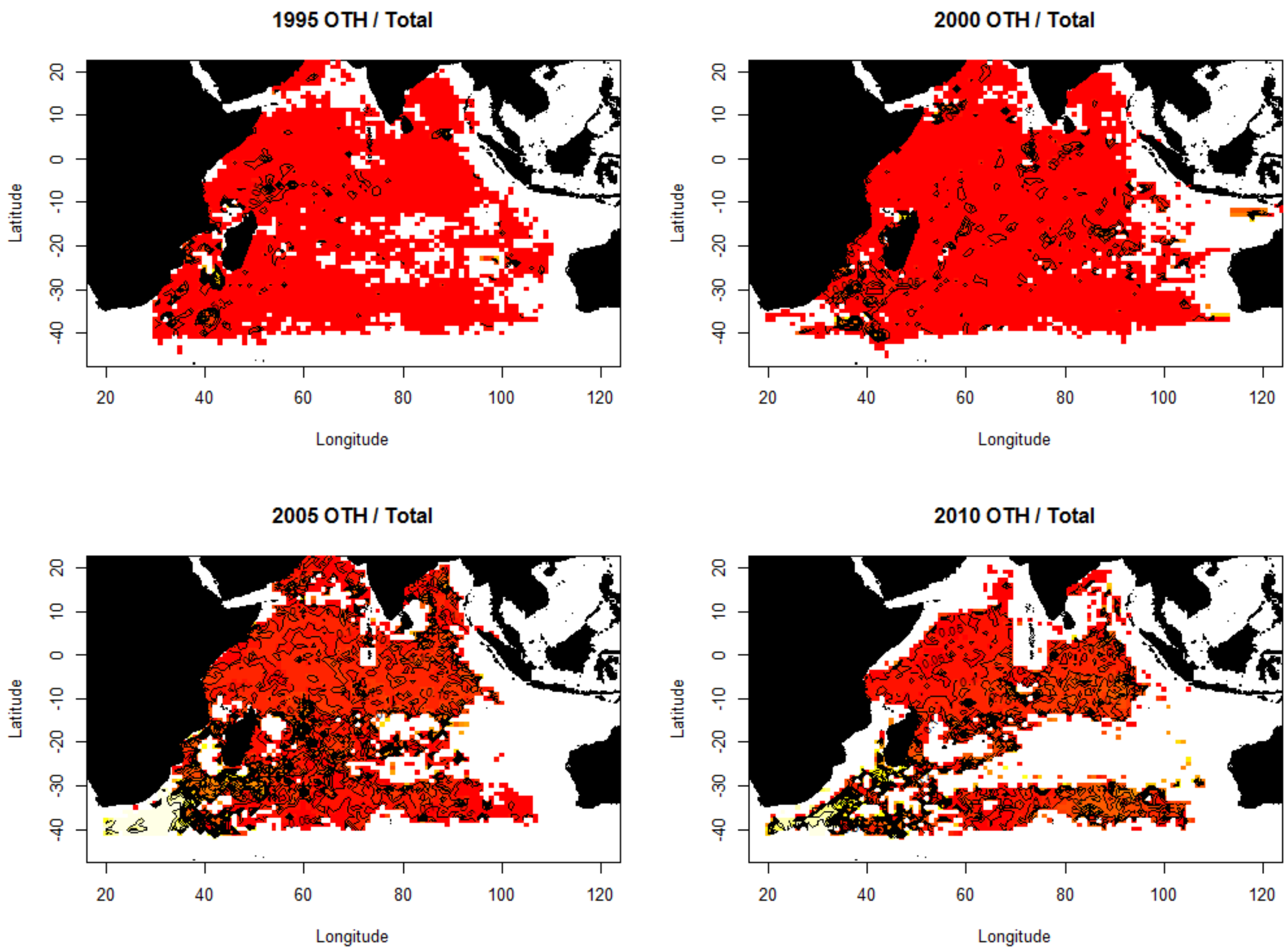


*Figure 8: Proportions of Taiwanese catch in number reported as 'other' species, by 5 year period, mapped by 1° square. More yellow indicates a higher percentage of 'other' species. Contour lines occur at 5% intervals.*

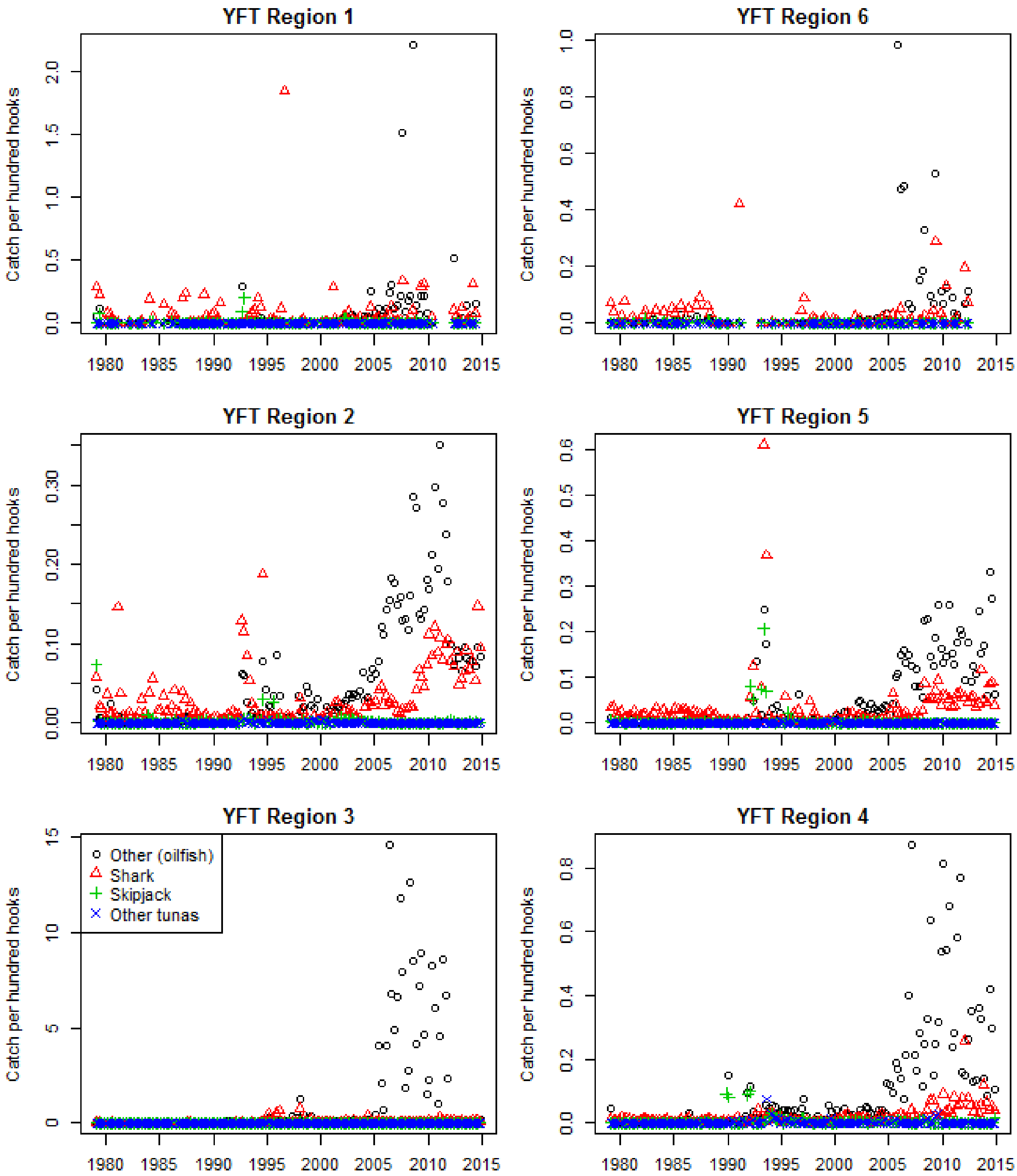


*Figure 9: Taiwanese catch rates per hundred hooks of oilfish, sharks, skipjack, and other tunas, by region and year-quarter.*

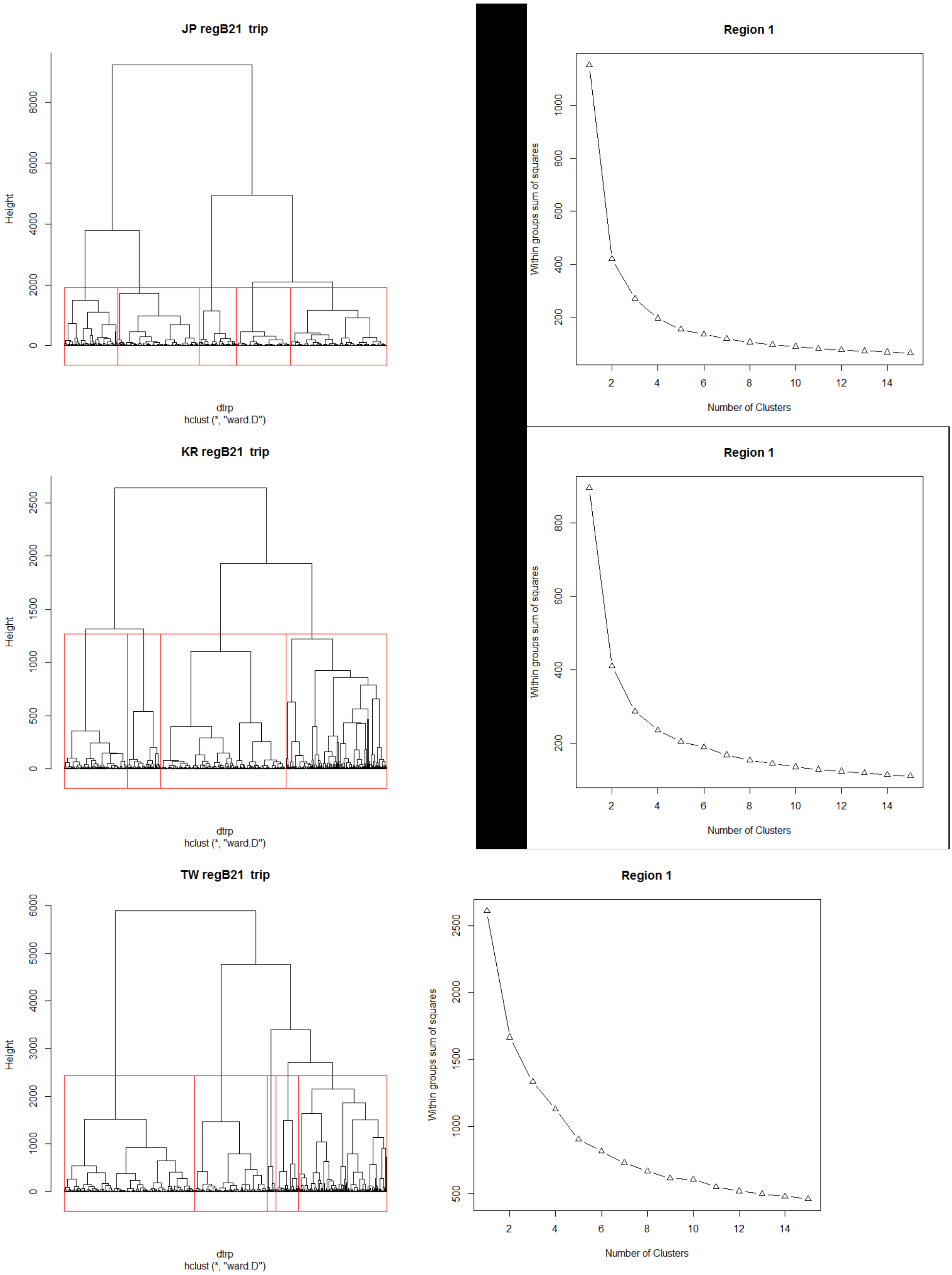


*Figure 10: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 1 (western tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

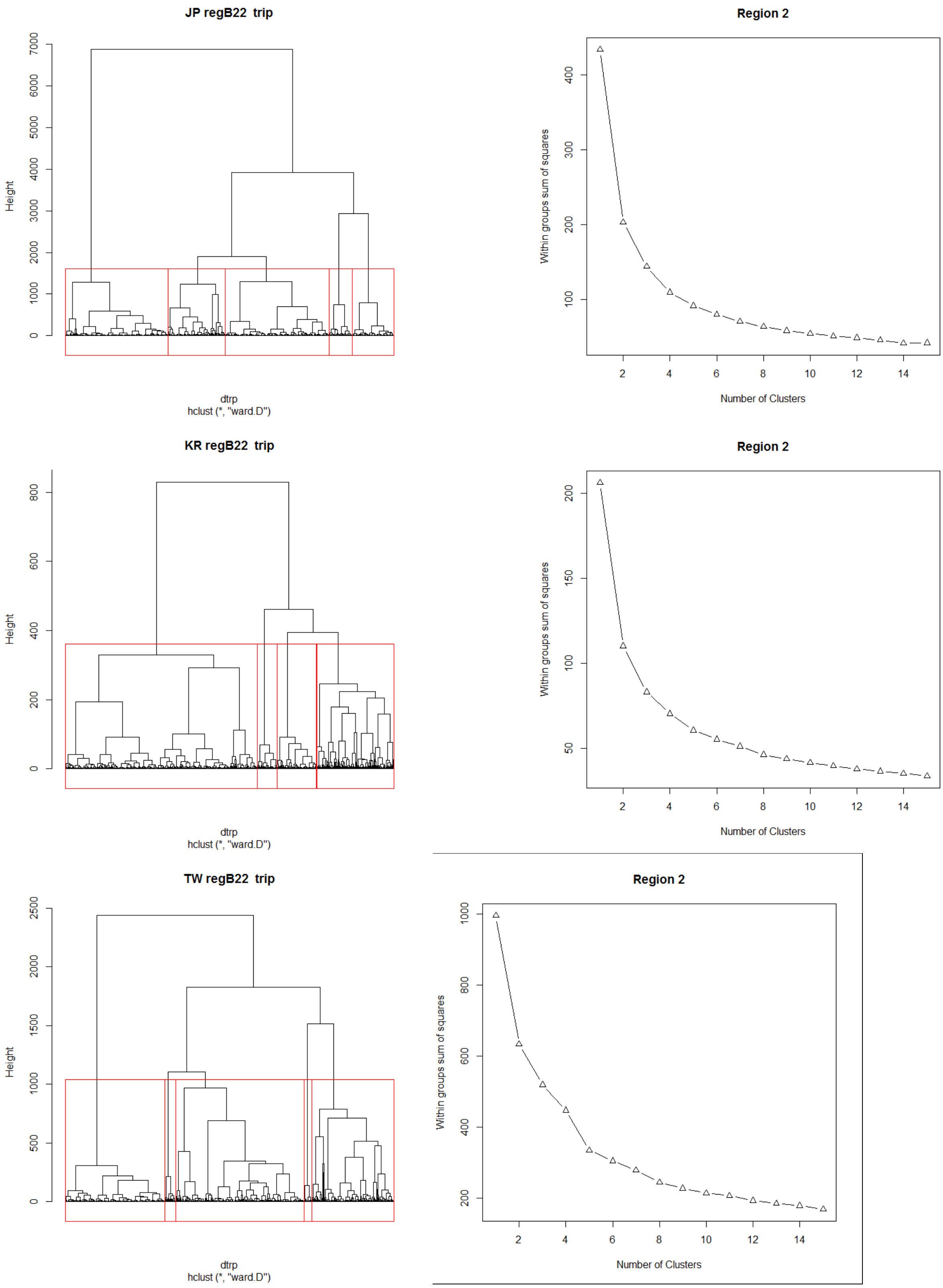


*Figure 11: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 2 (eastern tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

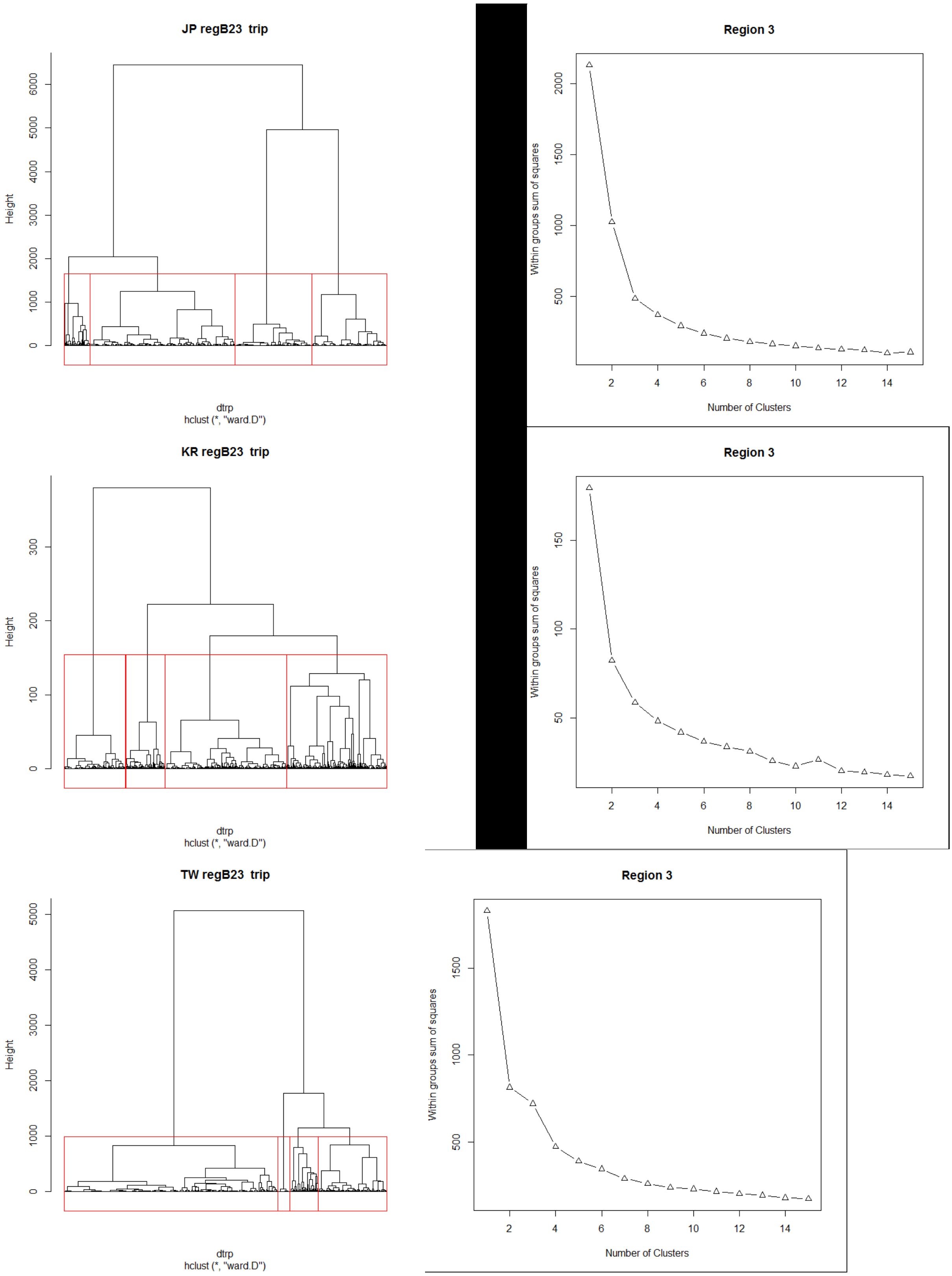


*Figure 12: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 3 (western temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

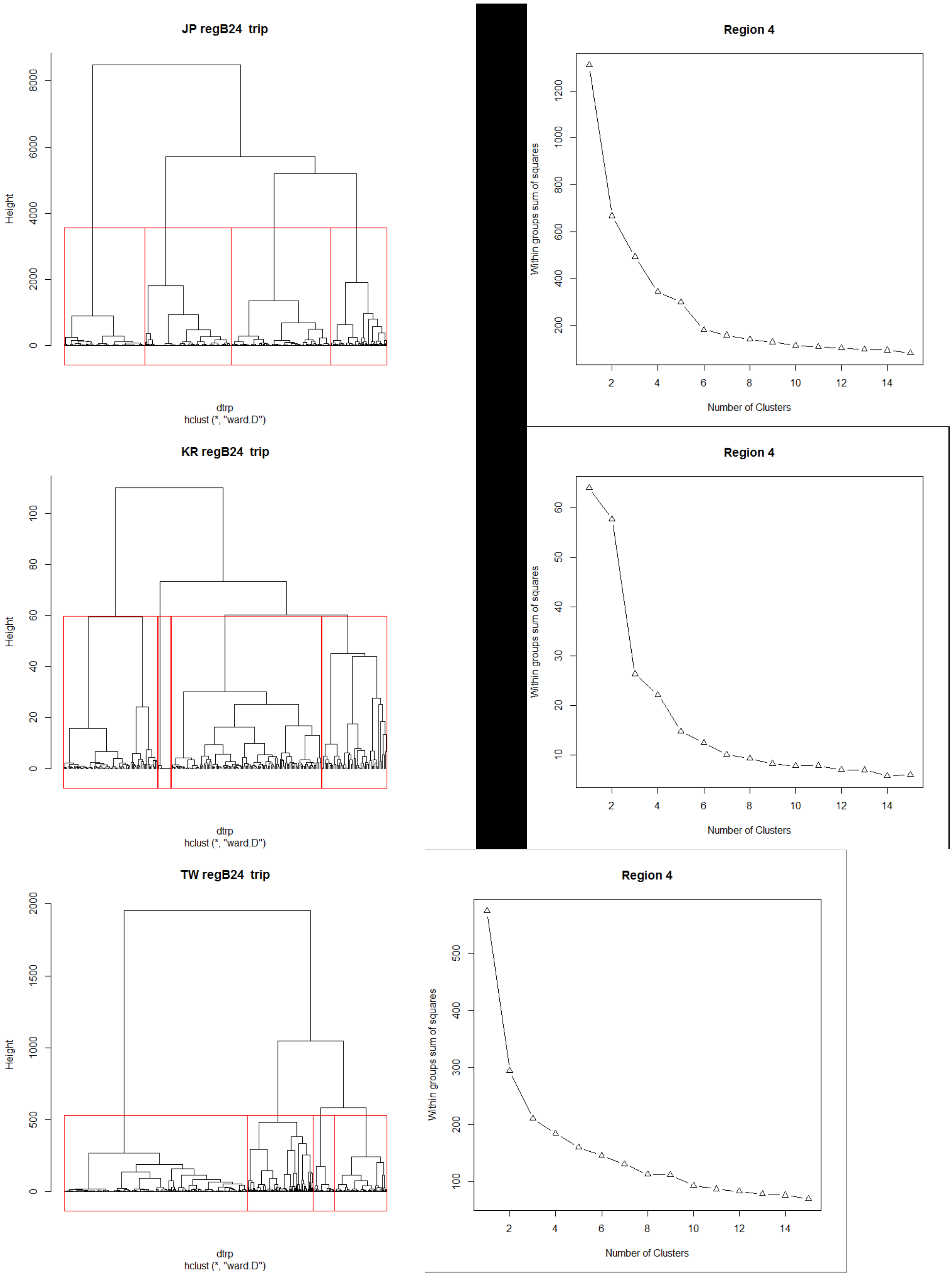


*Figure 13: Plots showing analyses to estimate the number of distinct classes of species composition in bigeye region 4 (eastern temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

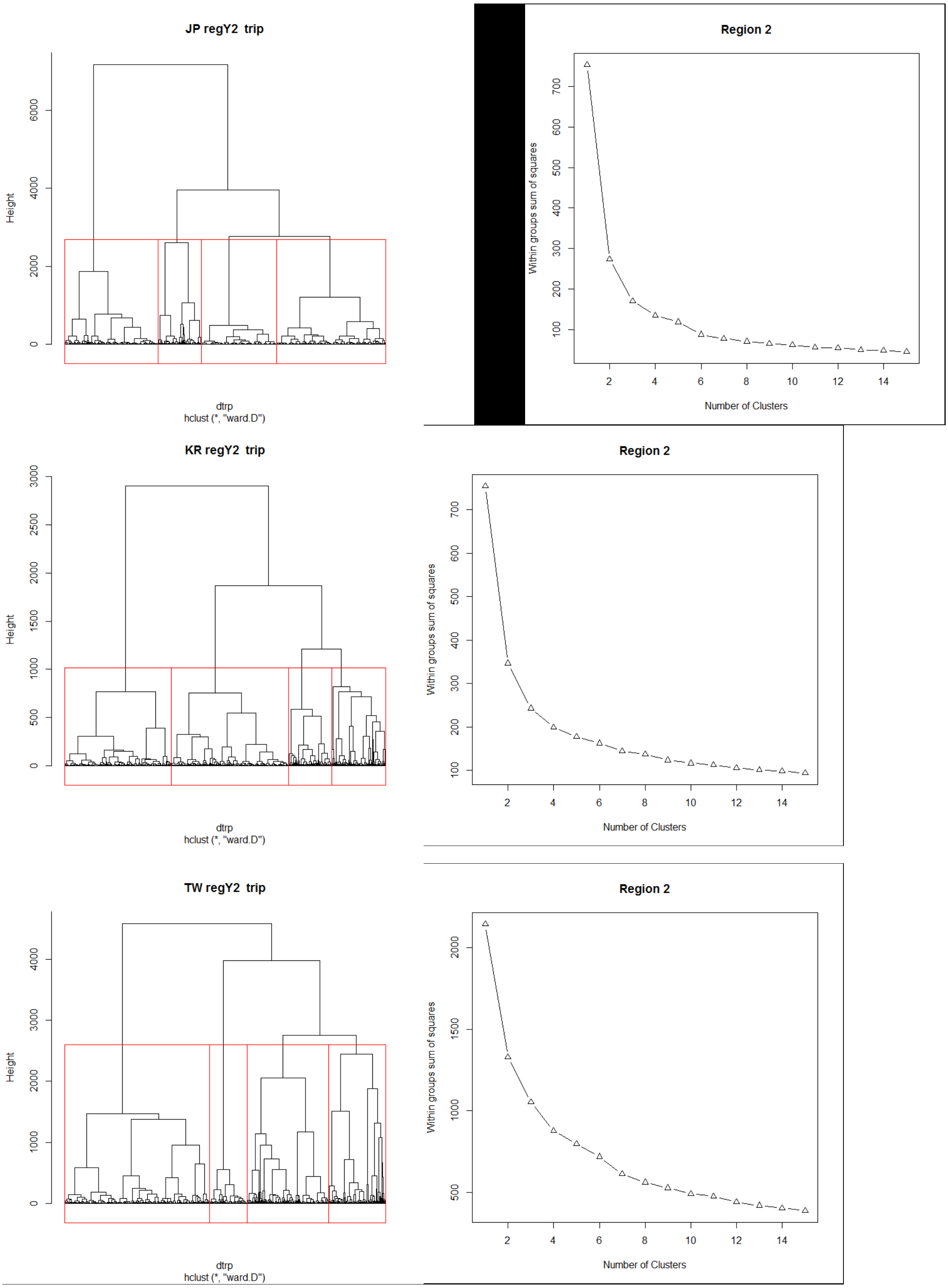


*Figure 14: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 2 (western tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

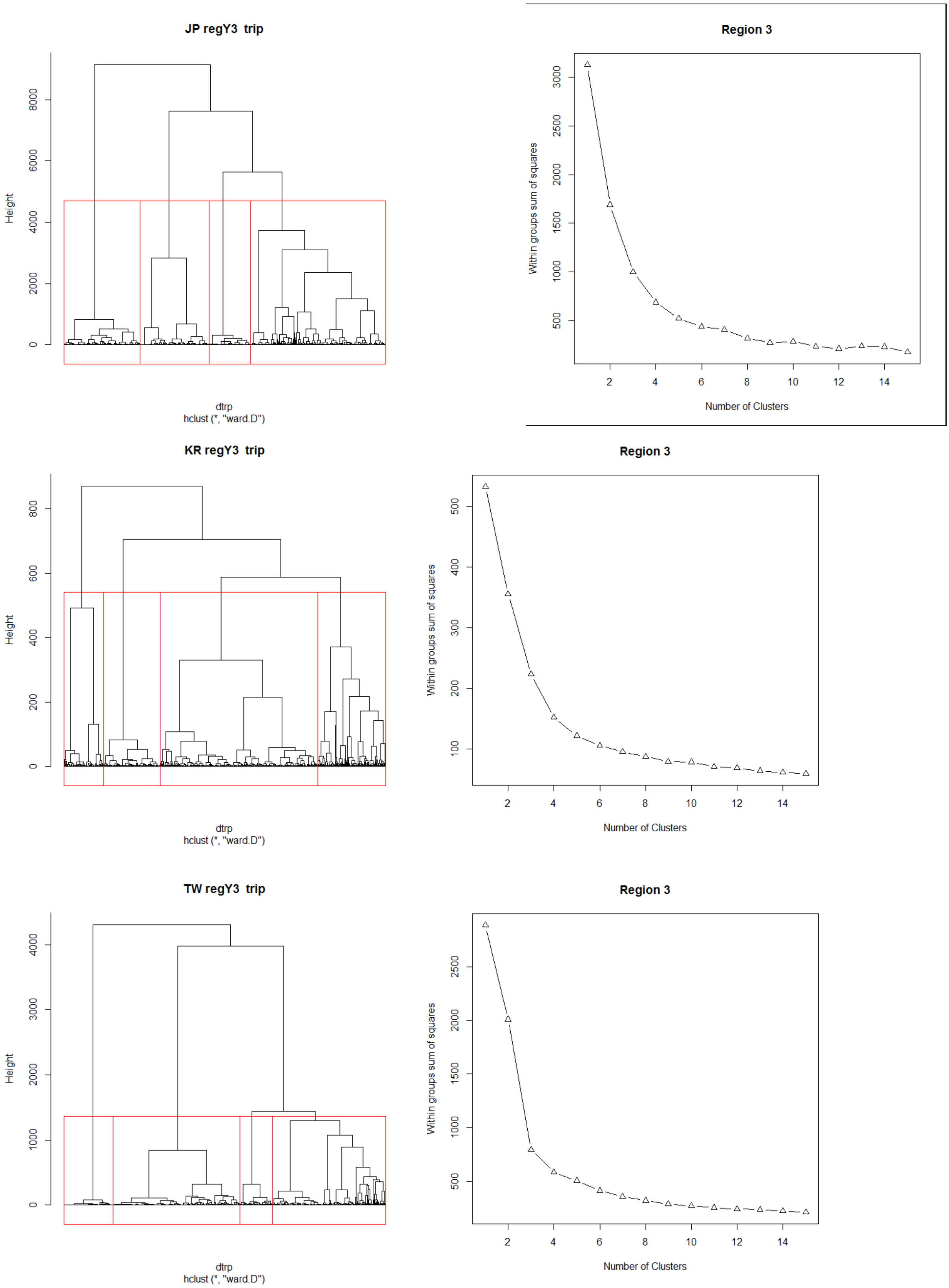


*Figure 15: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 3 (western temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

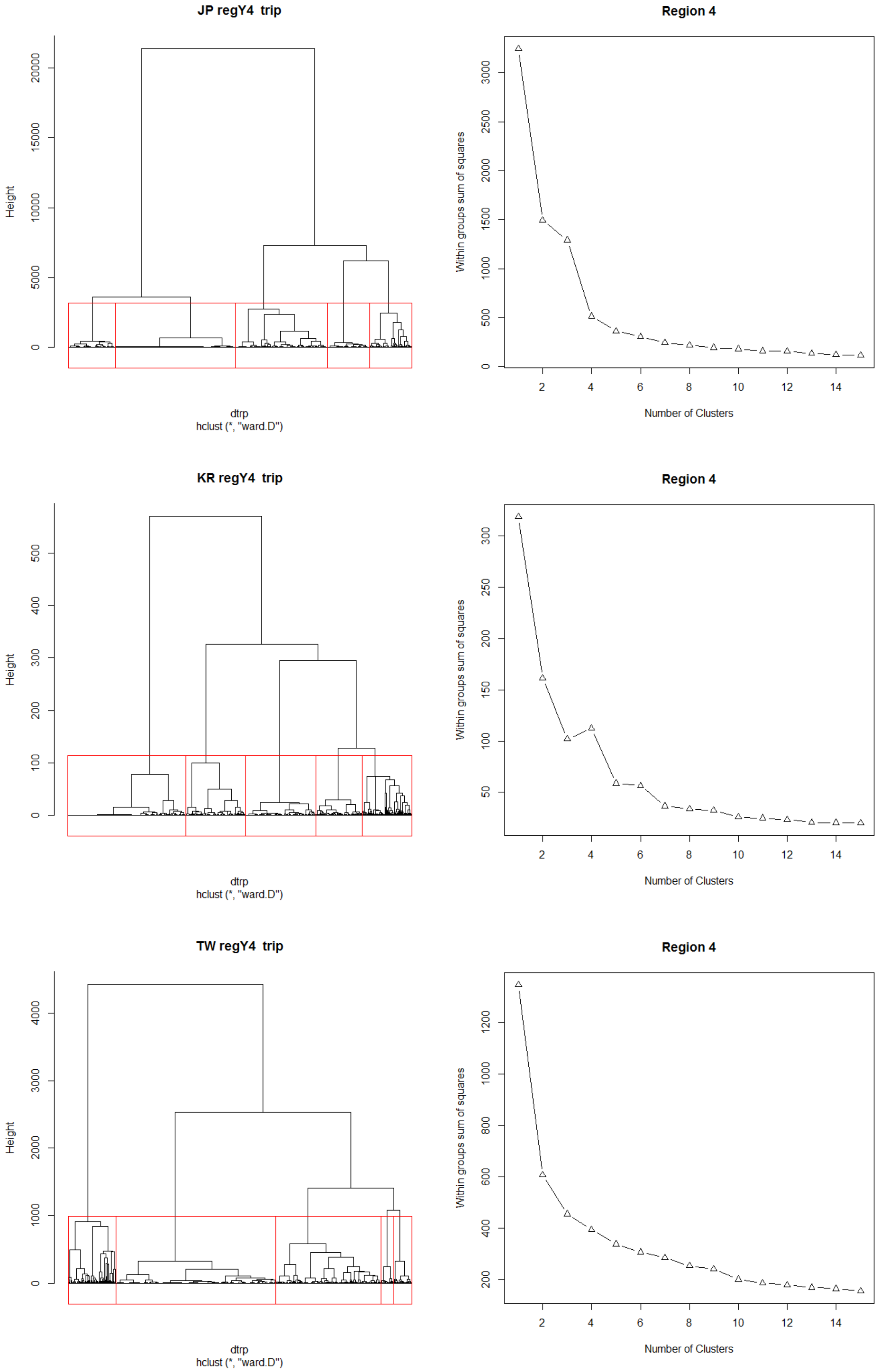


*Figure 16: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 4 (eastern temperate) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

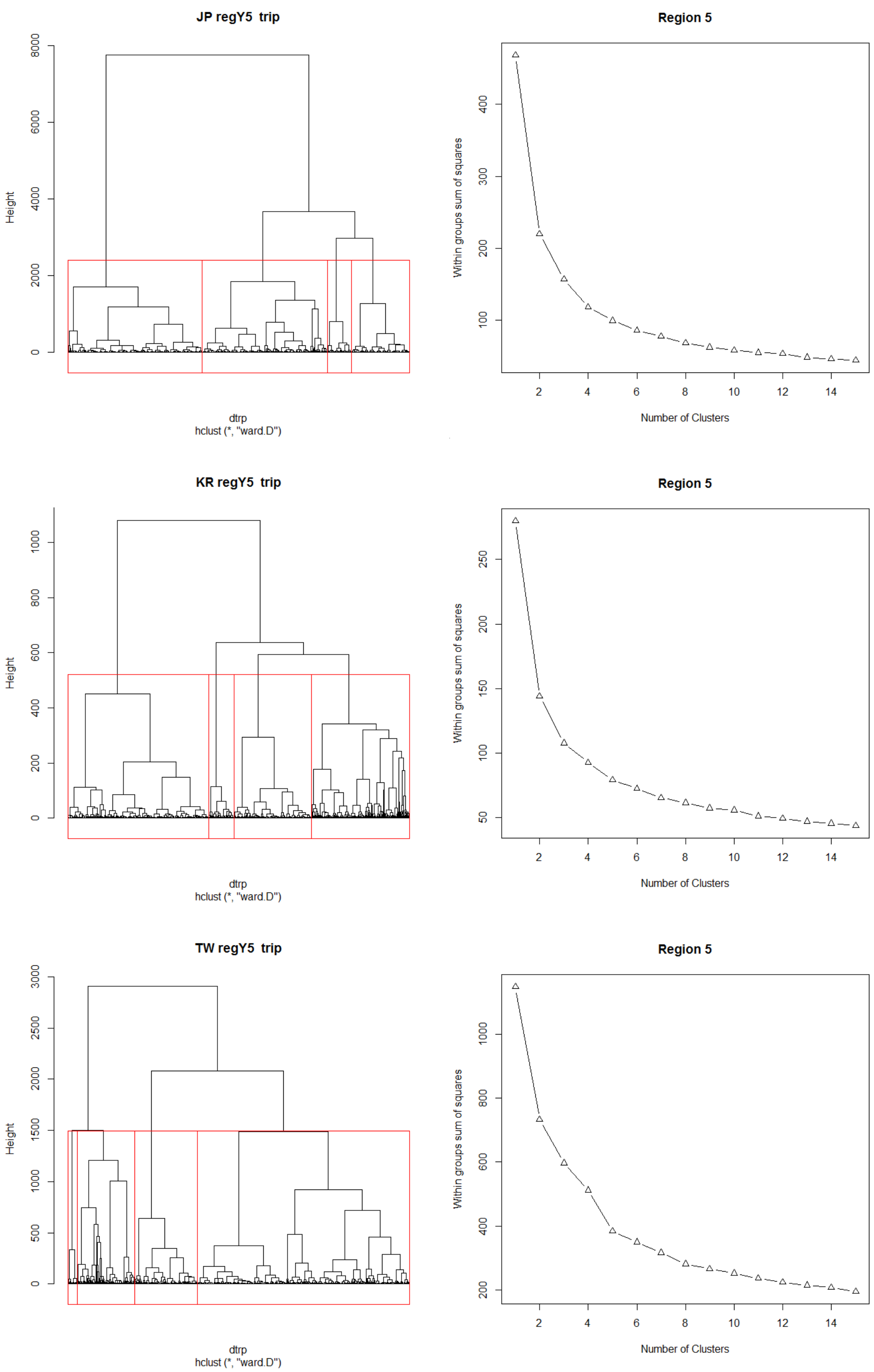


*Figure 17: Plots showing analyses to estimate the number of distinct classes of species composition in yellowfin region 5 (eastern tropical) for Japanese, Korean, and Taiwanese effort. These are based on a hierarchical Ward clustering analysis of trip-level data (left); and within-group sums of squares from kmeans analyses with a range of numbers of clusters (right).*

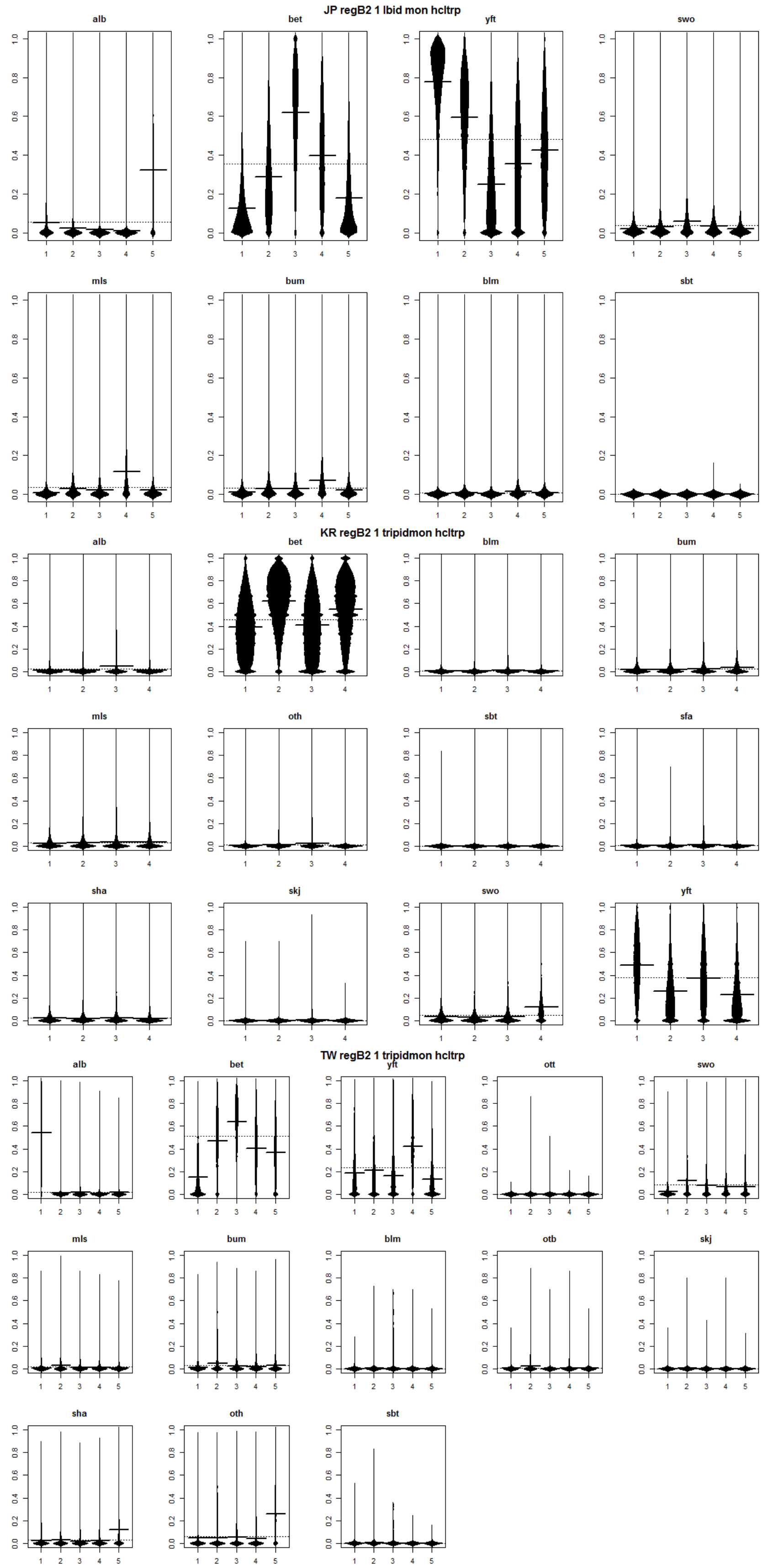


*Figure 18: Beanplots for bigeye region 1 (western tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

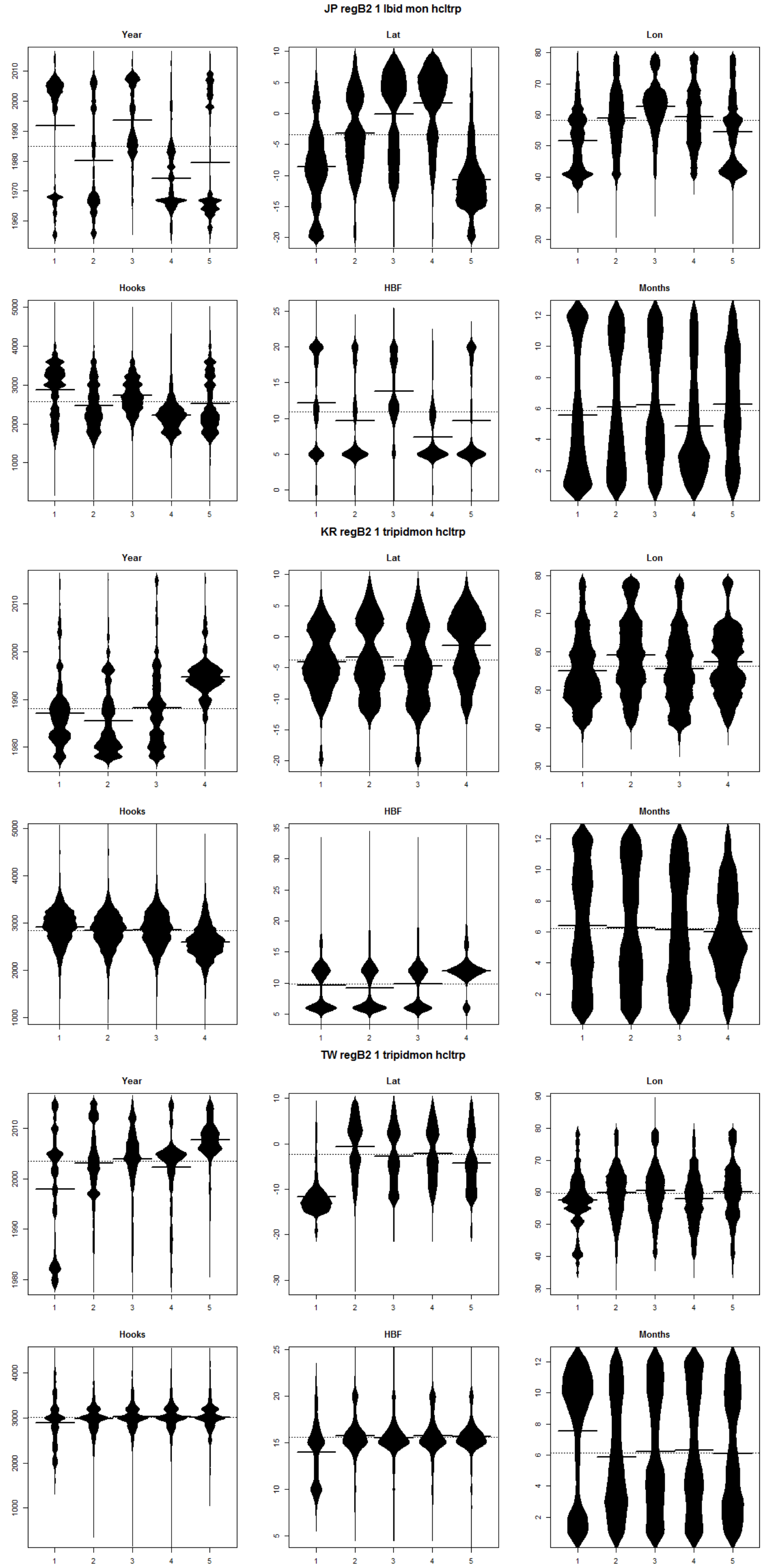


*Figure 19: Beanplots for bigeye region 1 (western tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

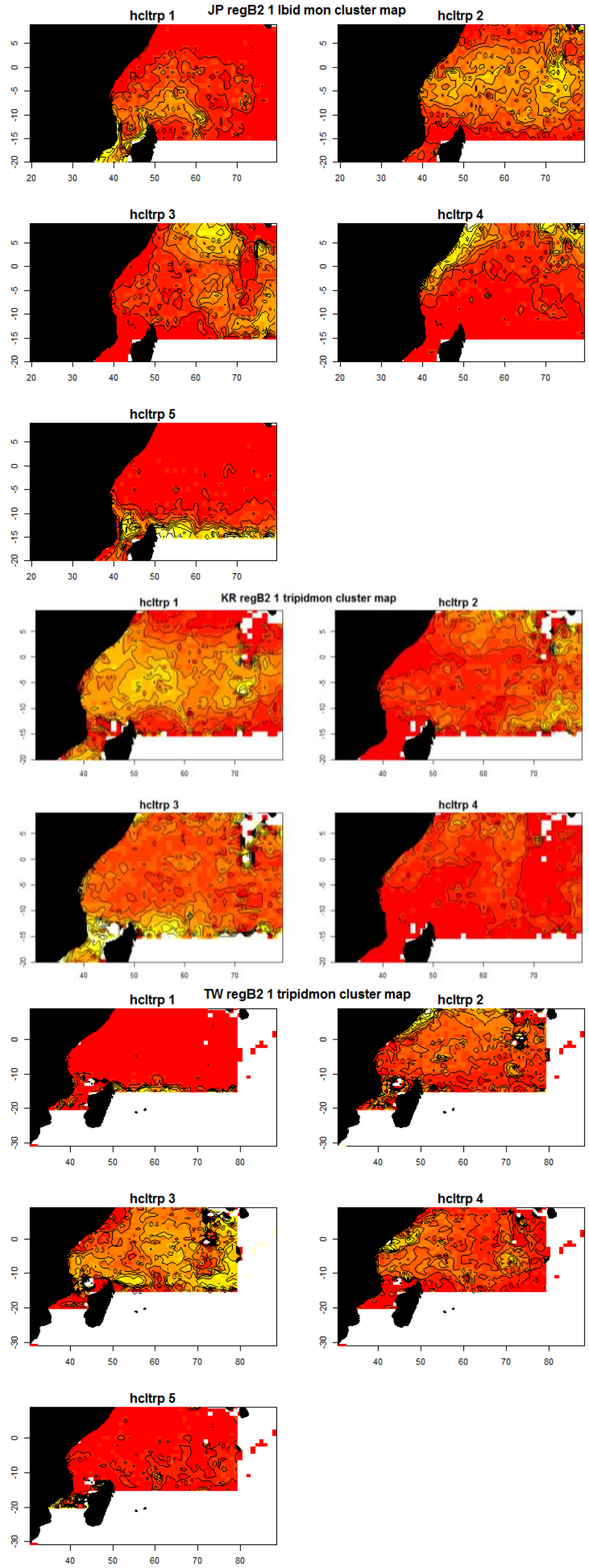


*Figure 20: Maps of the spatial distributions of clusters in region 1 (western tropical), for Japanese, Korean, and Taiwanese effort.*

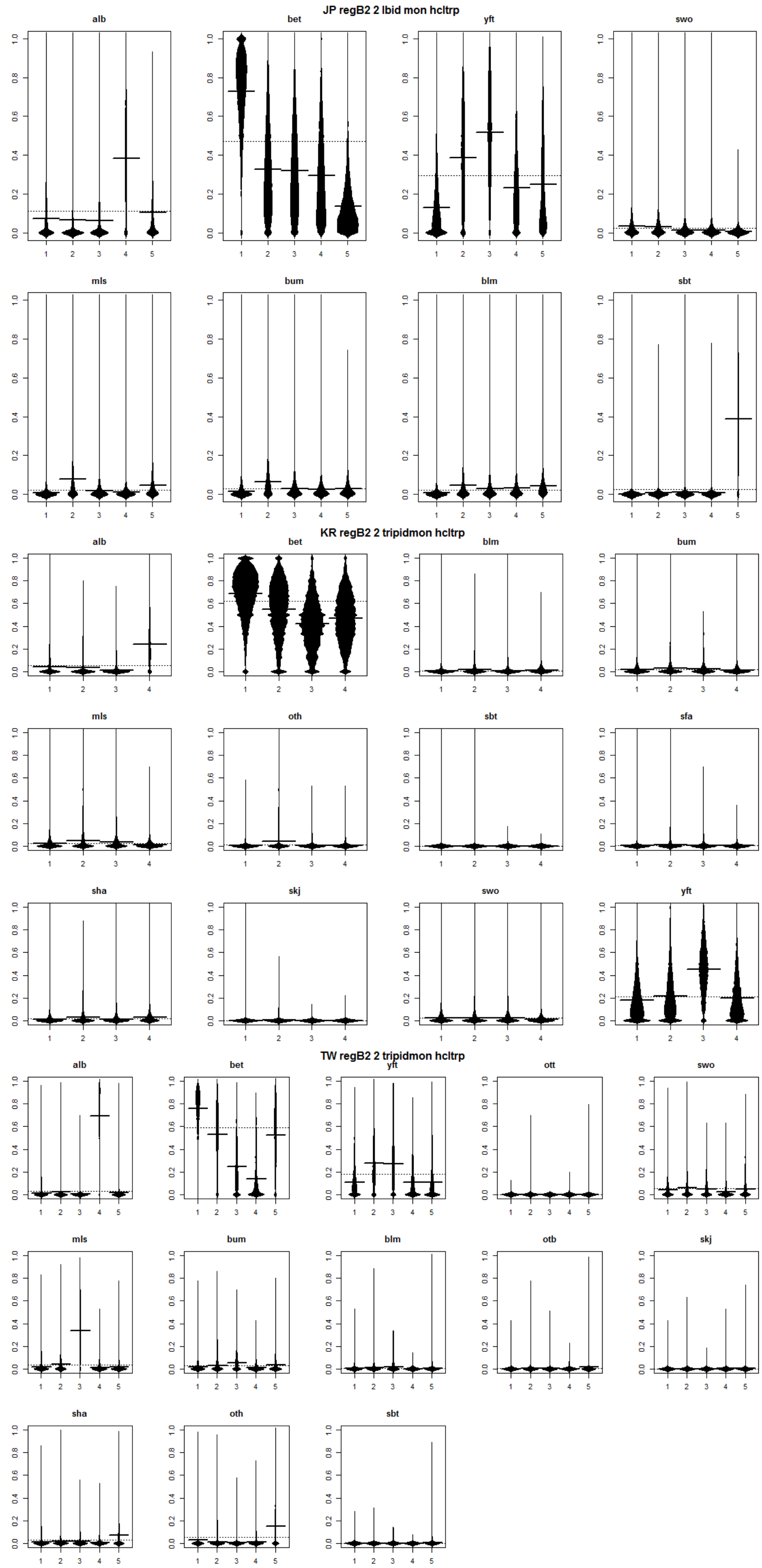


*Figure 21: Beanplots for bigeye region 2 (eastern tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

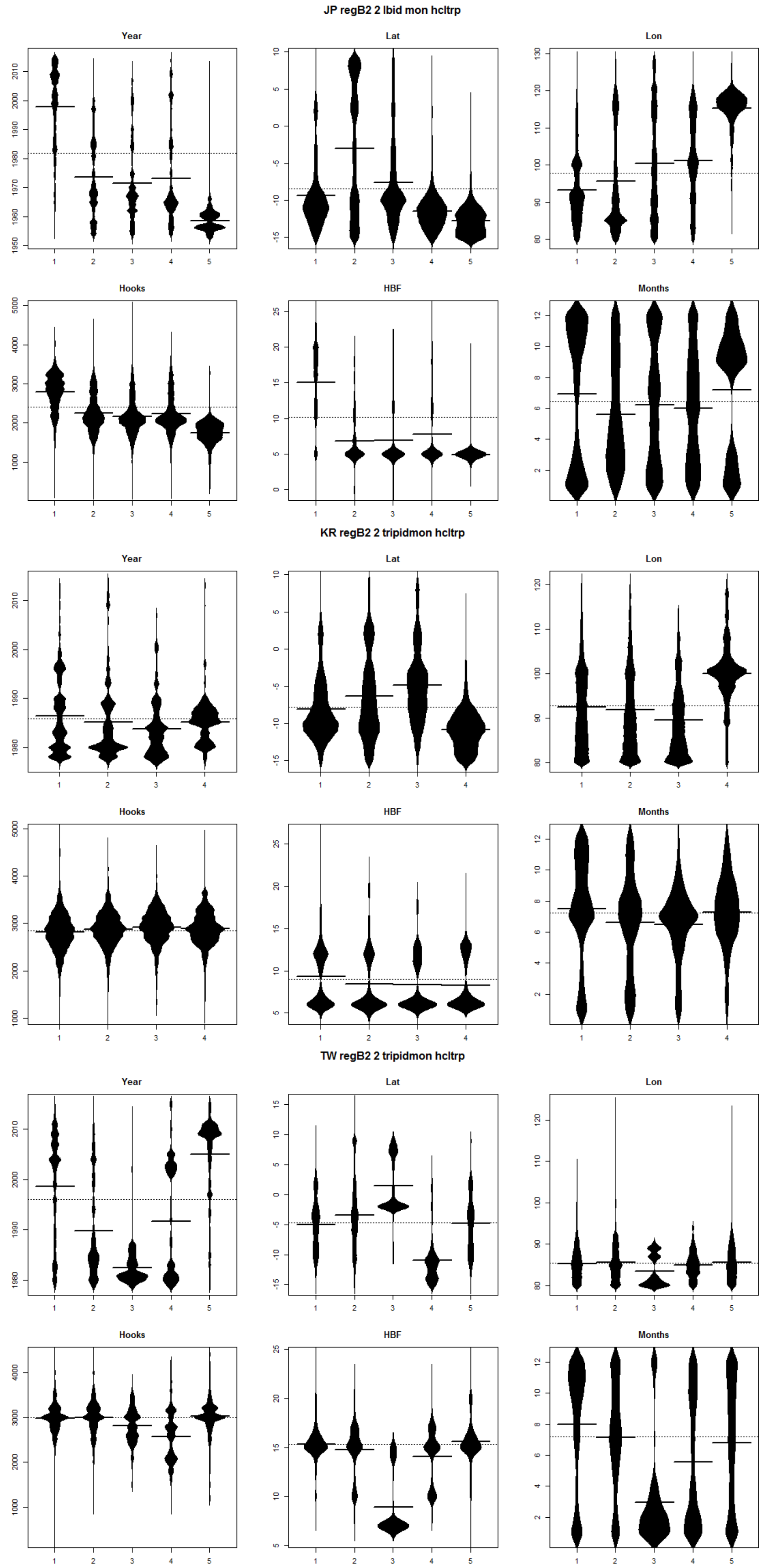


*Figure 22: Beanplots for bigeye region 2 (eastern tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

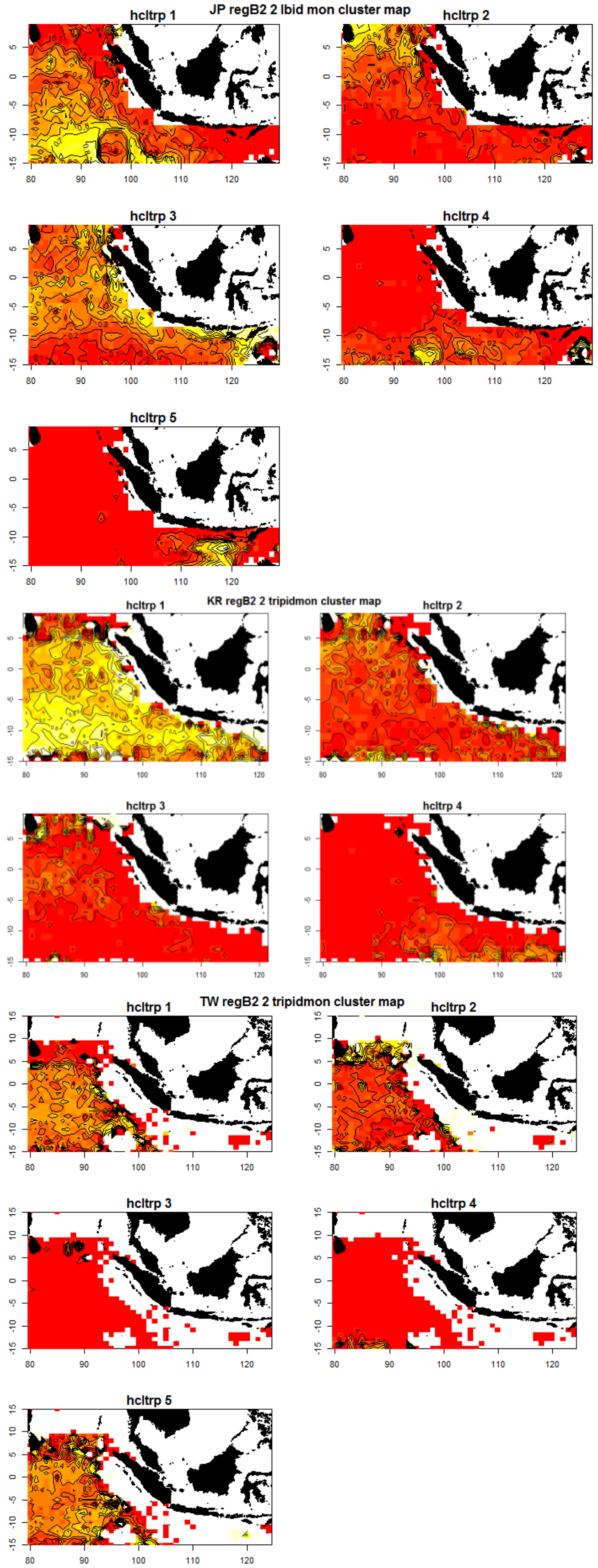


*Figure 23: Maps of the spatial distributions of clusters in region 2 (eastern tropical), for Japanese, Korean, and Taiwanese effort.*

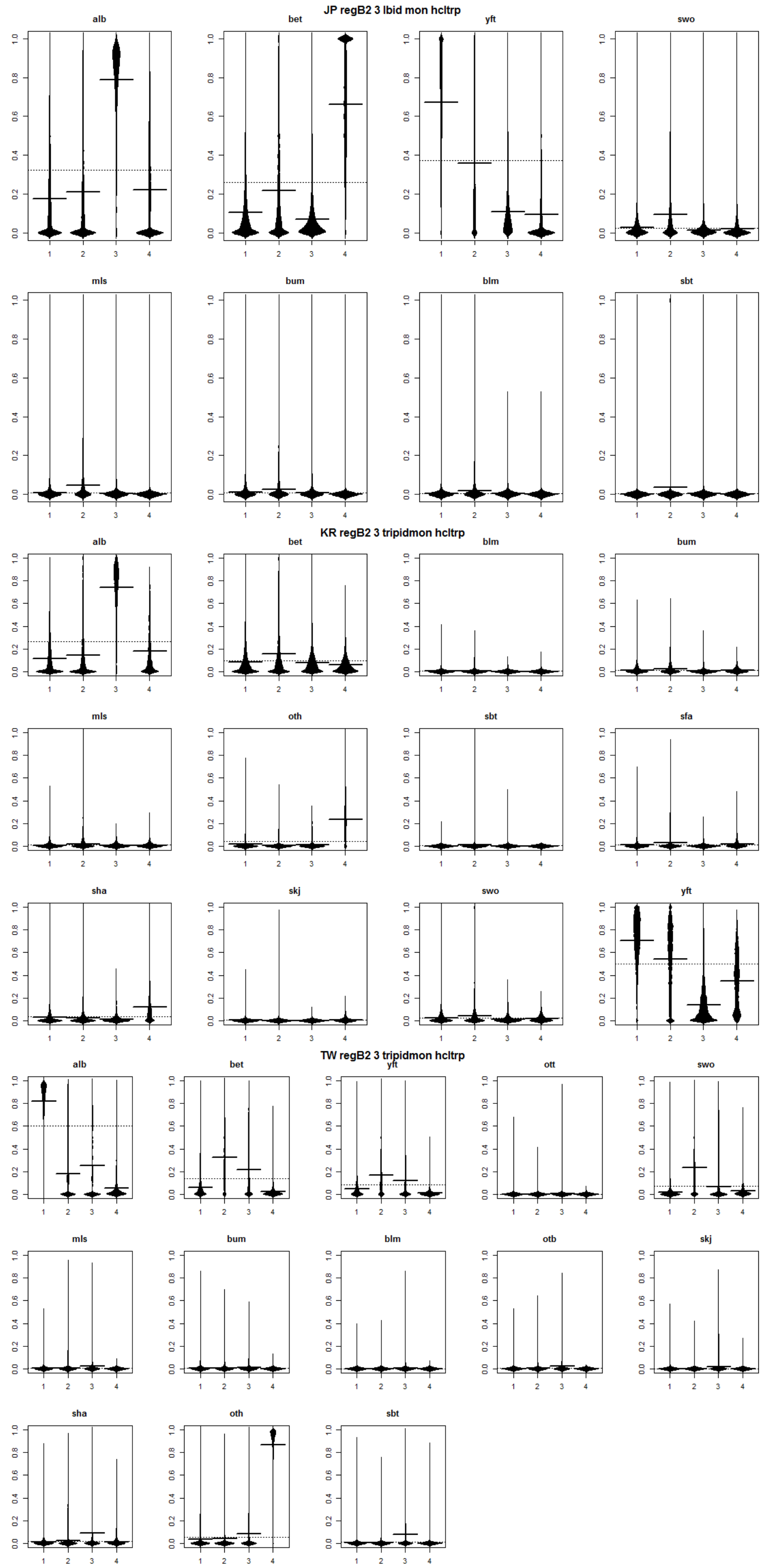


*Figure 24: Beanplots for bigeye region 3 (western temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

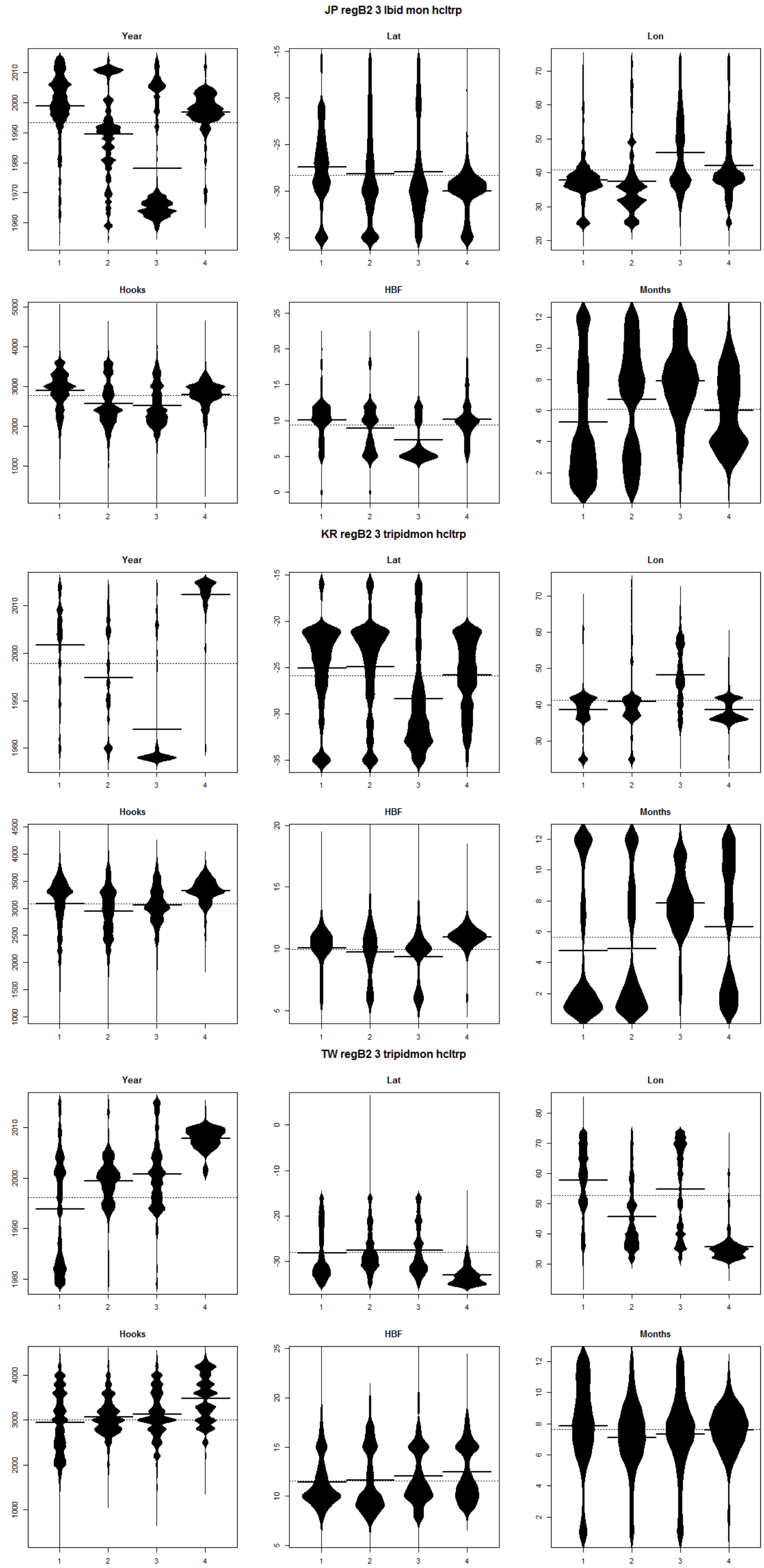


*Figure 25: Beanplots for bigeye region 3 (western temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

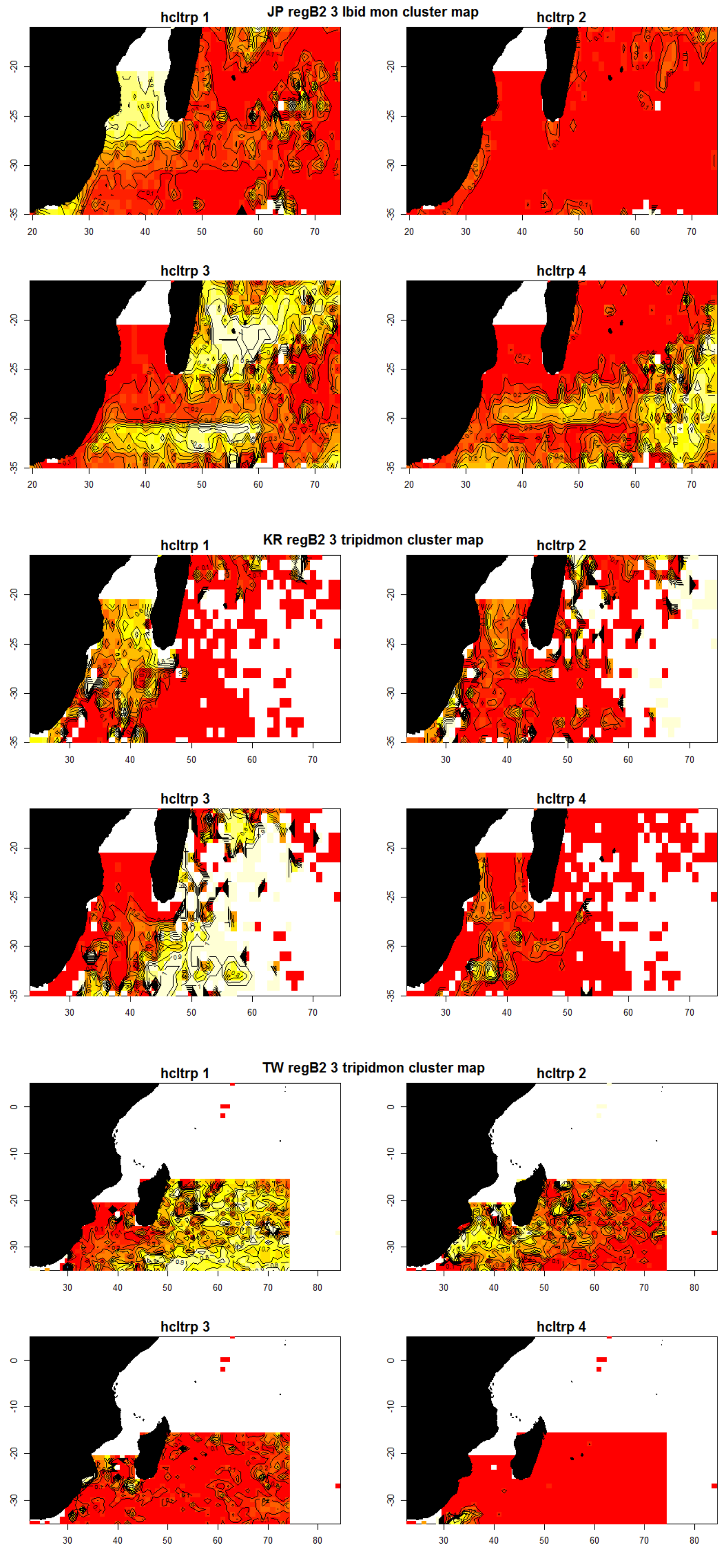


*Figure 26: Maps of the spatial distributions of clusters in region 3 (western temperate), for Japanese, Korean, and Taiwanese effort.*

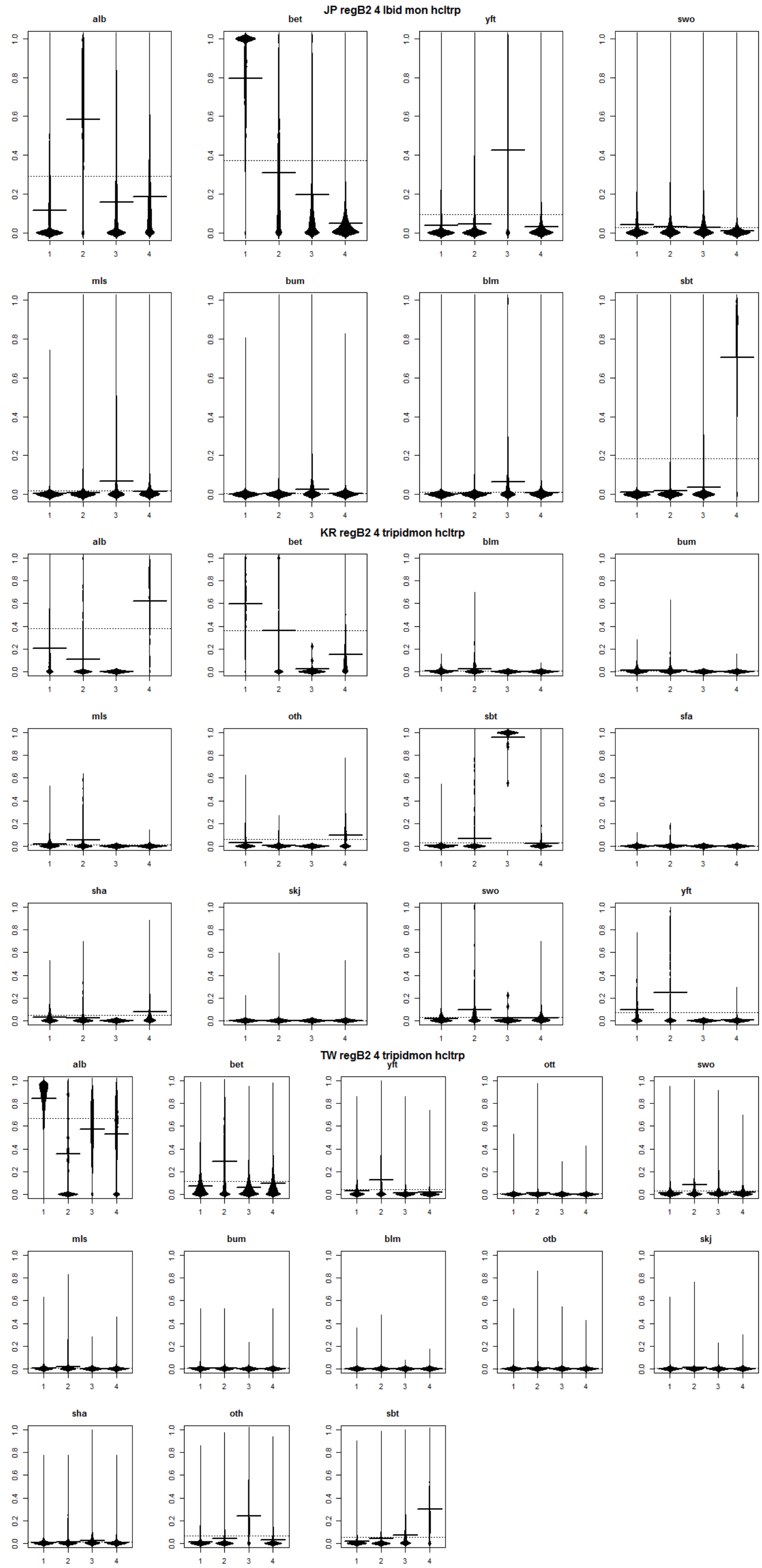


*Figure 27: Beanplots for bigeye region 4 (eastern temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

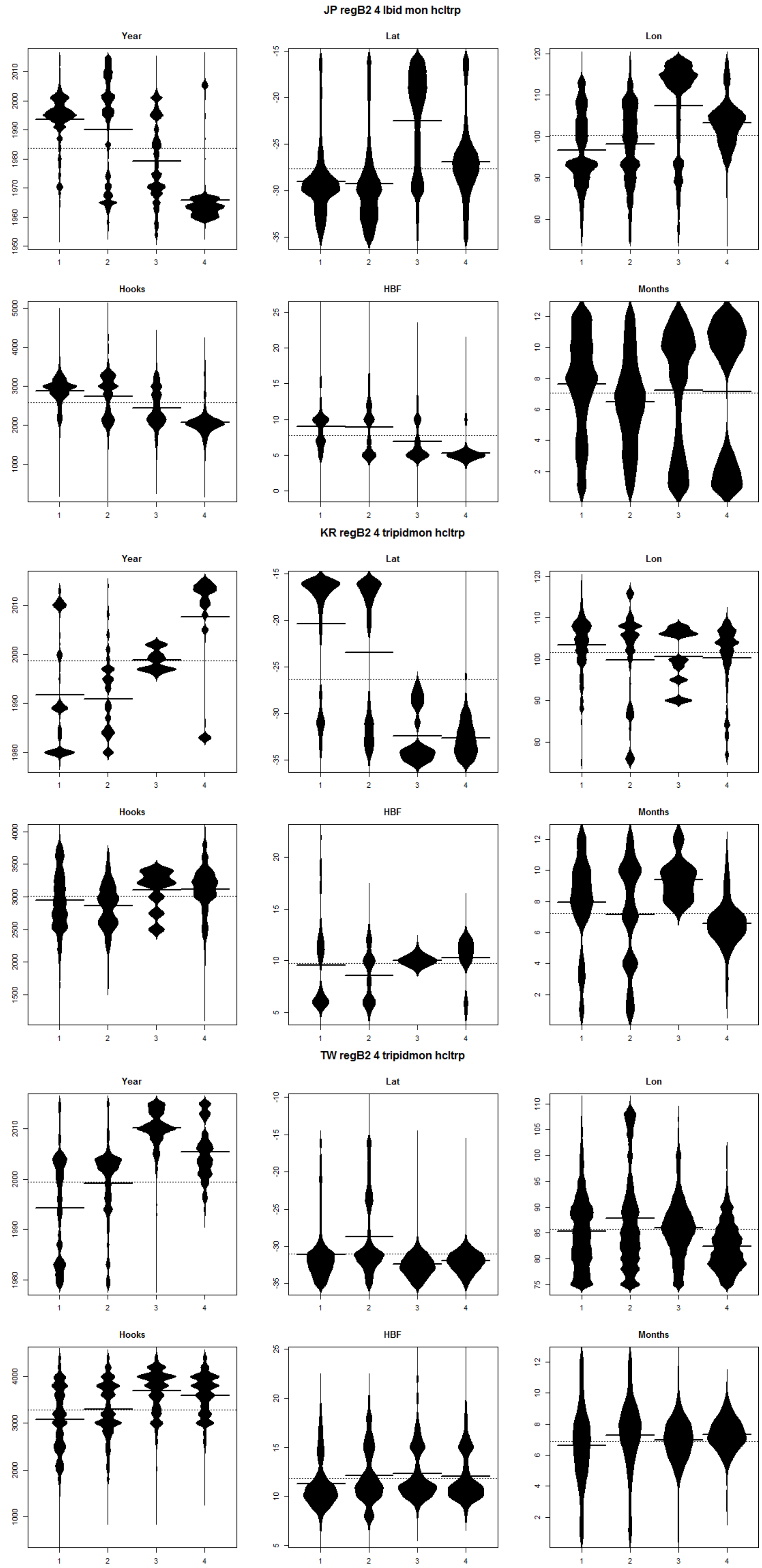


*Figure 28: Beanplots for bigeye region 4 (eastern temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

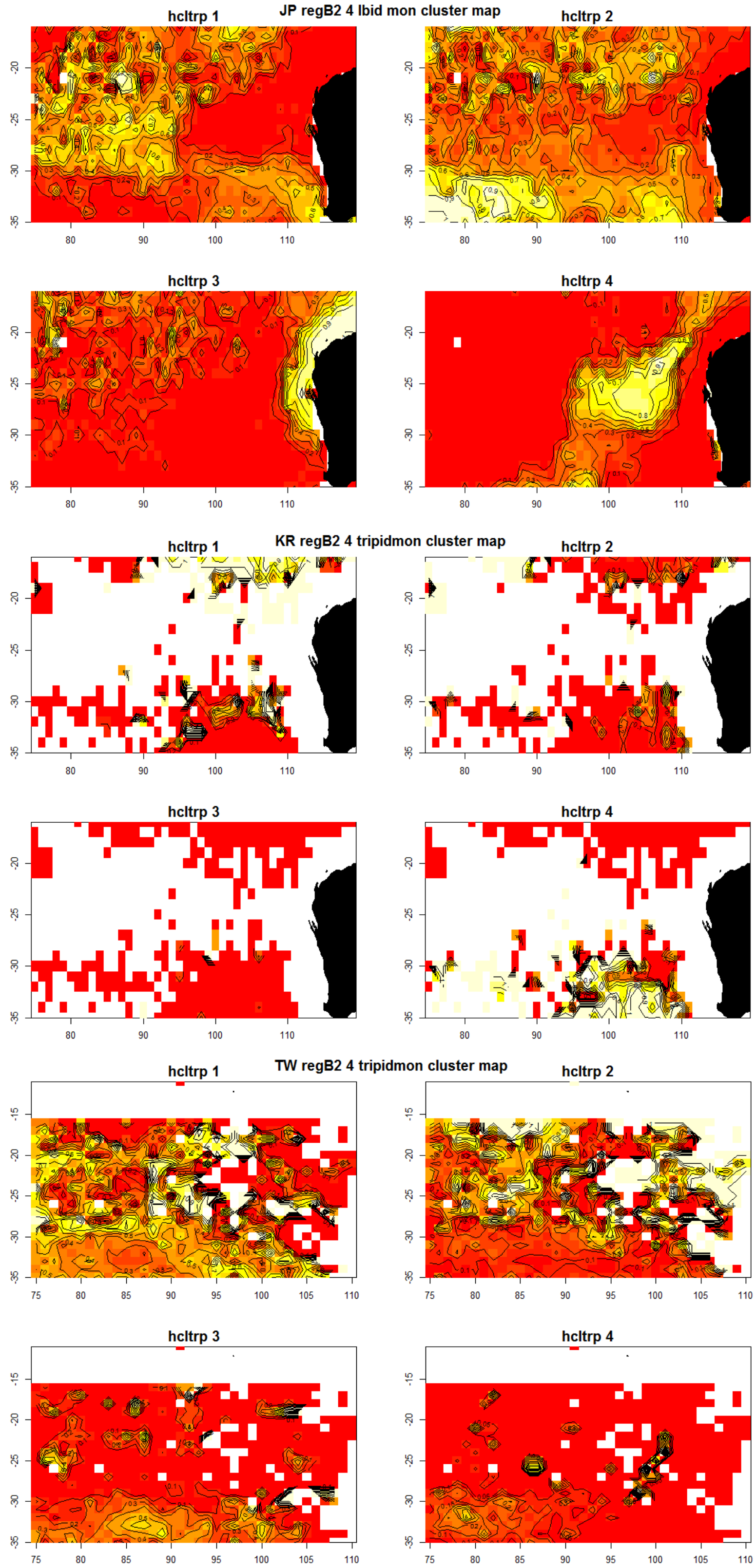


*Figure 29: Maps of the spatial distributions of clusters in region 4 (eastern temperate), for Japanese, Korean, and Taiwanese effort.*

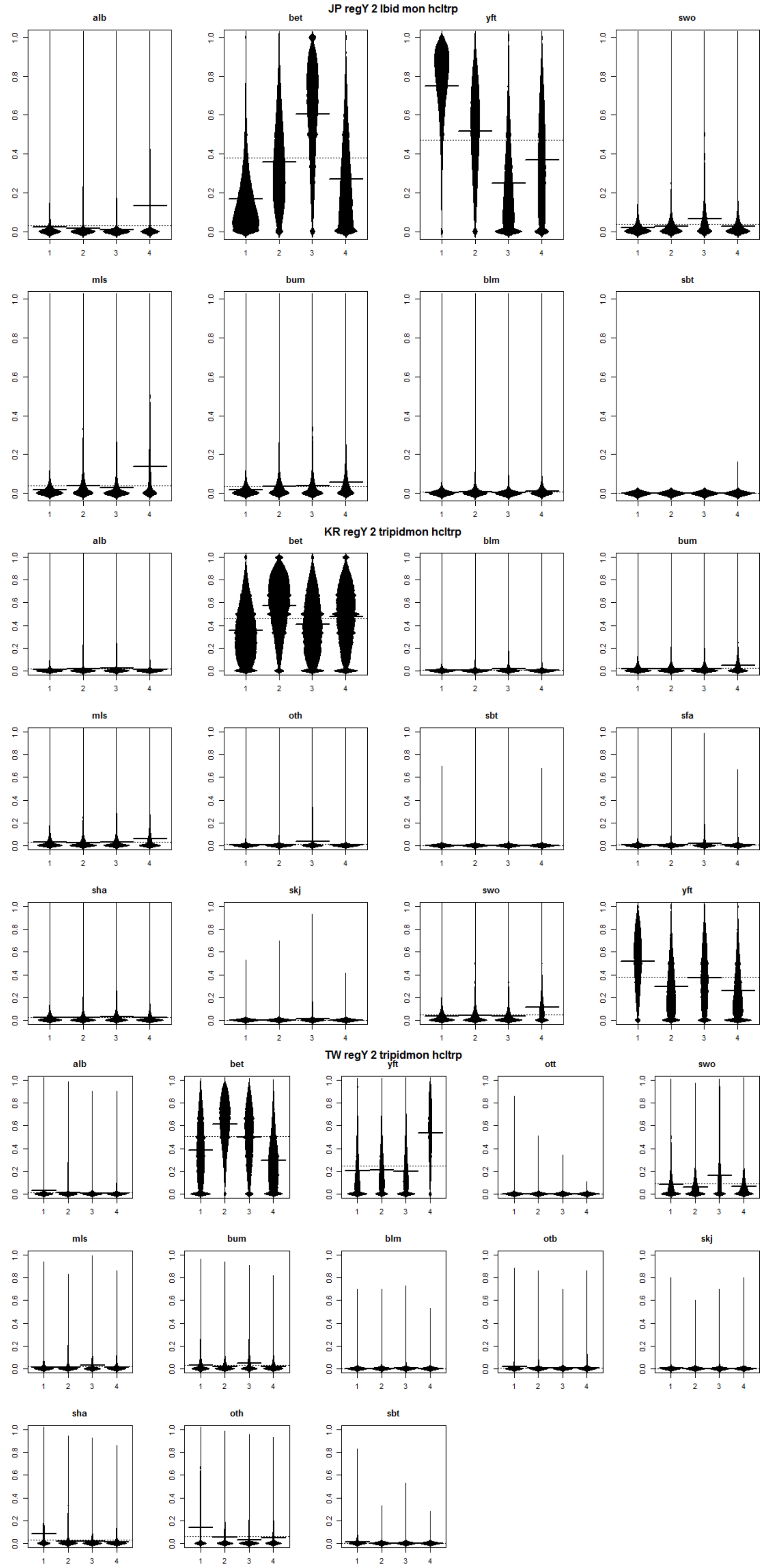


*Figure 30: Beanplots for yellowfin region 2 (western tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

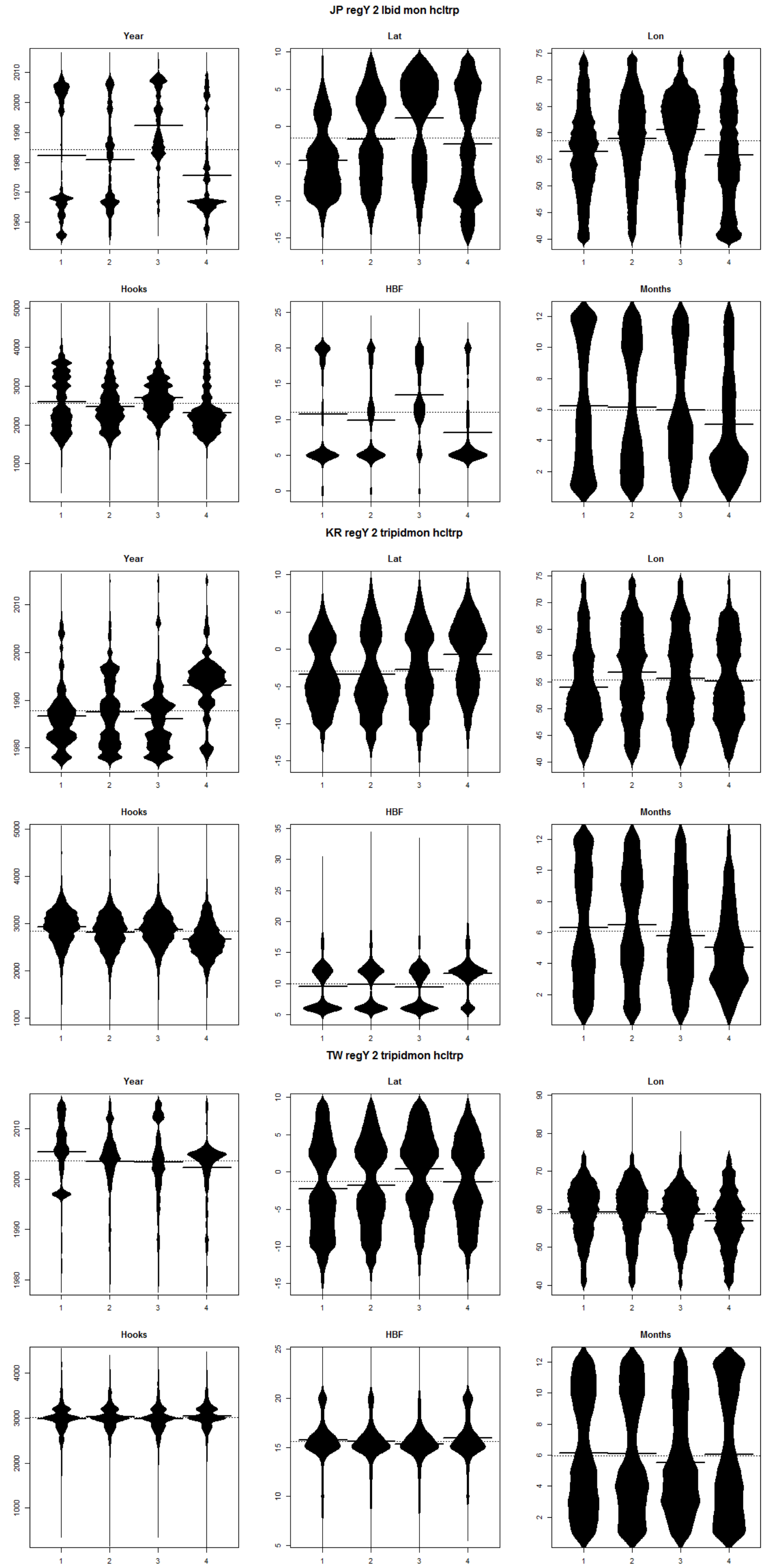


*Figure 31: Beanplots for yellowfin region 2 (western tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

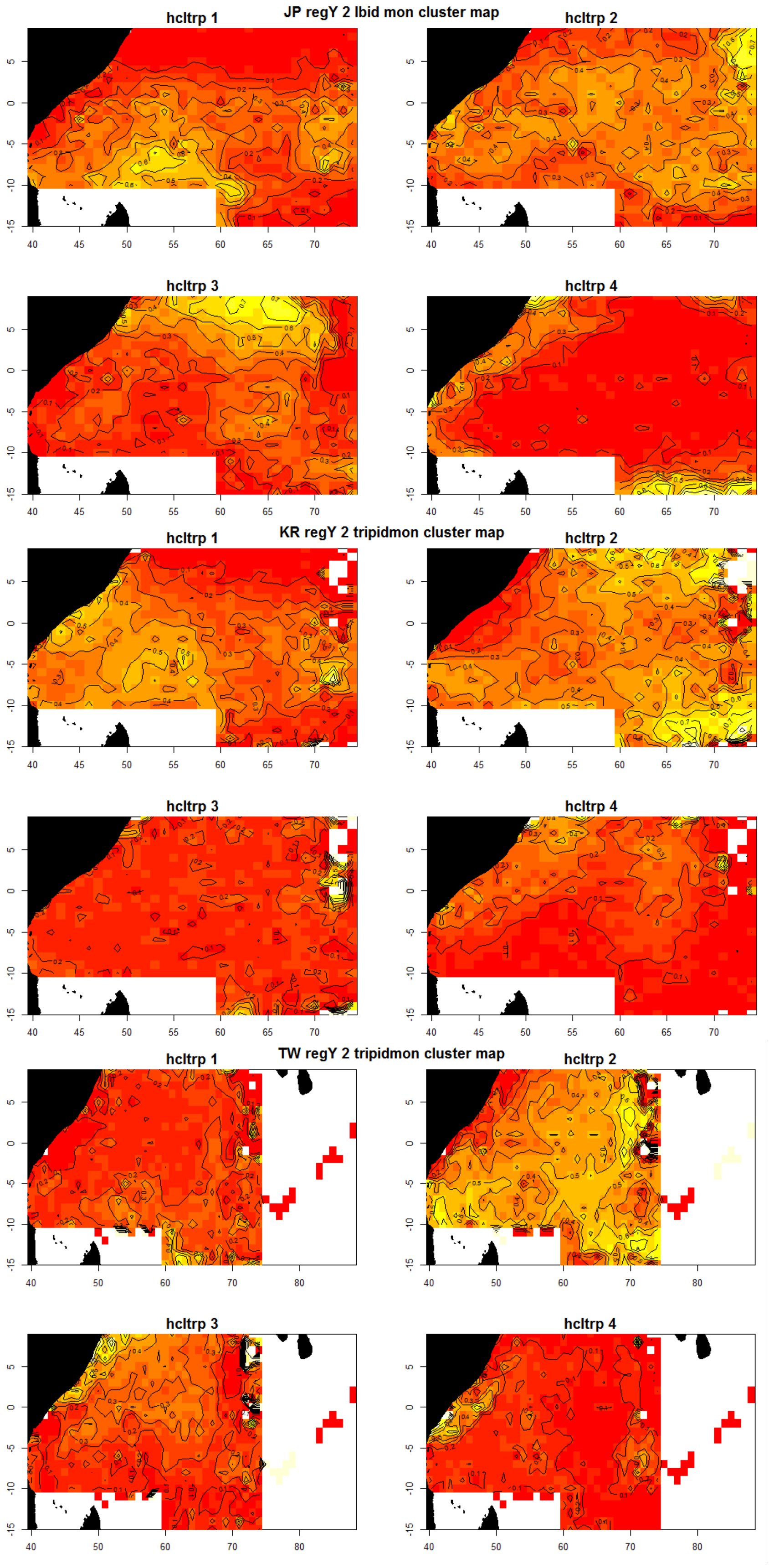


*Figure 32: Maps of the spatial distributions of clusters in region 2 (western tropical), for Japanese, Korean, and Taiwanese effort.*

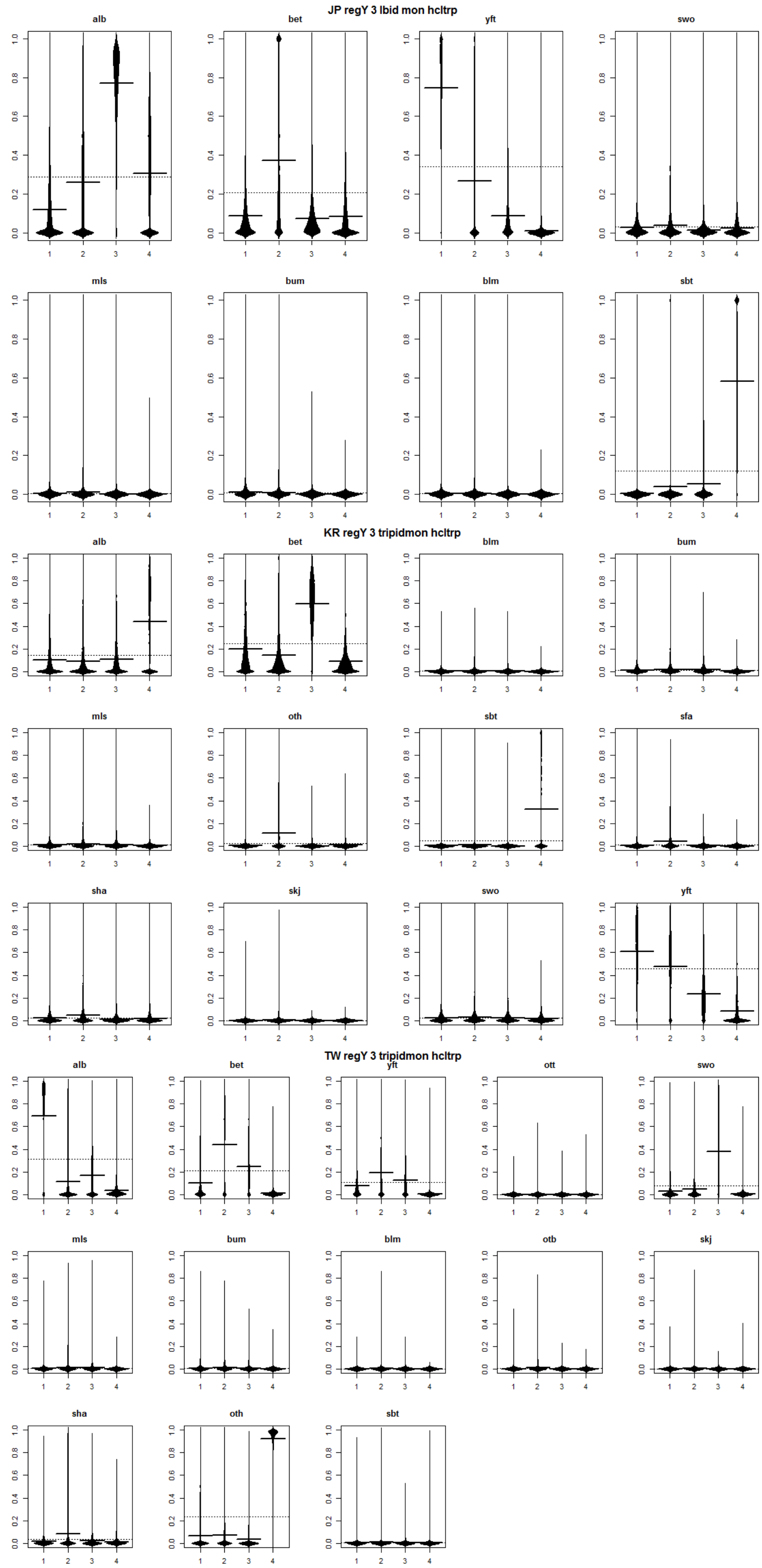


*Figure 33: Beanplots for yellowfin region 3 (western temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

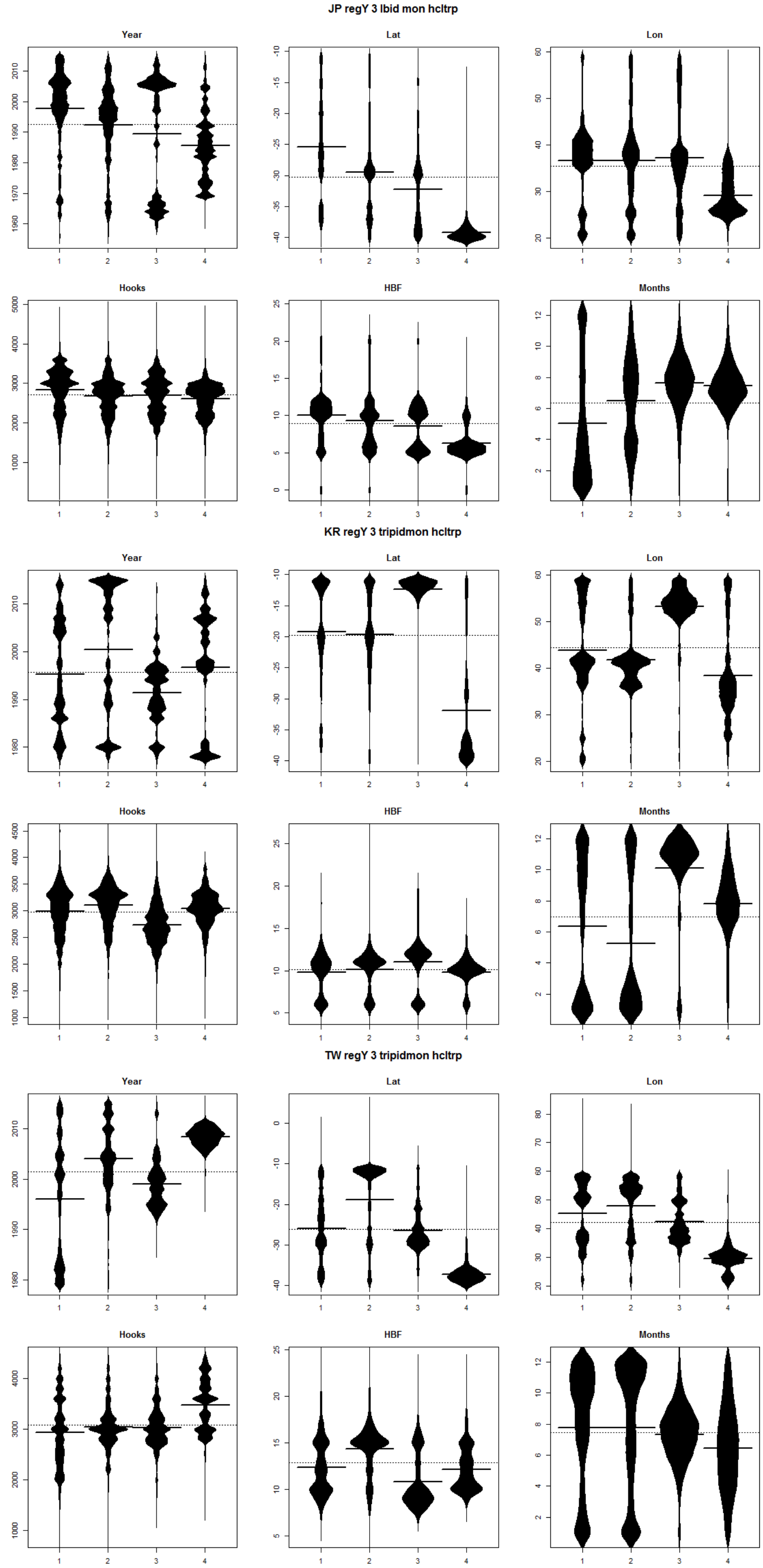


*Figure 34: Beanplots for yellowfin region 3 (western temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

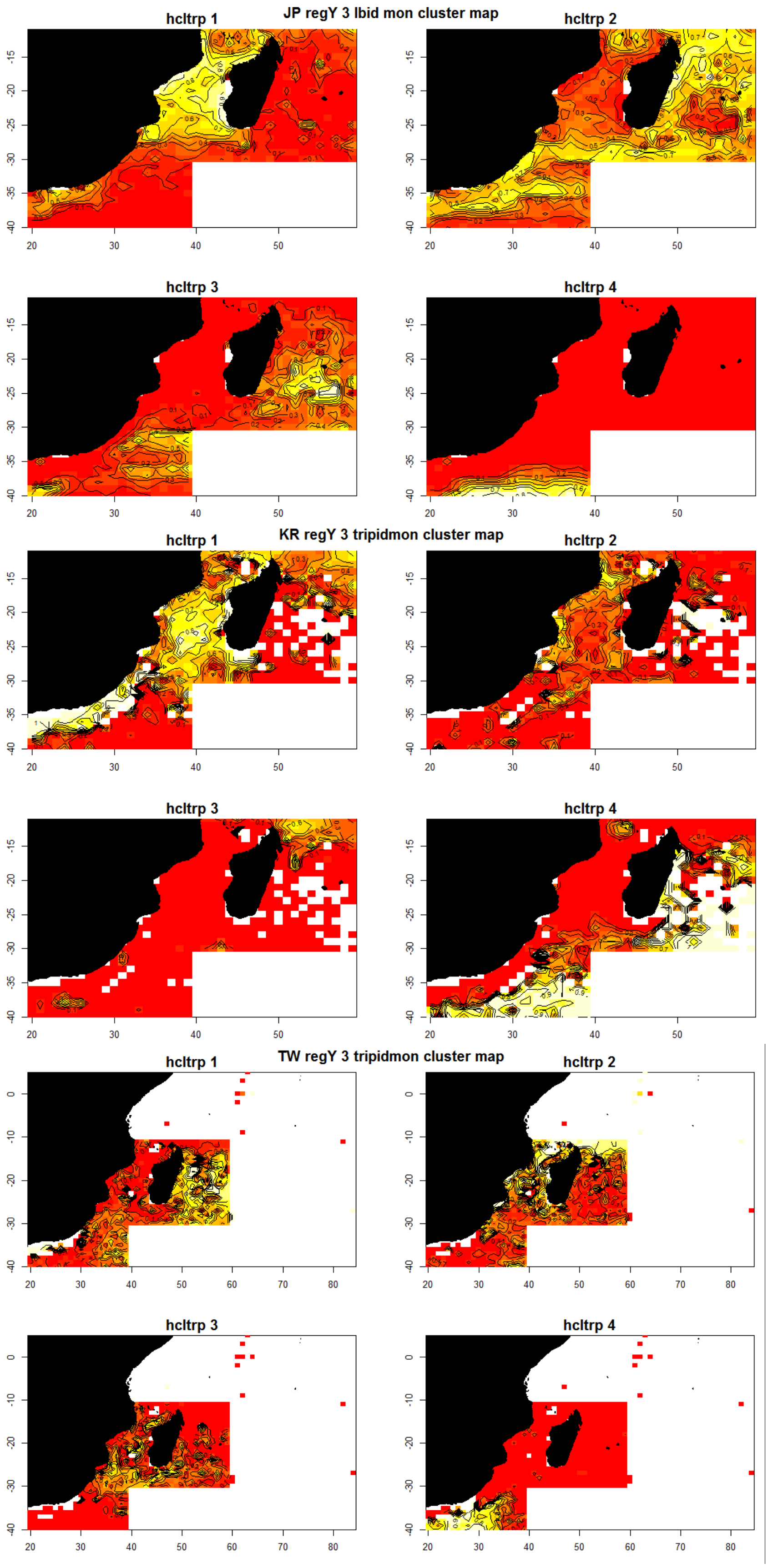


*Figure 35: Maps of the spatial distributions of clusters in region 3 (western temperate), for Japanese, Korean, and Taiwanese effort.*

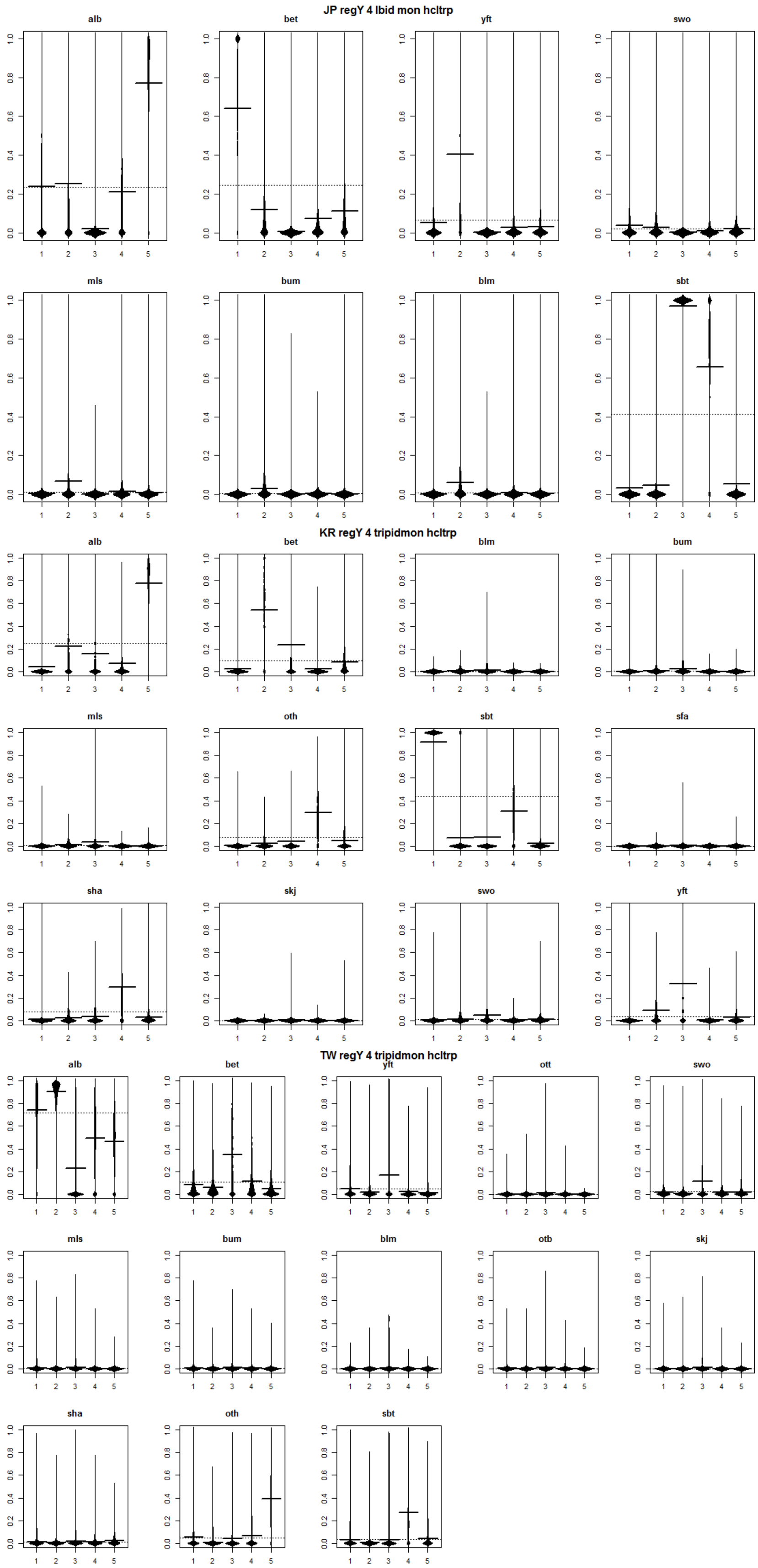


*Figure 36: Beanplots for yellowfin region 4 (eastern temperate) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

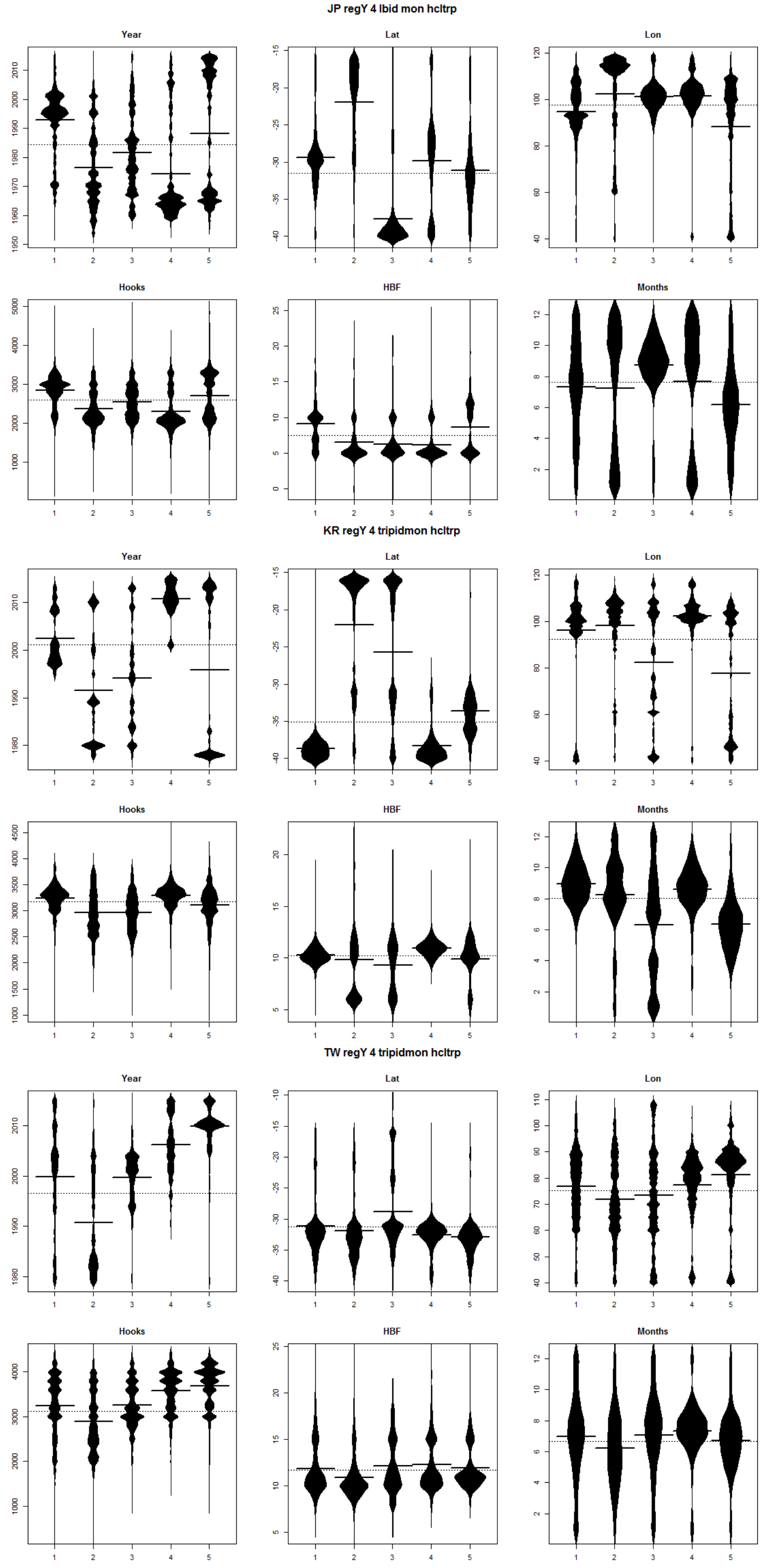


*Figure 37: Beanplots for yellowfin region 4 (eastern temperate) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

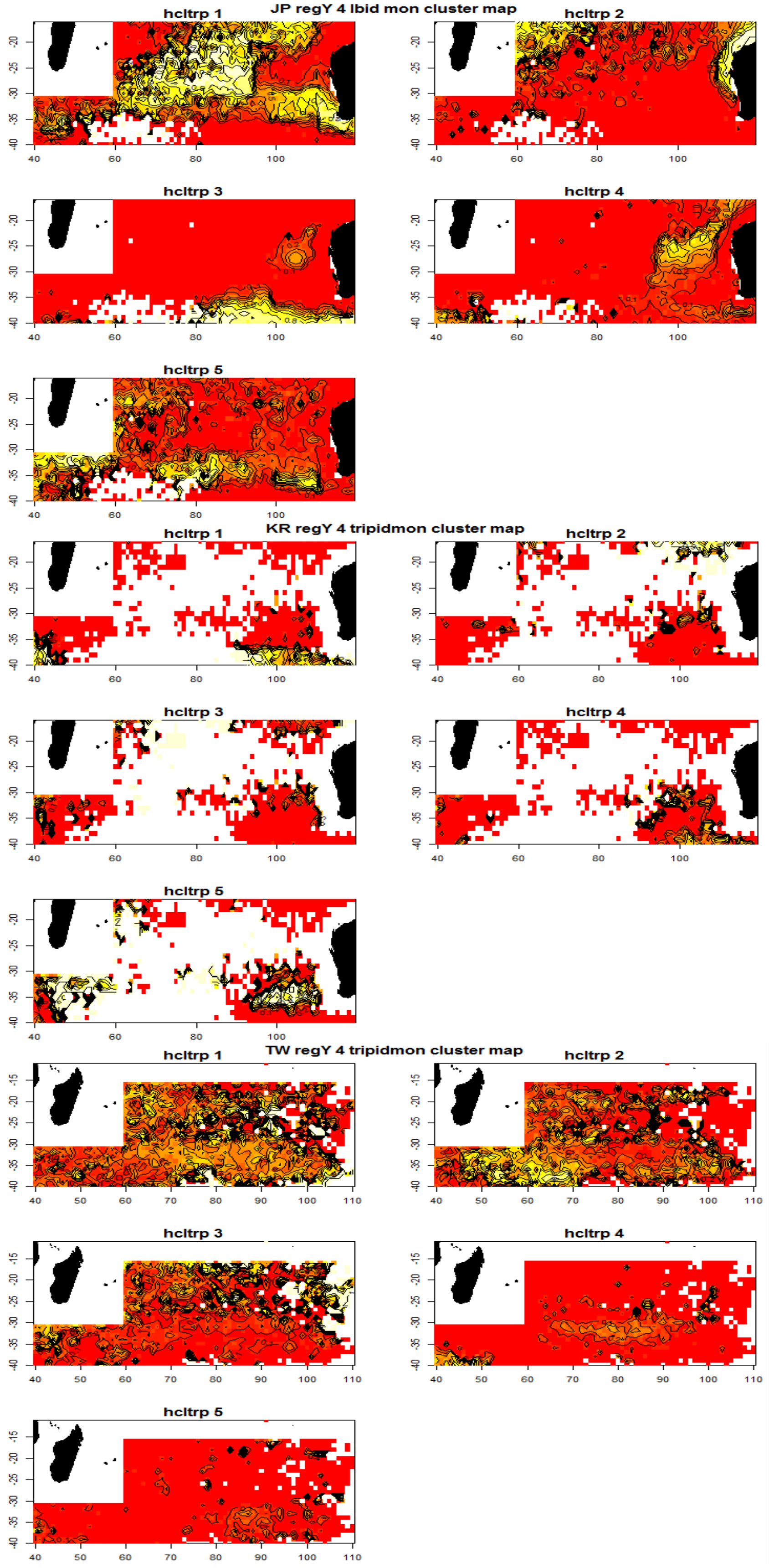


*Figure 38: Maps of the spatial distributions of clusters in region 4 (eastern temperate), for Japanese, Korean, and Taiwanese effort.*

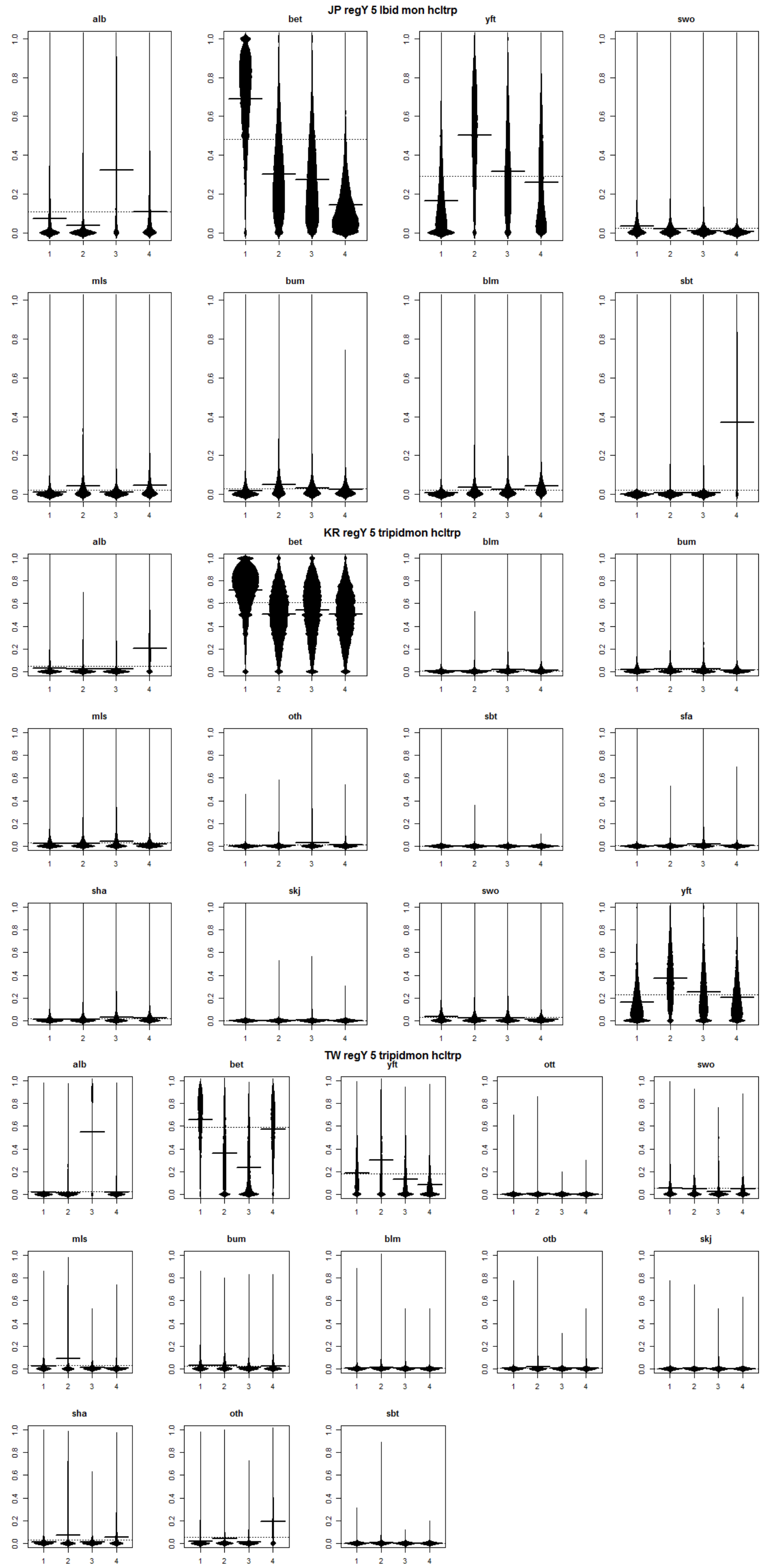


*Figure 39: Beanplots for yellowfin region 5 (eastern tropical) showing species composition by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

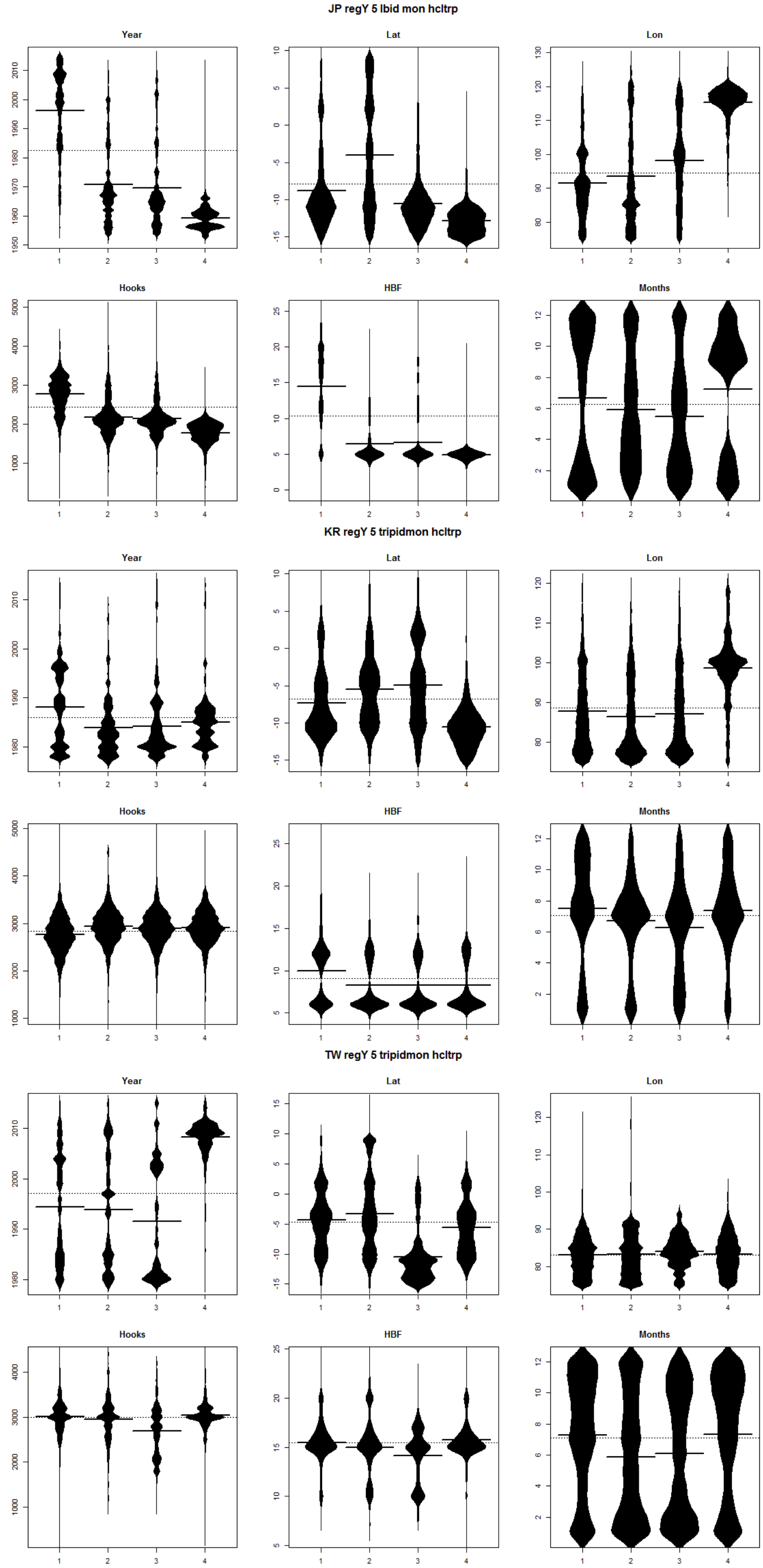


*Figure 40: Beanplots for yellowfin region 5 (eastern tropical) showing number of sets versus covariate by cluster for Japanese (top), Korean (middle) and Taiwanese (bottom) effort. The horizontal bars indicate the medians.*

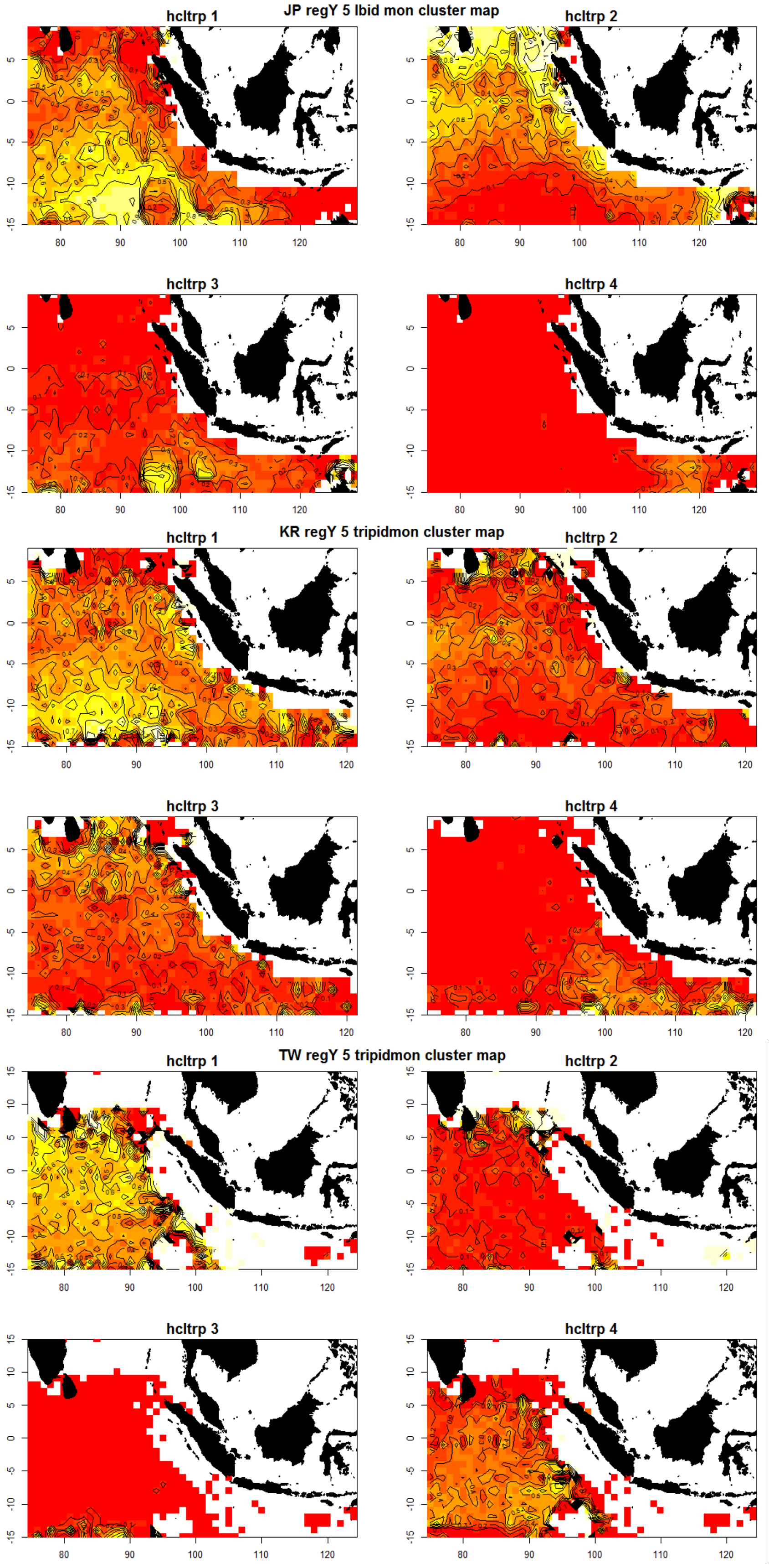


*Figure 41: Maps of the spatial distributions of clusters in region 5 (eastern tropical), for Japanese, Korean, and Taiwanese effort.*

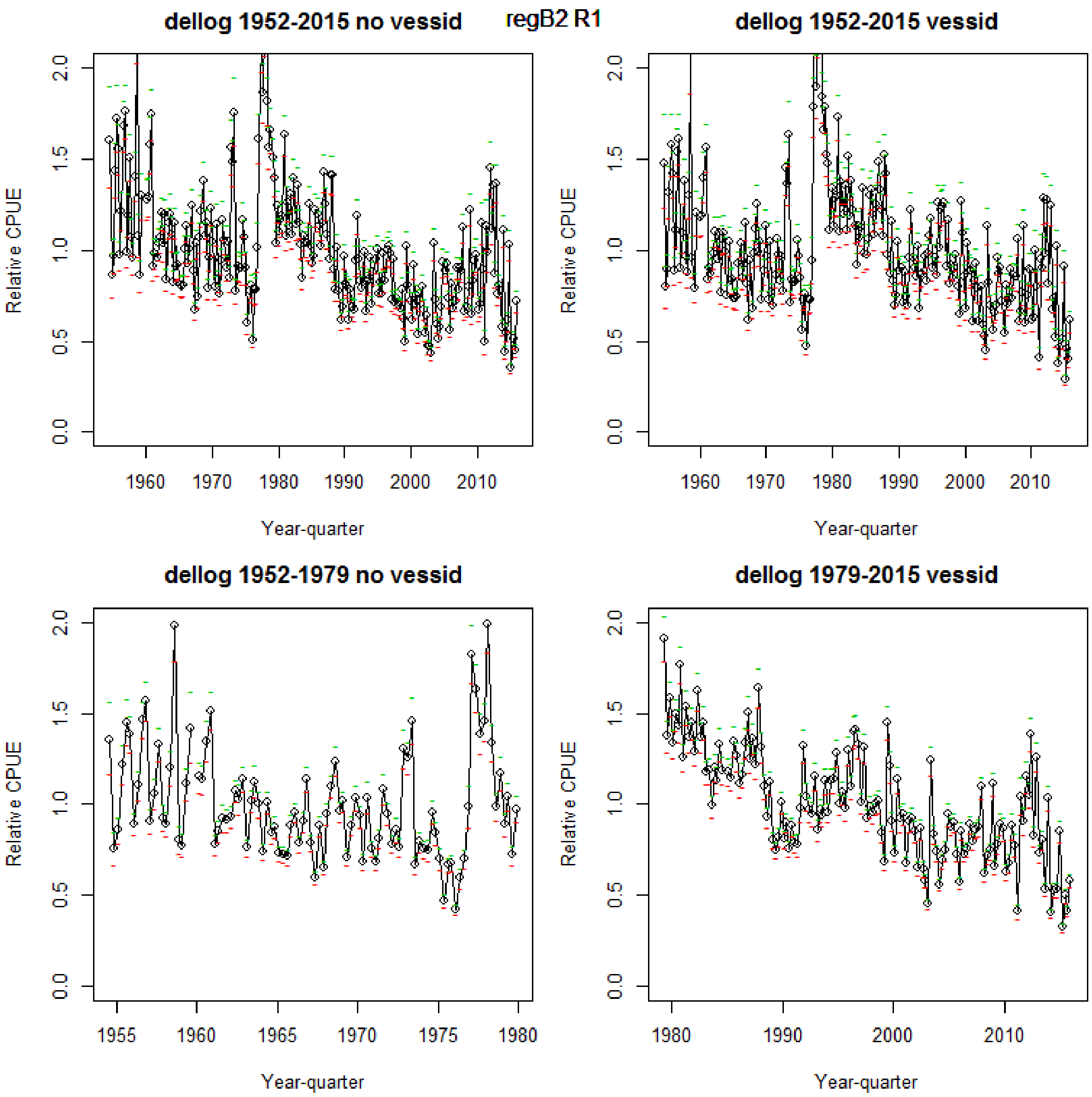


*Figure 42: Estimated CPUE series for bigeye region 1 (western tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

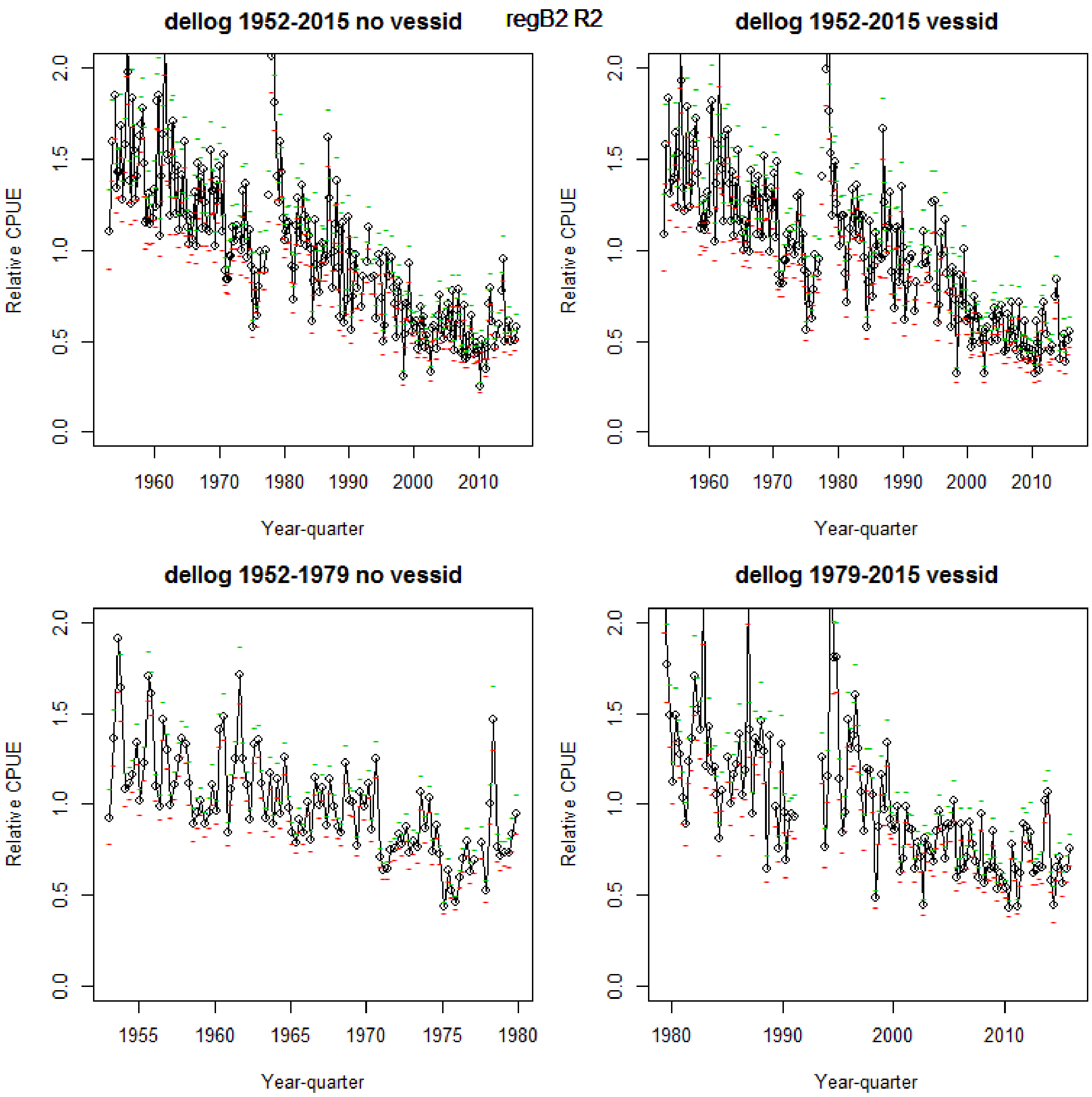


*Figure 43: Estimated CPUE series for bigeye region 2 (eastern tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

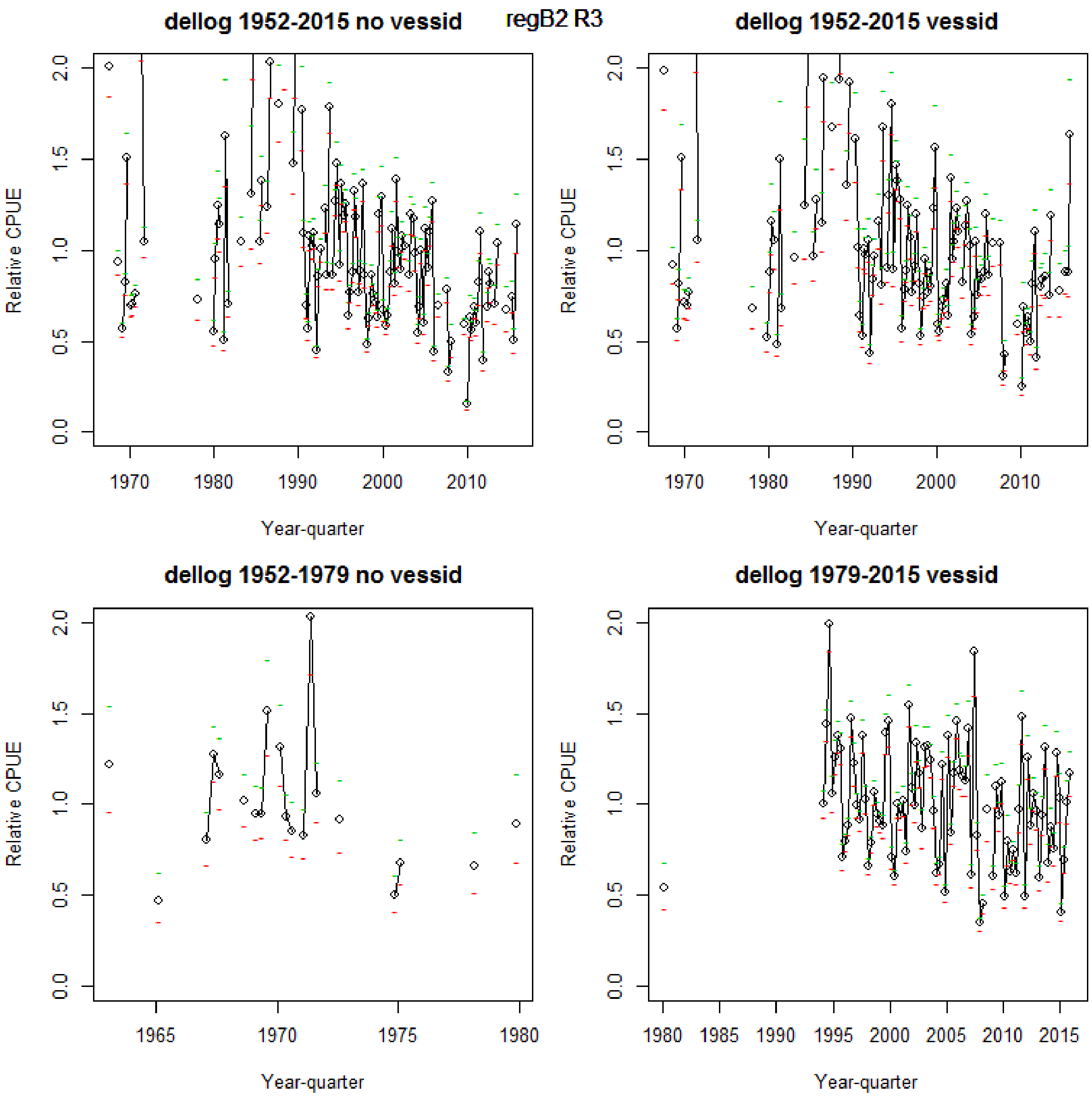


*Figure 44: Estimated CPUE series for bigeye region 3 (western temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

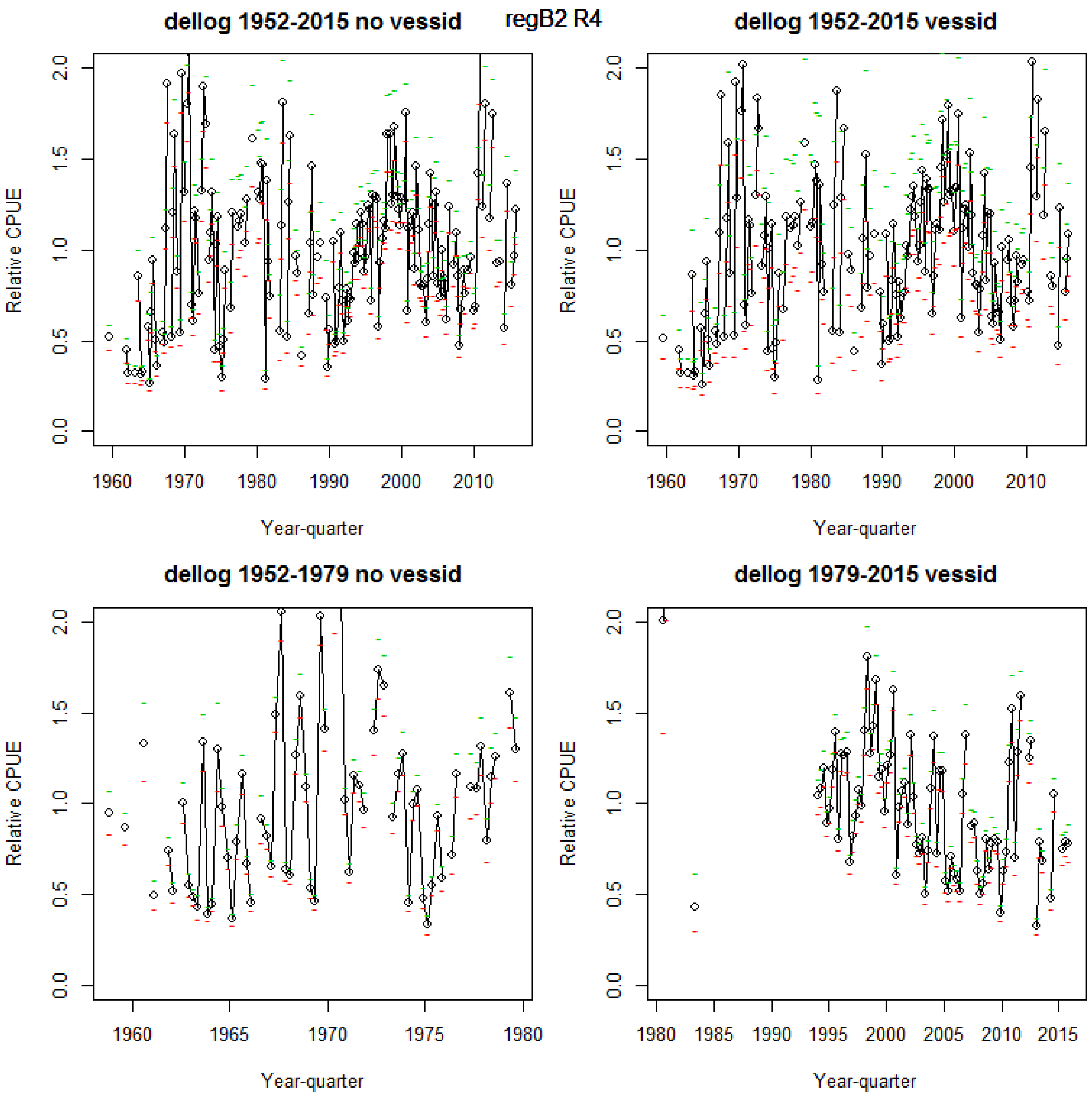


*Figure 45: Estimated CPUE series for bigeye region 4 (eastern temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

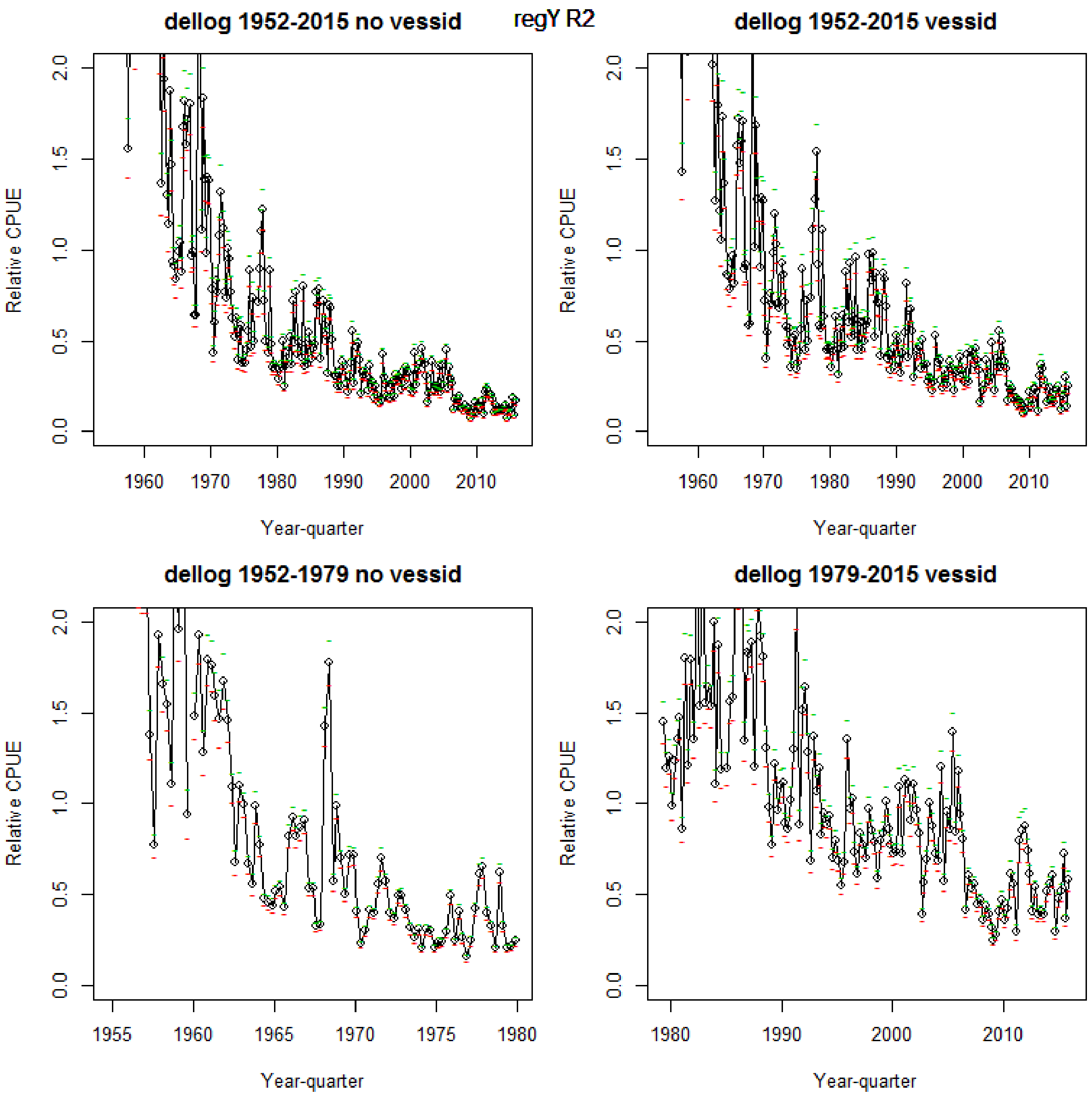


*Figure 46: Estimated CPUE series for yellowfin region 2 (western tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

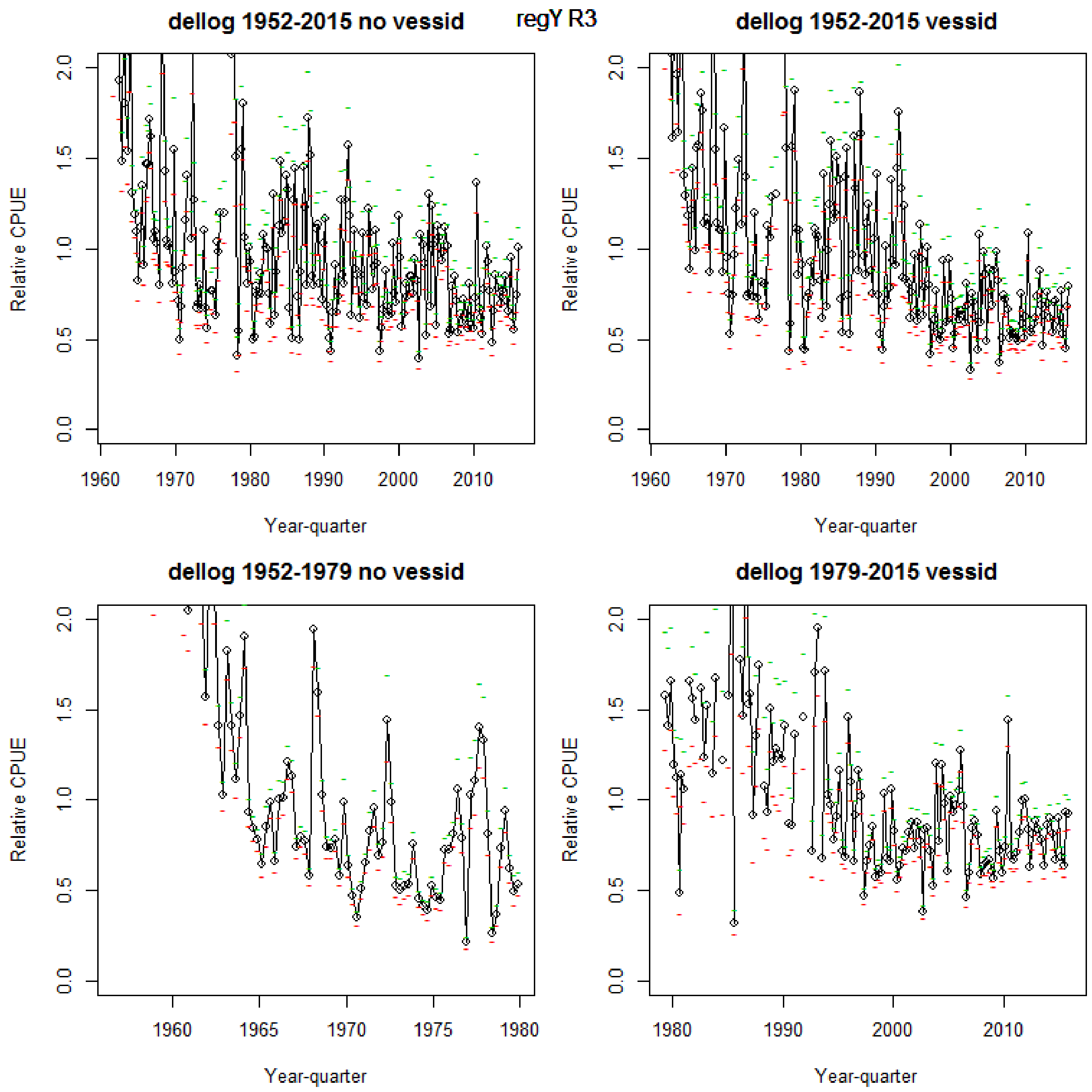


*Figure 47: Estimated CPUE series for yellowfin region 3 (western temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

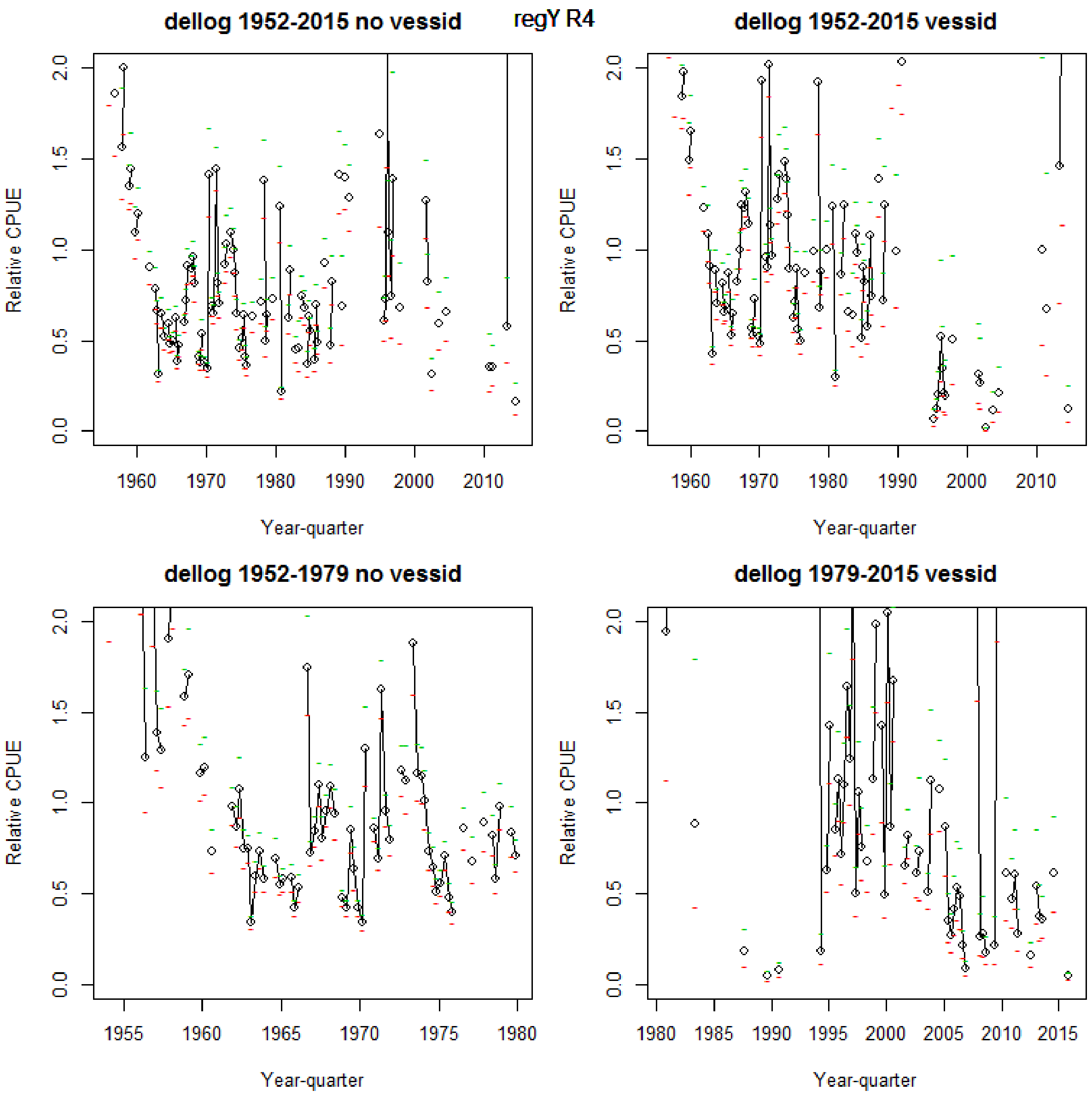


*Figure 48: Estimated CPUE series for yellowfin region 4 (eastern temperate), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

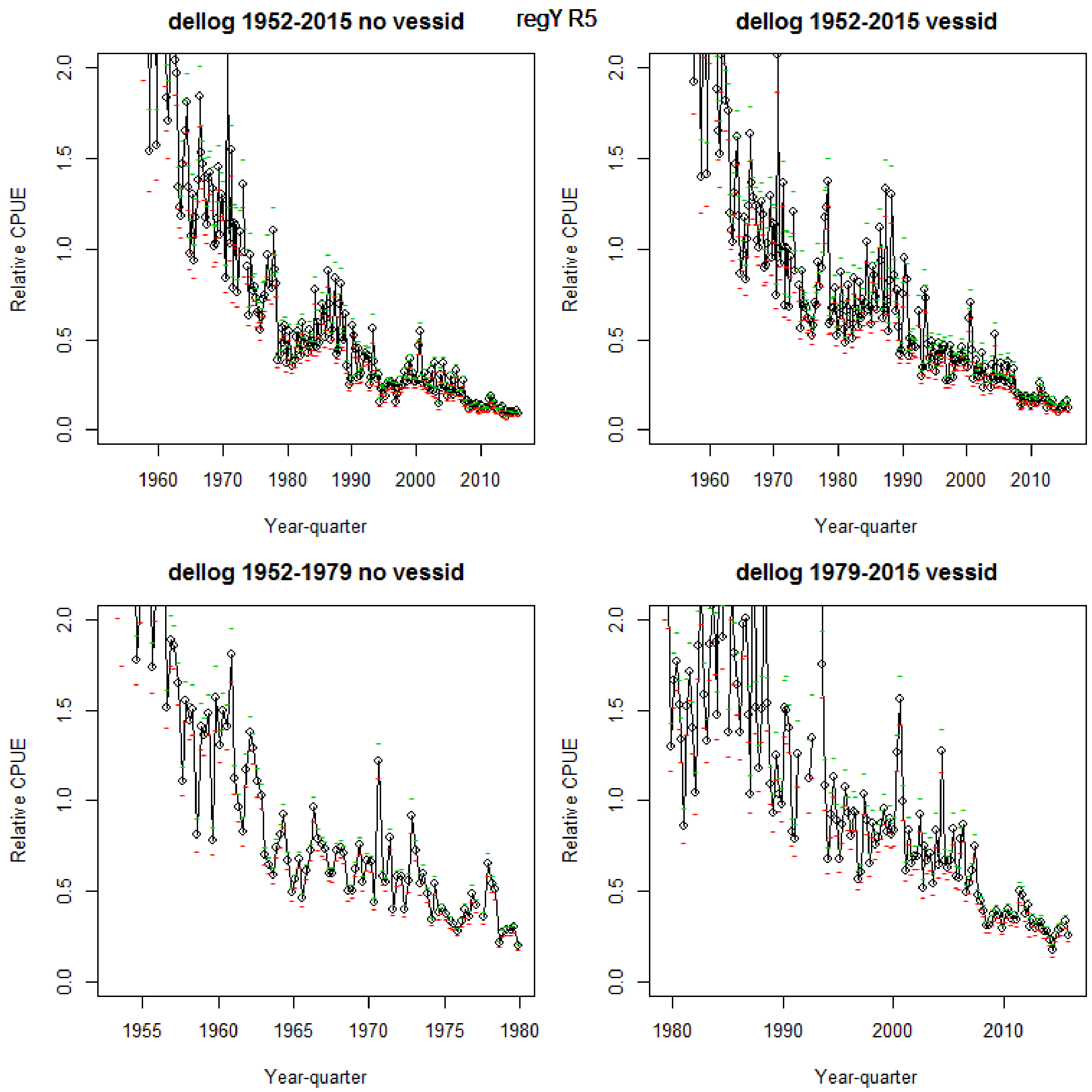


*Figure 49: Estimated CPUE series for yellowfin region 5 (eastern tropical), including time series for all years (top) both with (right) and without (left) vessel effects, and time series for 1952-79 without vessel effects, and 1979-2015 with vessel effects.*

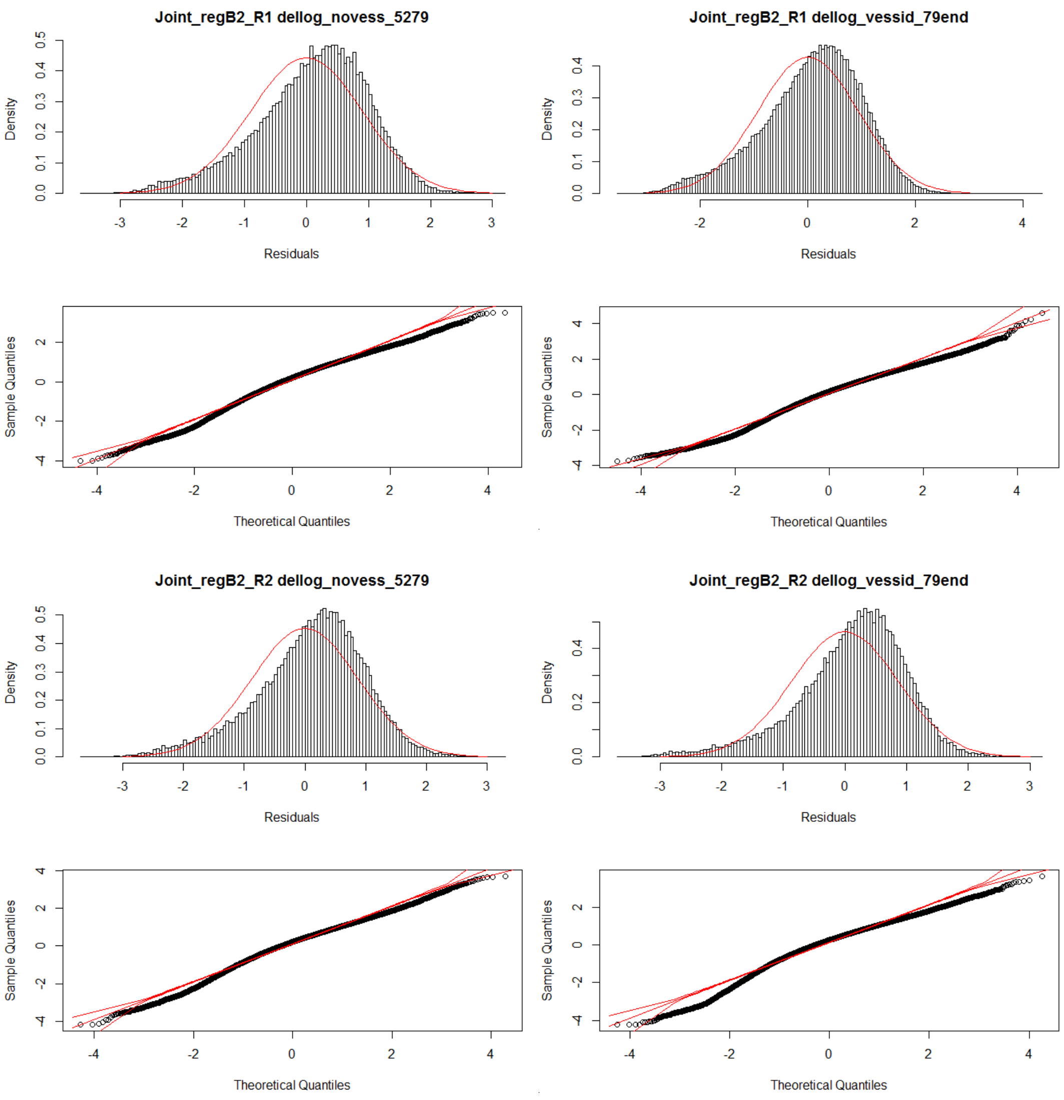


*Figure 50: Diagnostics plots for bigeye lognormal positive models in tropical regions 1 and 2, for 1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right).*

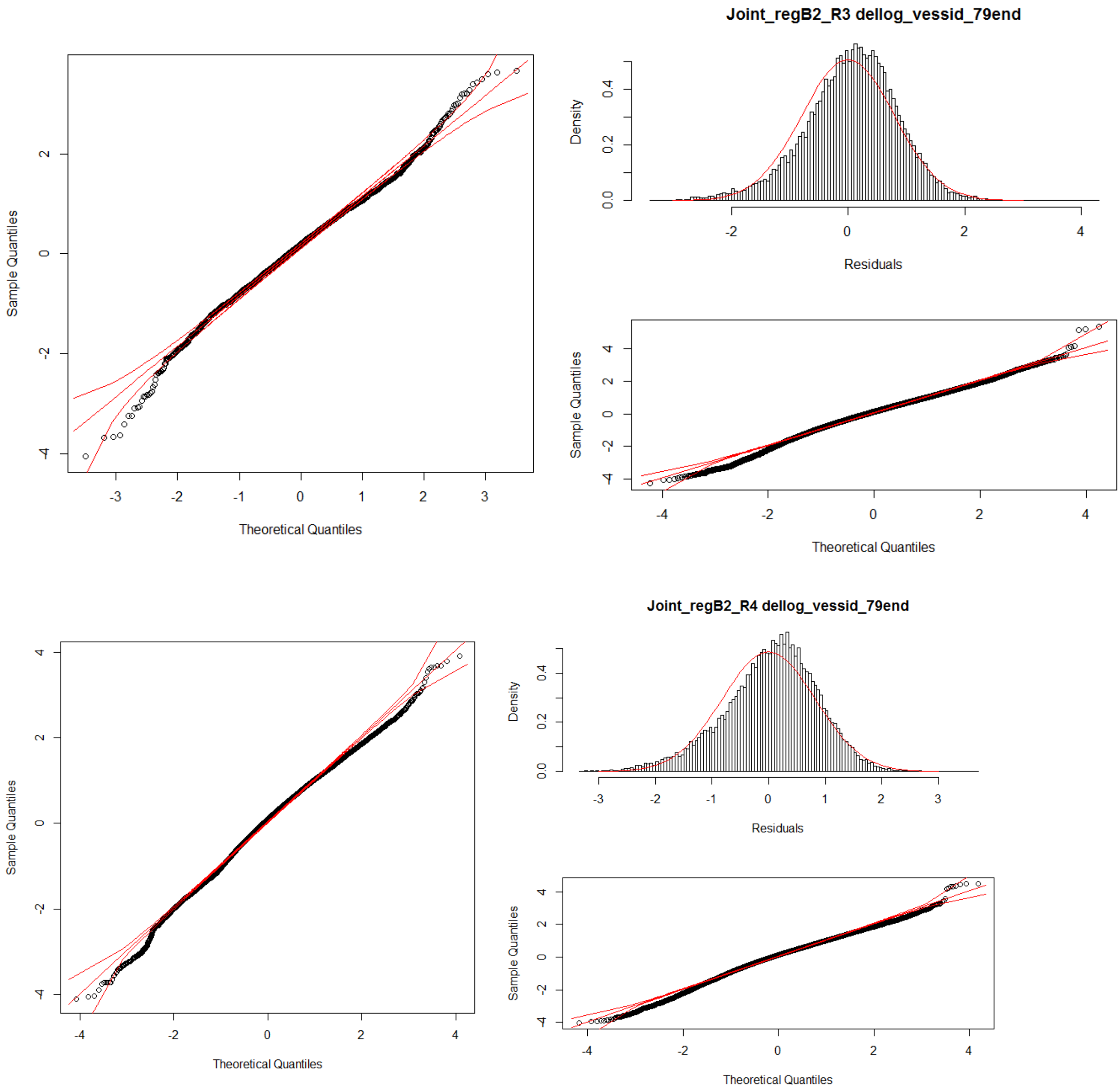


*Figure 51: Diagnostics plots for bigeye lognormal positive models in temperate regions 3 and 4, for1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right).*

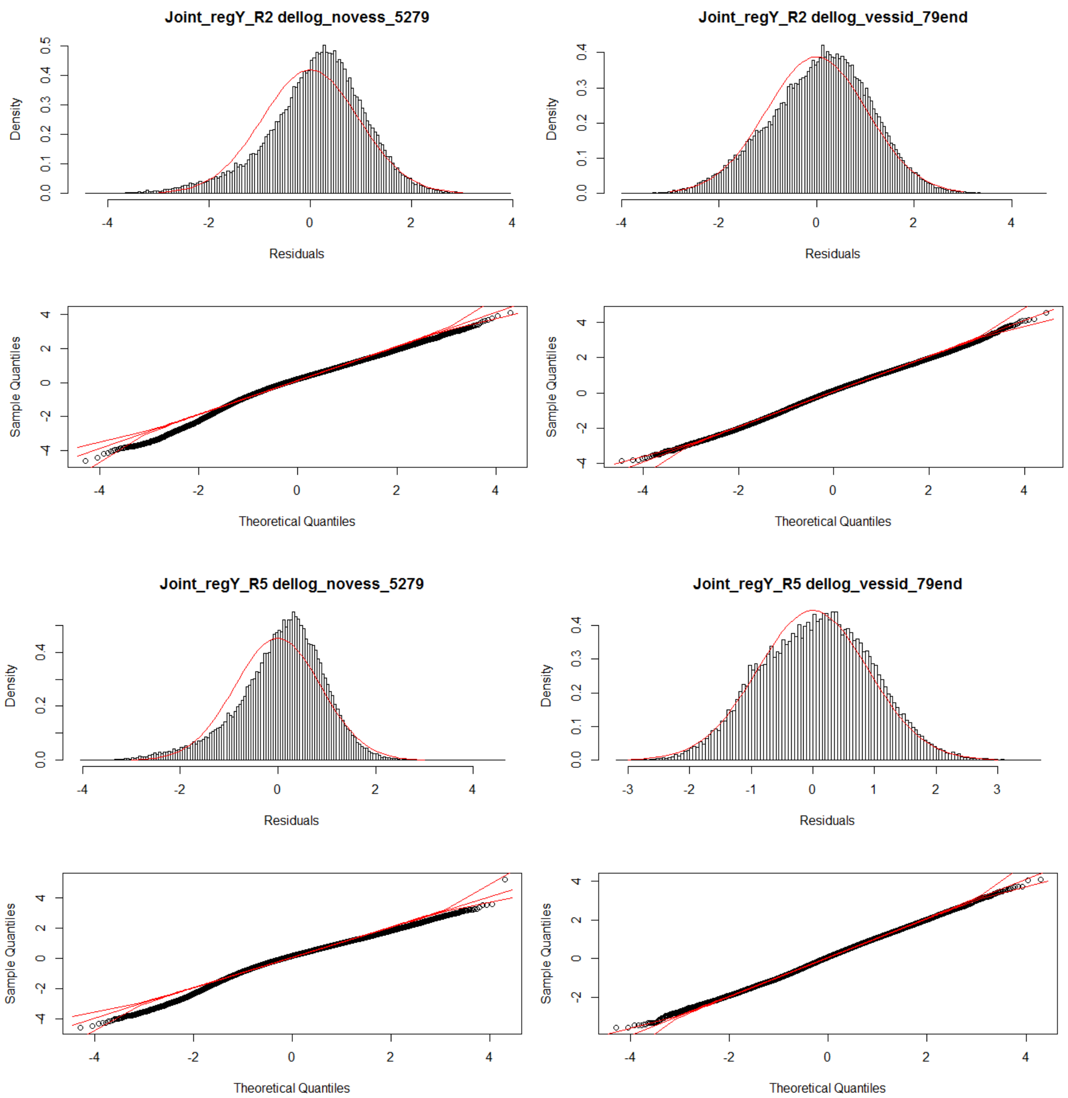


*Figure 52: Diagnostics plots for yellowfin lognormal positive models in tropical regions 2 and 5, for 1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right).*

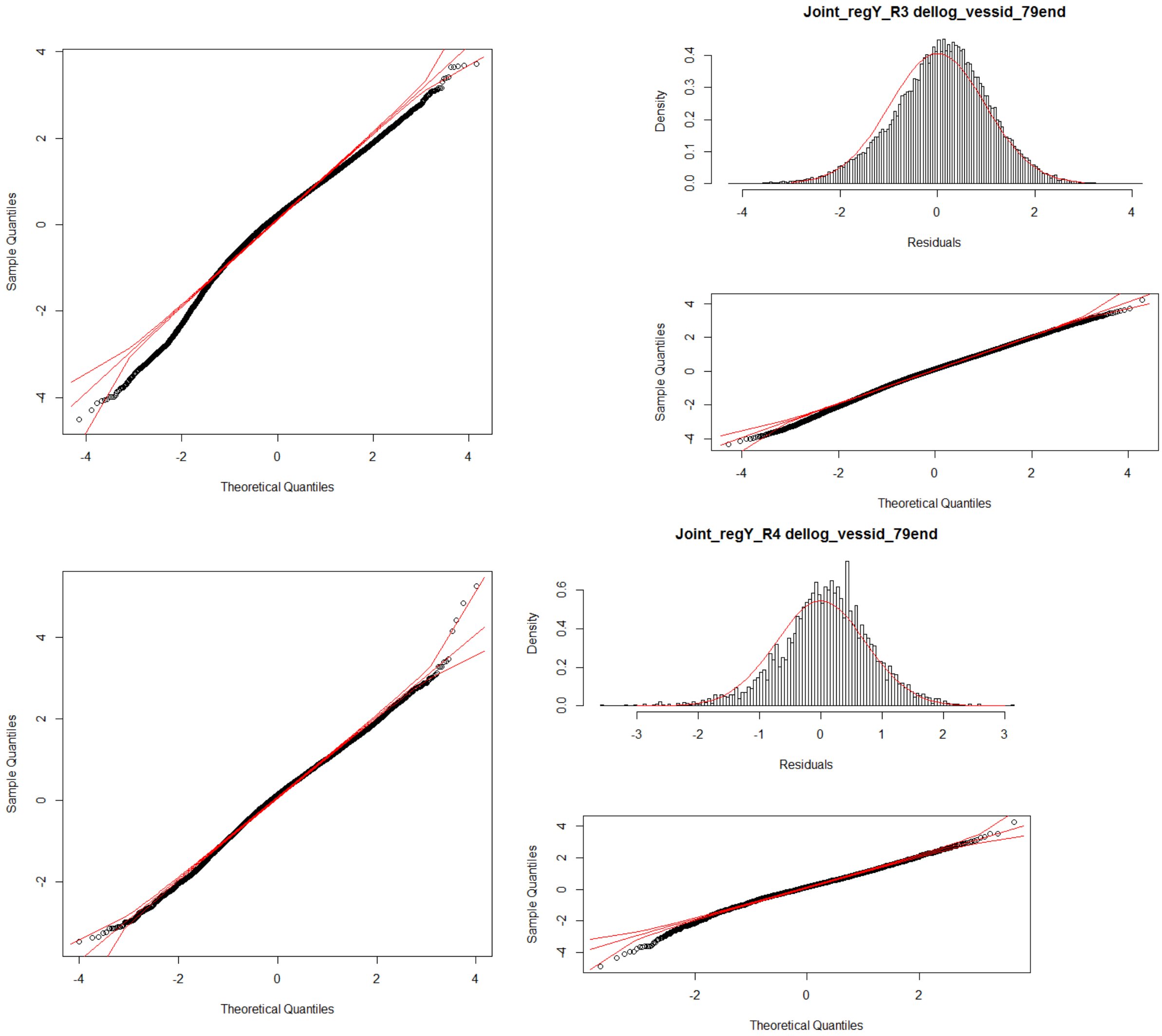


*Figure 53: Diagnostics plots for yellowfin lognormal positive models in temperate regions 3 and 4, for1952-79 without vessel effects (left) and for 1979-2015 with vessel effects (right).*

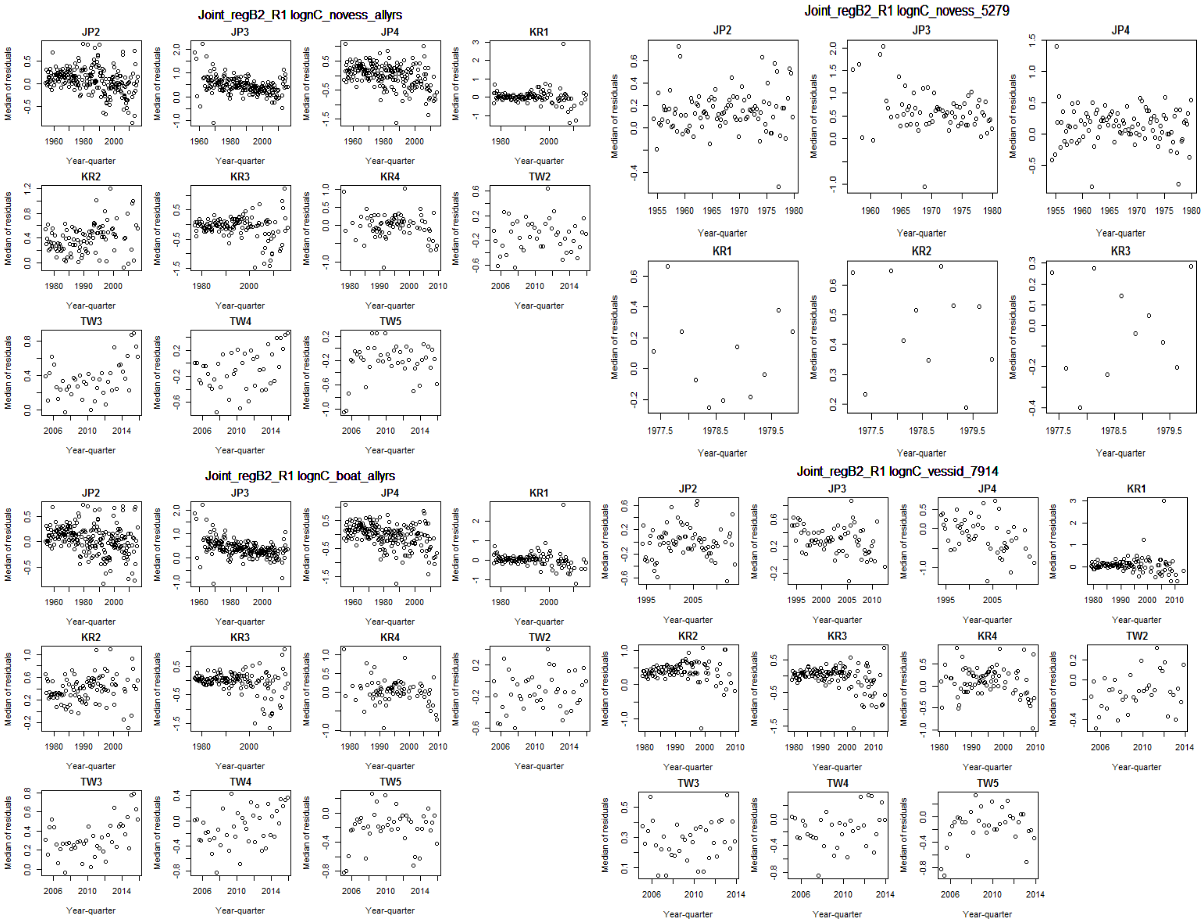


*Figure 54: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 1 (western tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

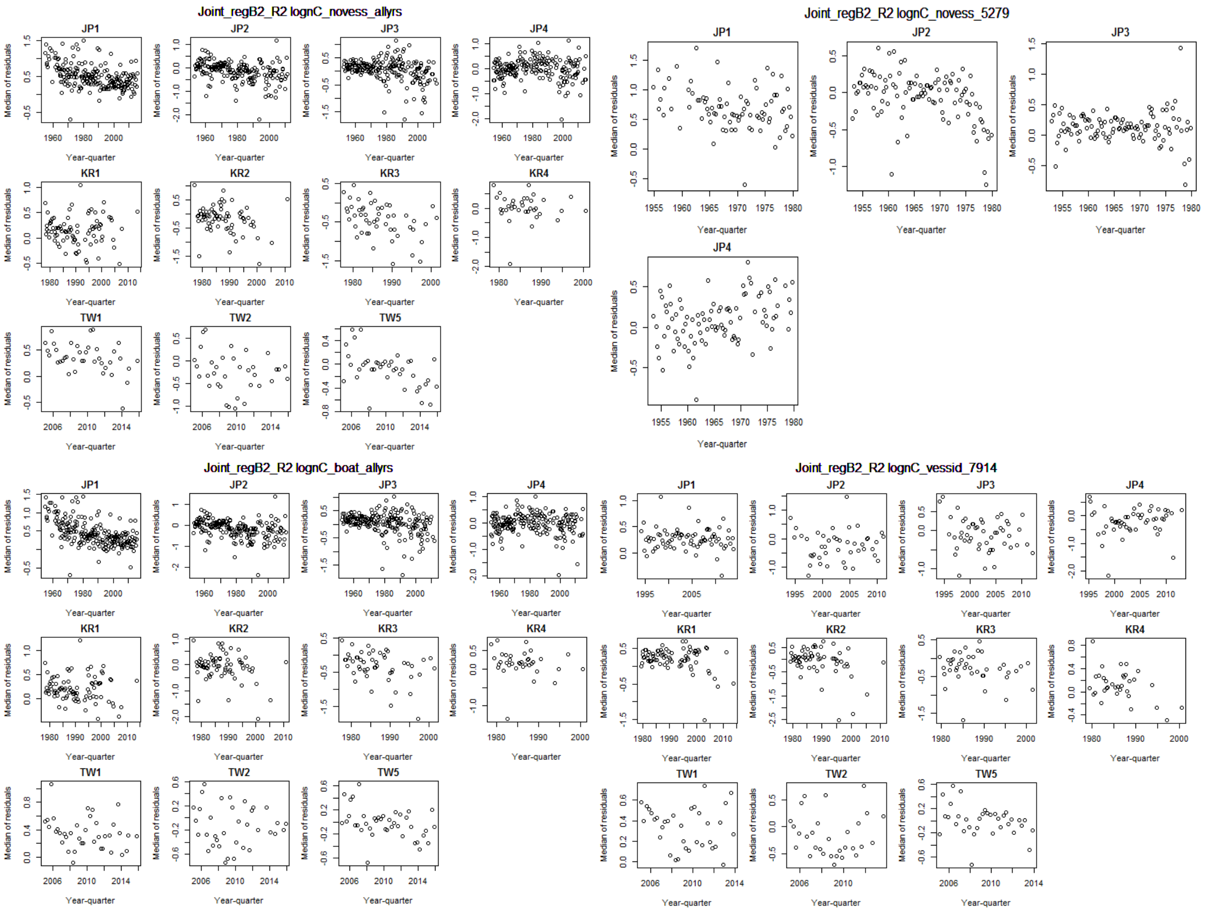


*Figure 55: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 2 (eastern tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

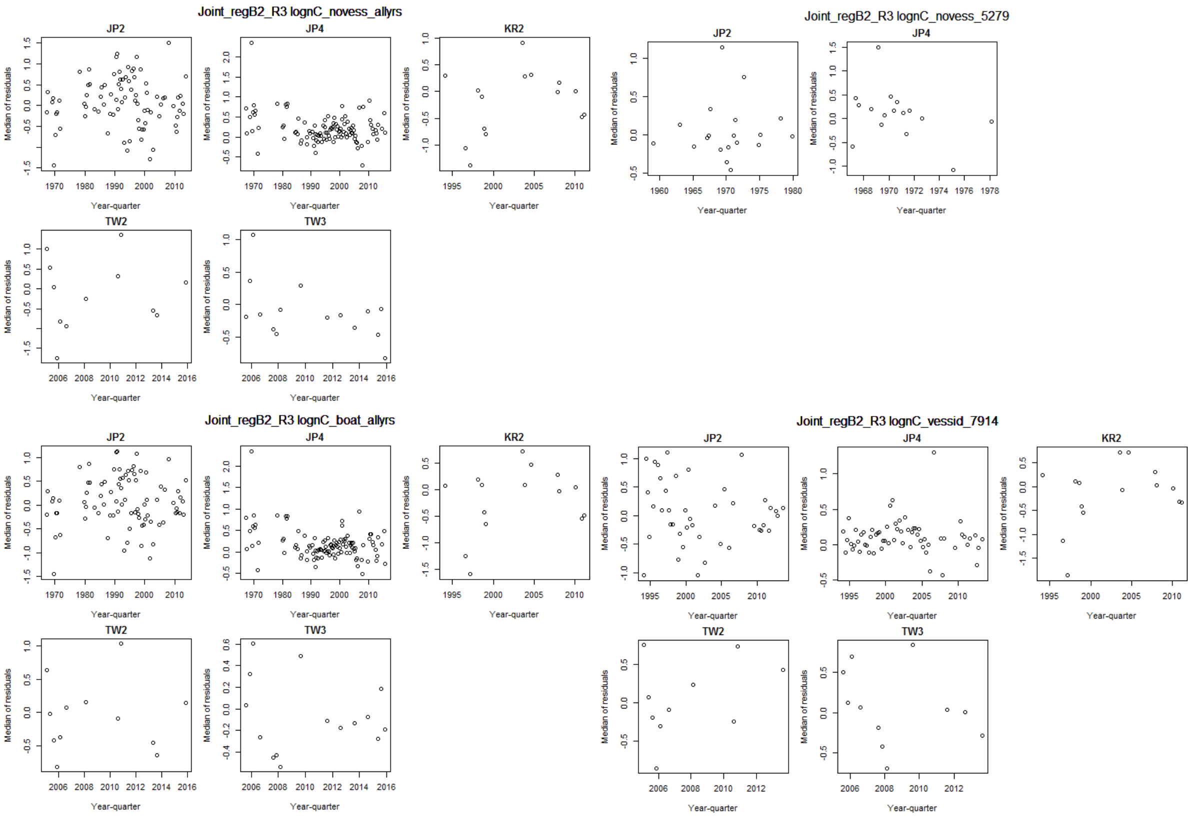


*Figure 56: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 3 (western temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

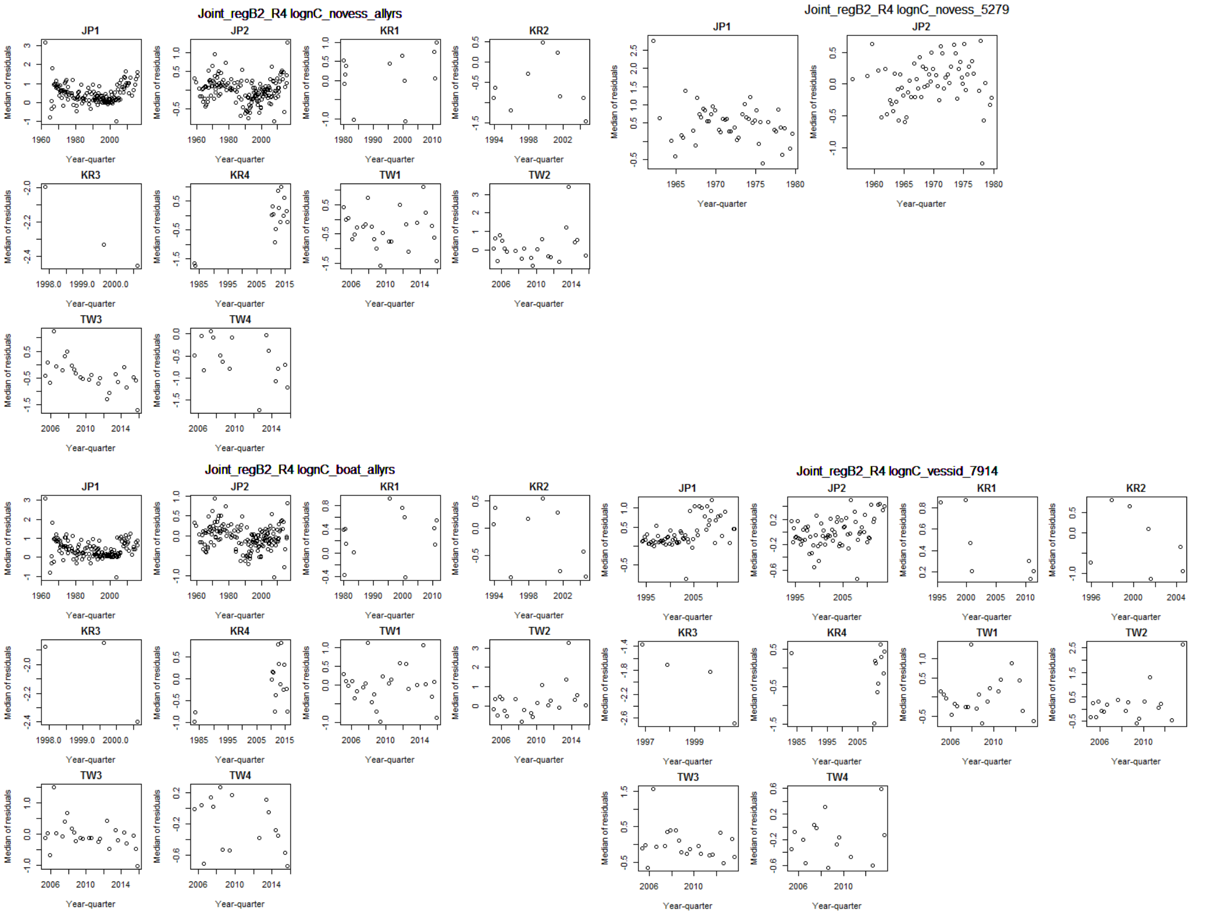


*Figure 57: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for bigeye in region 4 (eastern temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

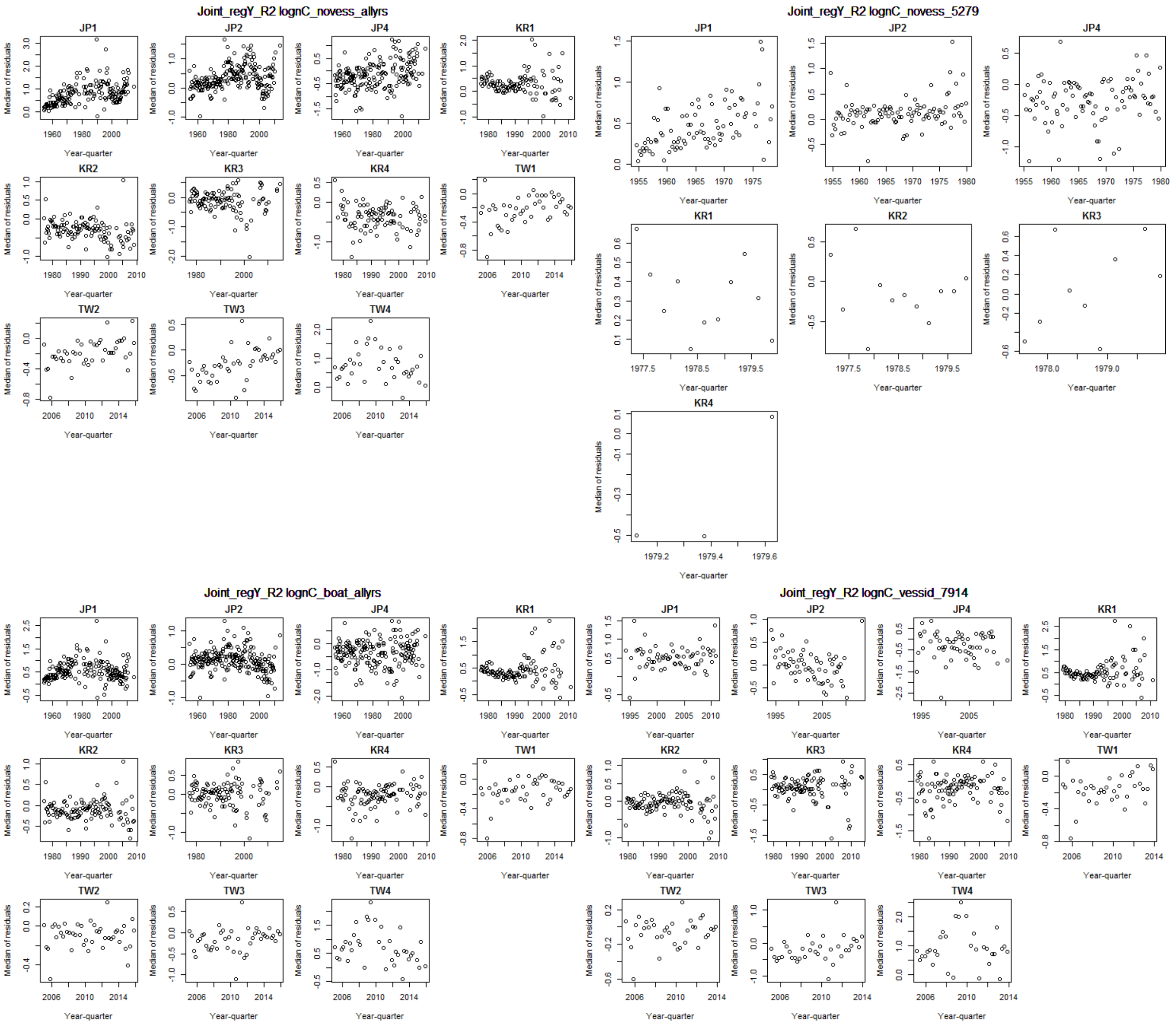


*Figure 58: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 2 (western tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

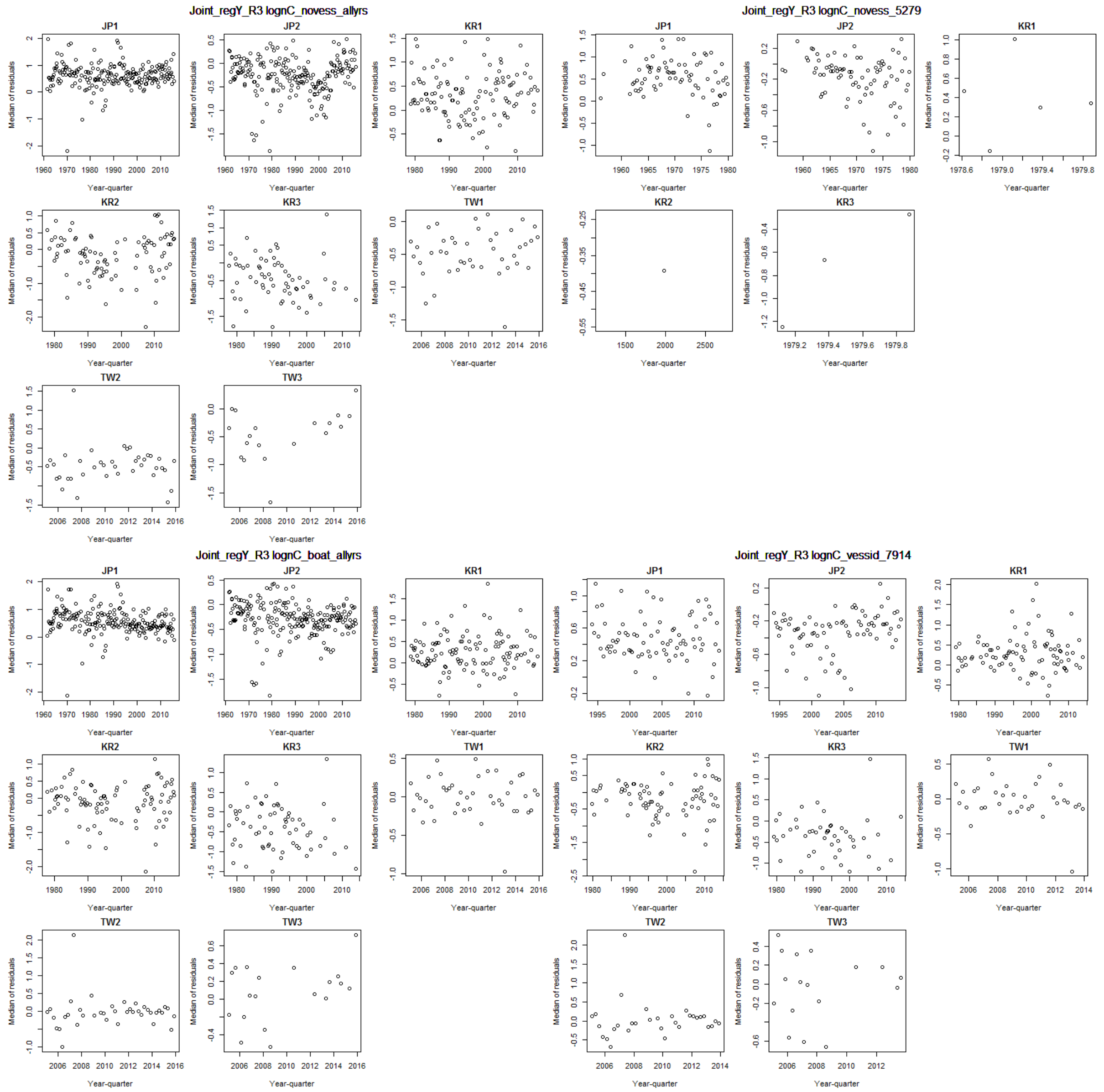

*Figure 59: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 3 (western temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

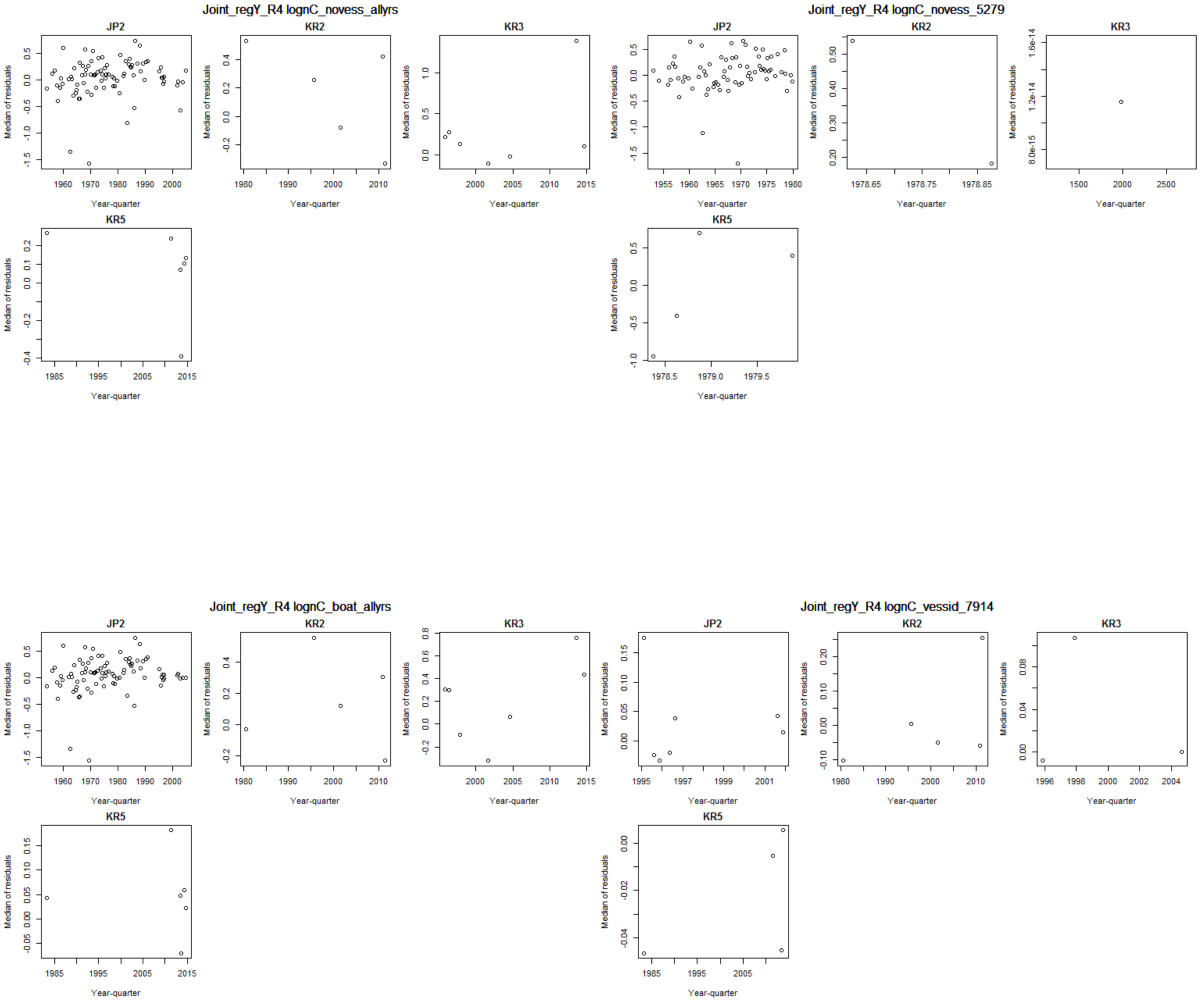


*Figure 60: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 4 (eastern temperate). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

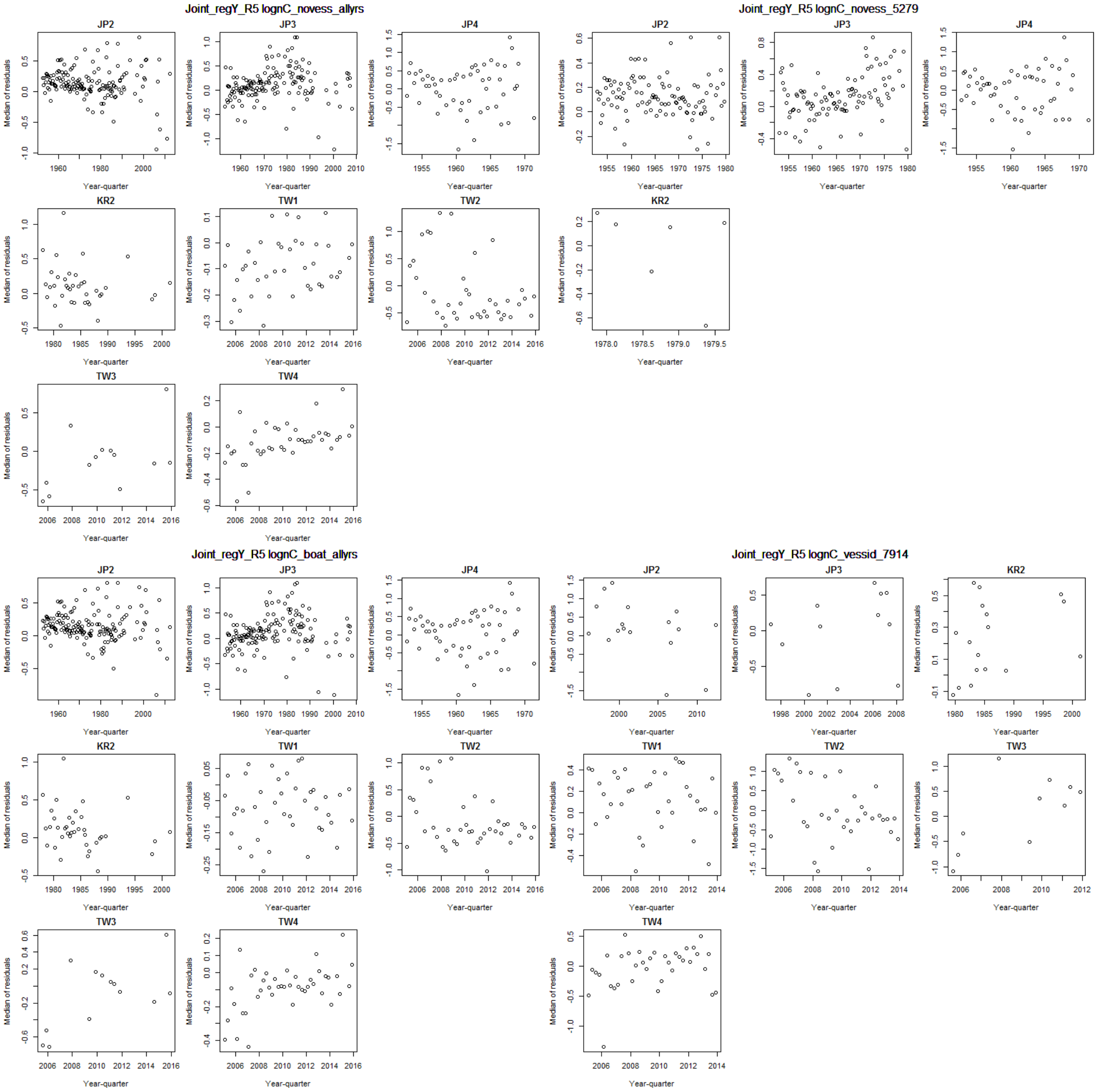


*Figure 61: Median residuals from the lognormal constant model per year-quarter (x-axis), by cluster (subplots), for yellowfin in region 5 (eastern tropics). Residuals are shown for 4 models: 1952-2015 without vessel effects (top left), 1952-1979 without vessel effects (top right), 1952-2015 with vessel effects (bottom left) and 1979-2015 with vessel effects (bottom right).*

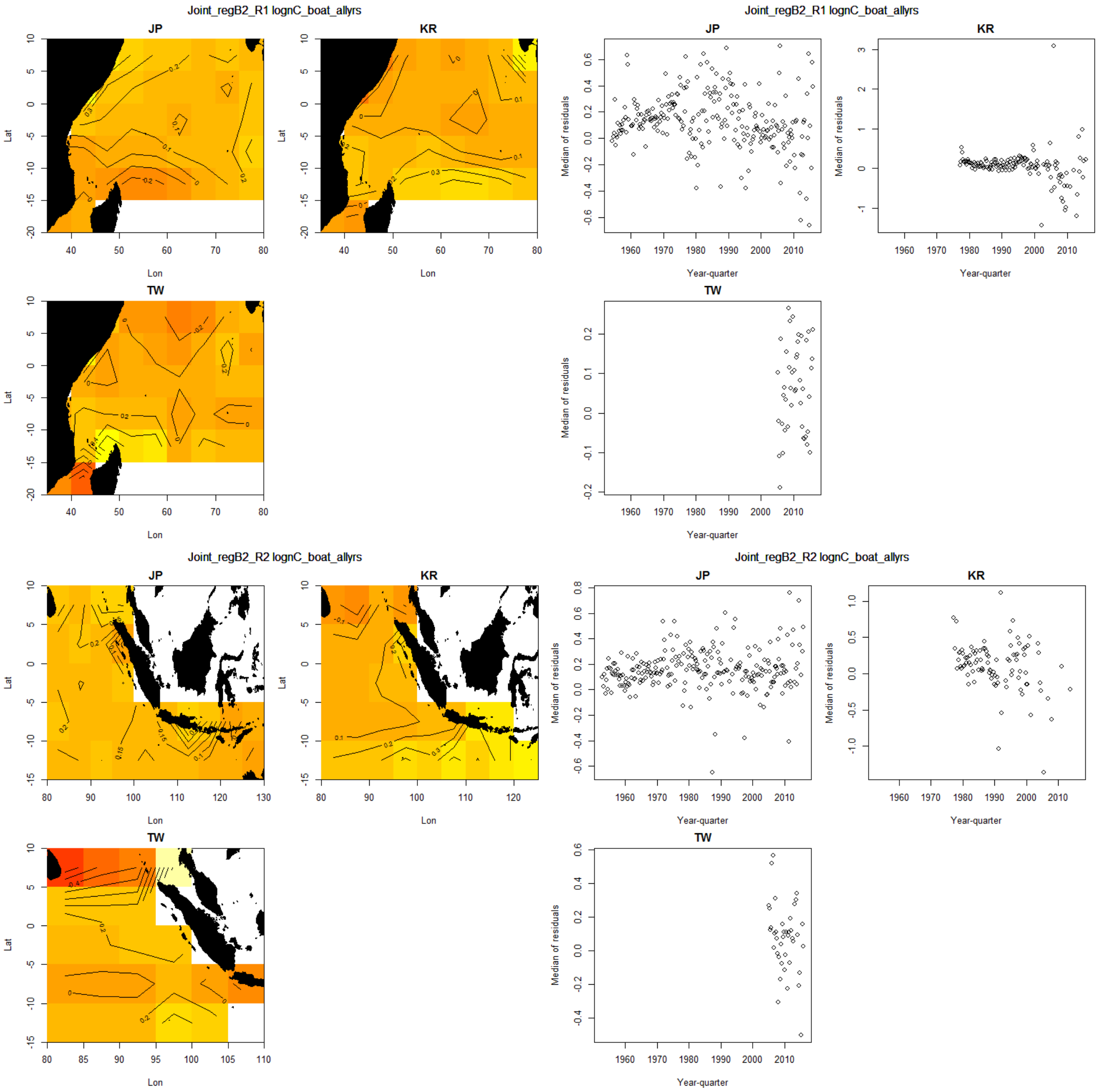


*Figure 62: Bigeye residuals for tropical regions 1 (above) and 2 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right).*

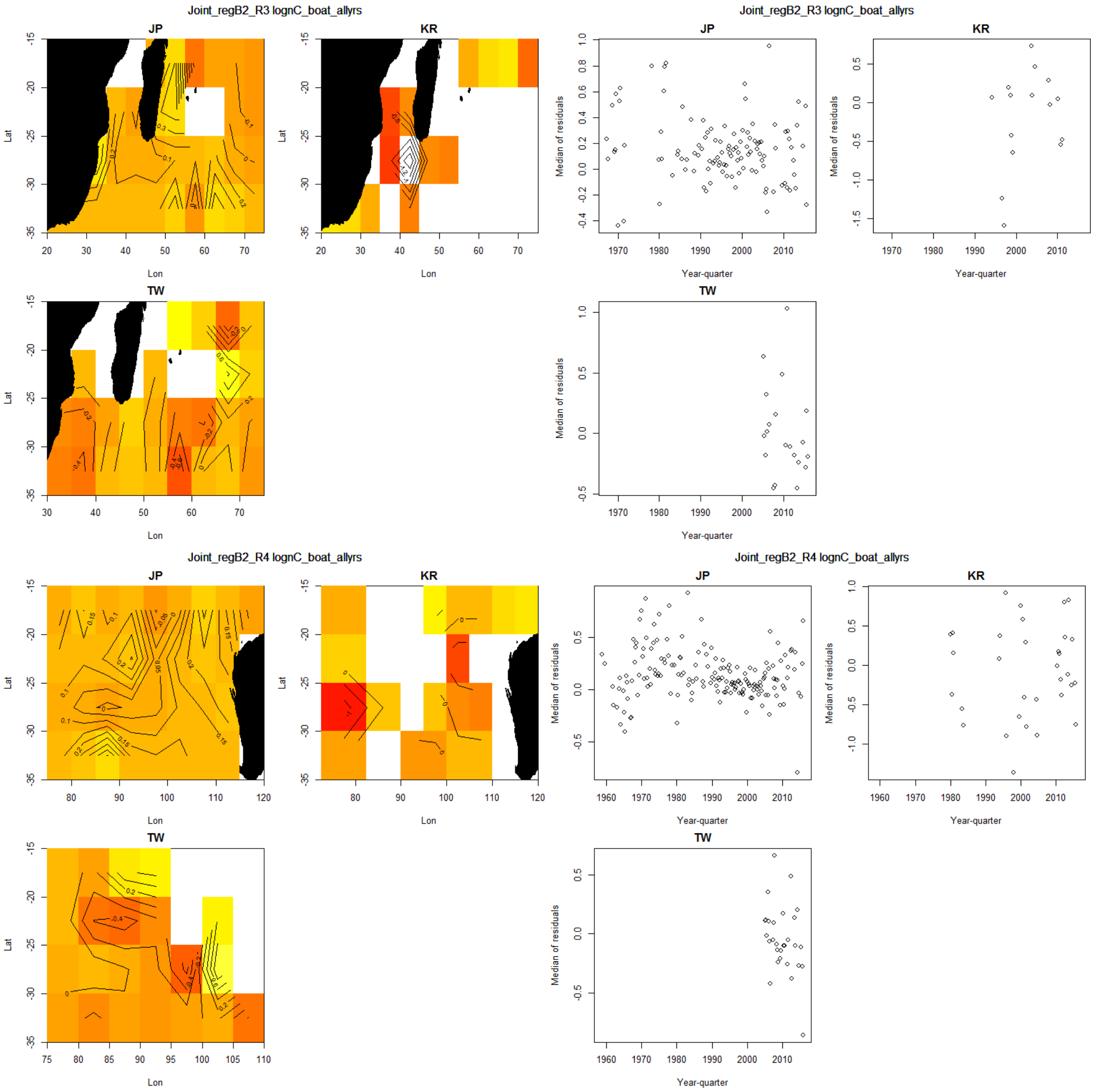


*Figure 63: Bigeye residuals for temperate regions 3 (above) and 4 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right).*

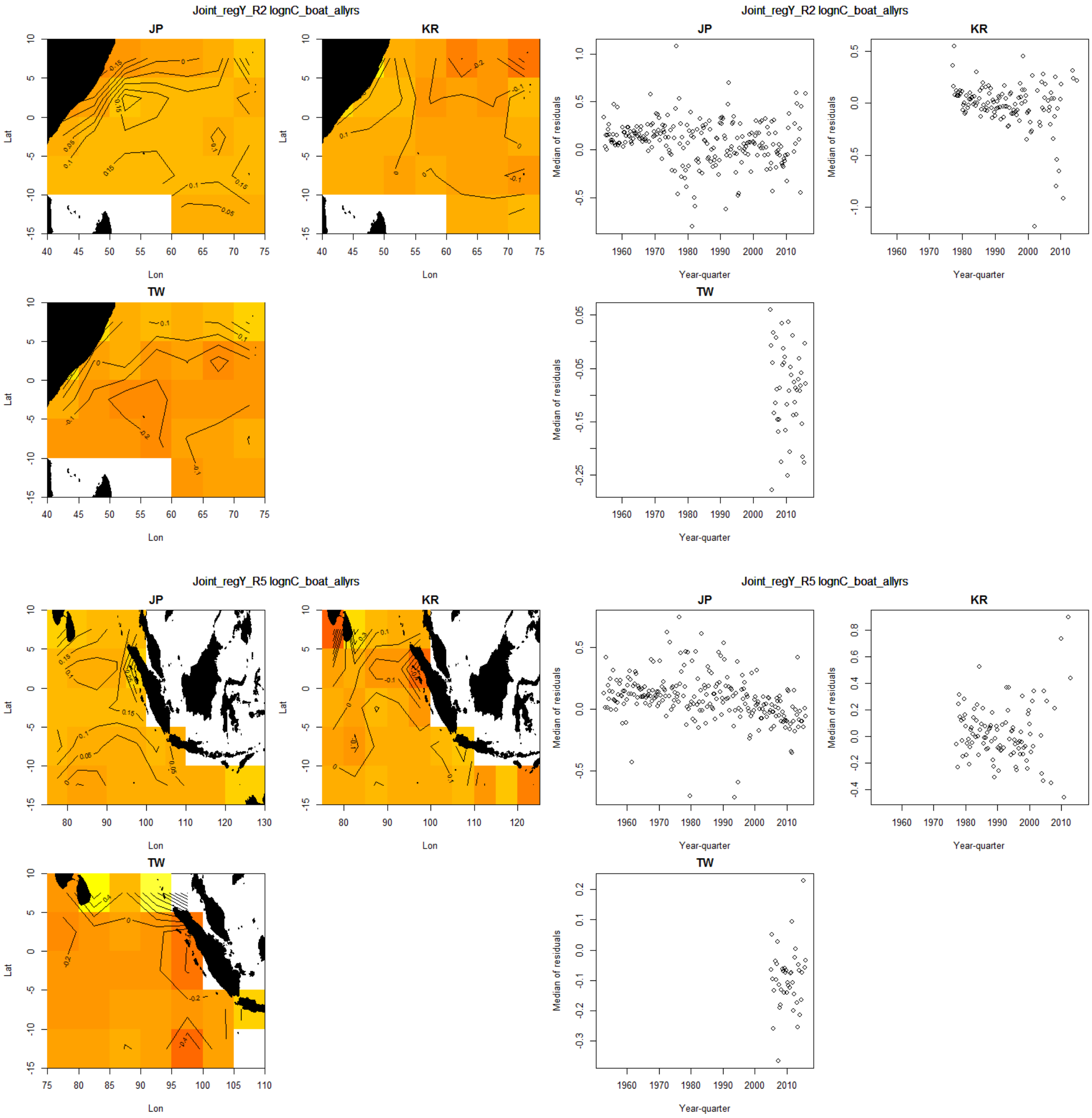


*Figure 64: Yellowfin residuals for tropical regions 2 (above) and 5 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right).*

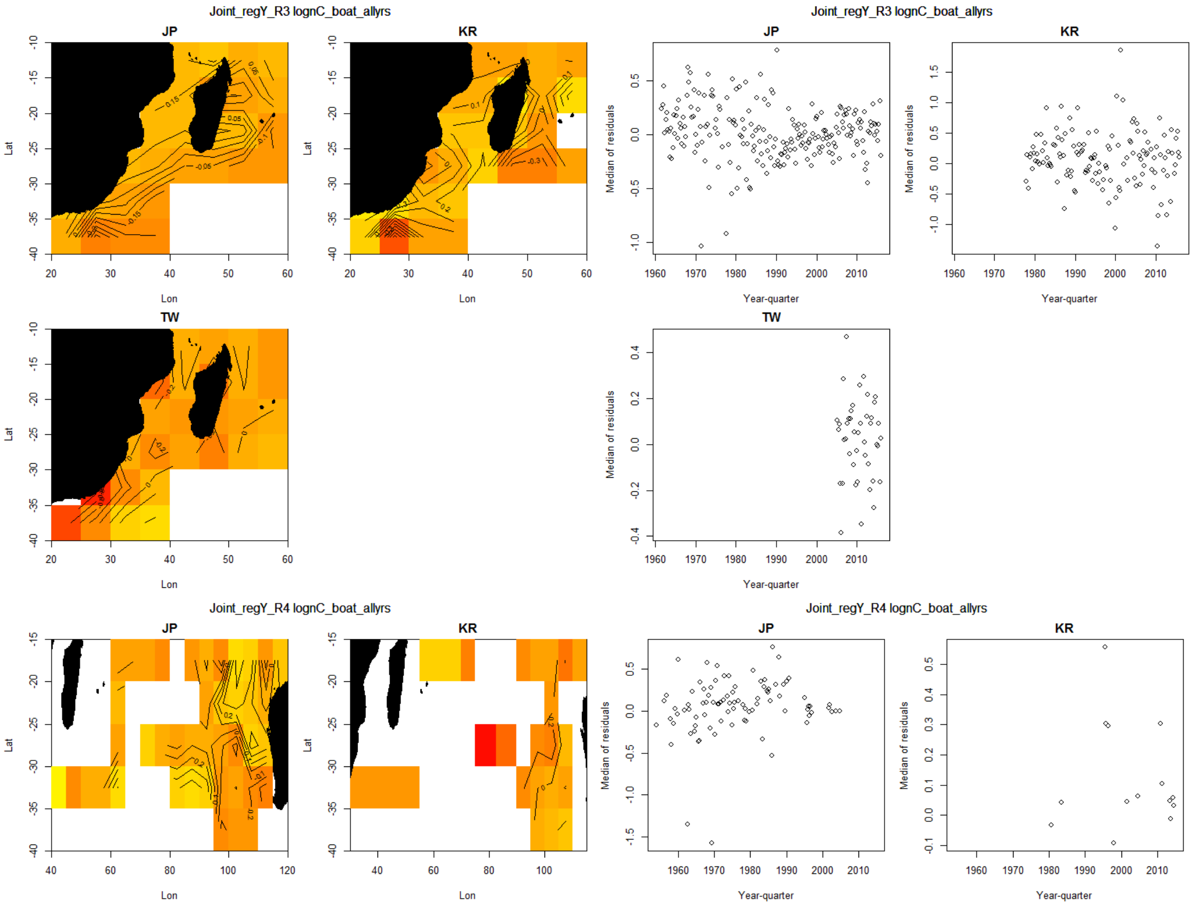


*Figure 65: Yellowfin residuals for temperate regions 3 (above) and 4 (below), by flag. Median residuals are mapped by 5 cell (left) and plotted by year-quarter (right).*

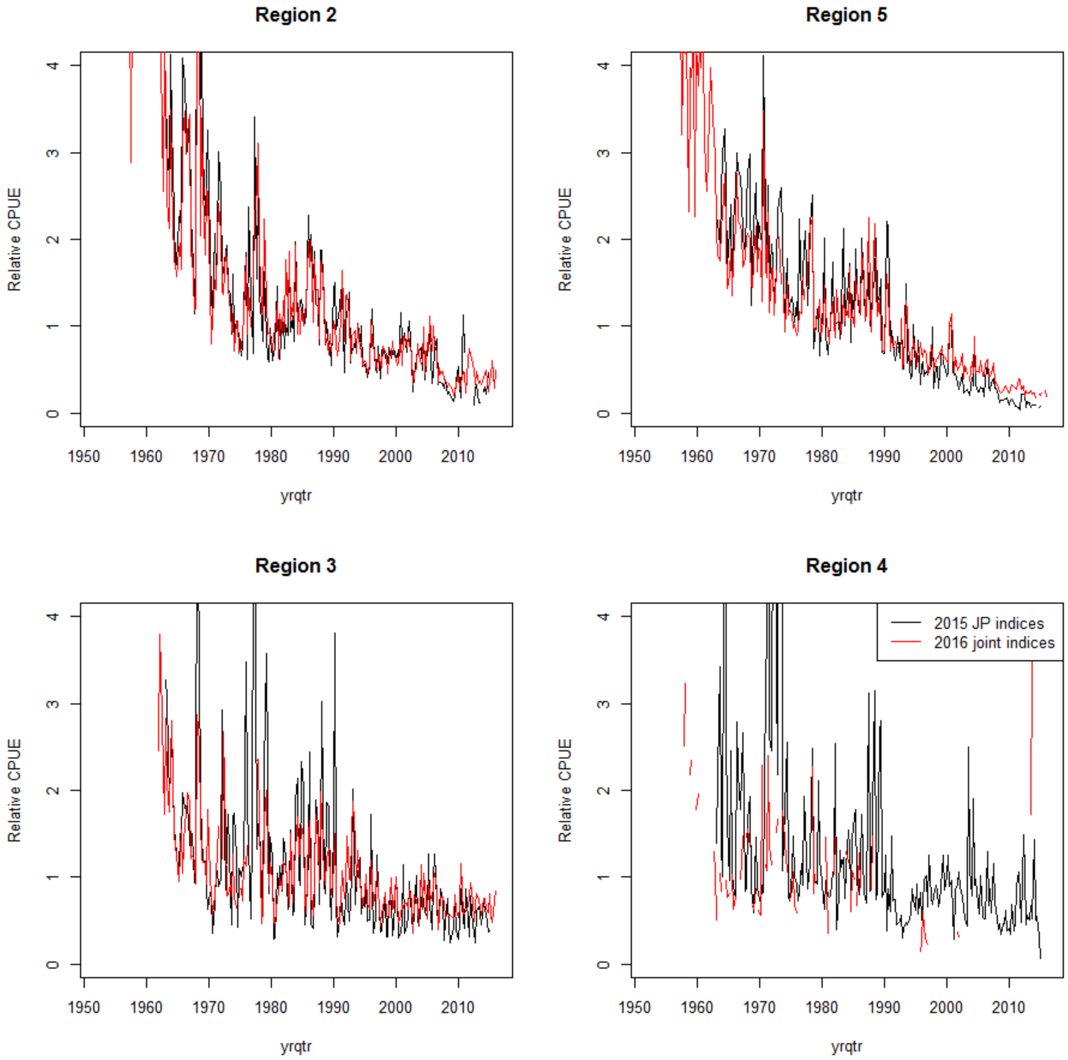


*Figure 66: Comparison of the joint indices described in this paper (red) with the Japanese indices developed in 2015 and used in the 2015 yellowfin stock assessment.*

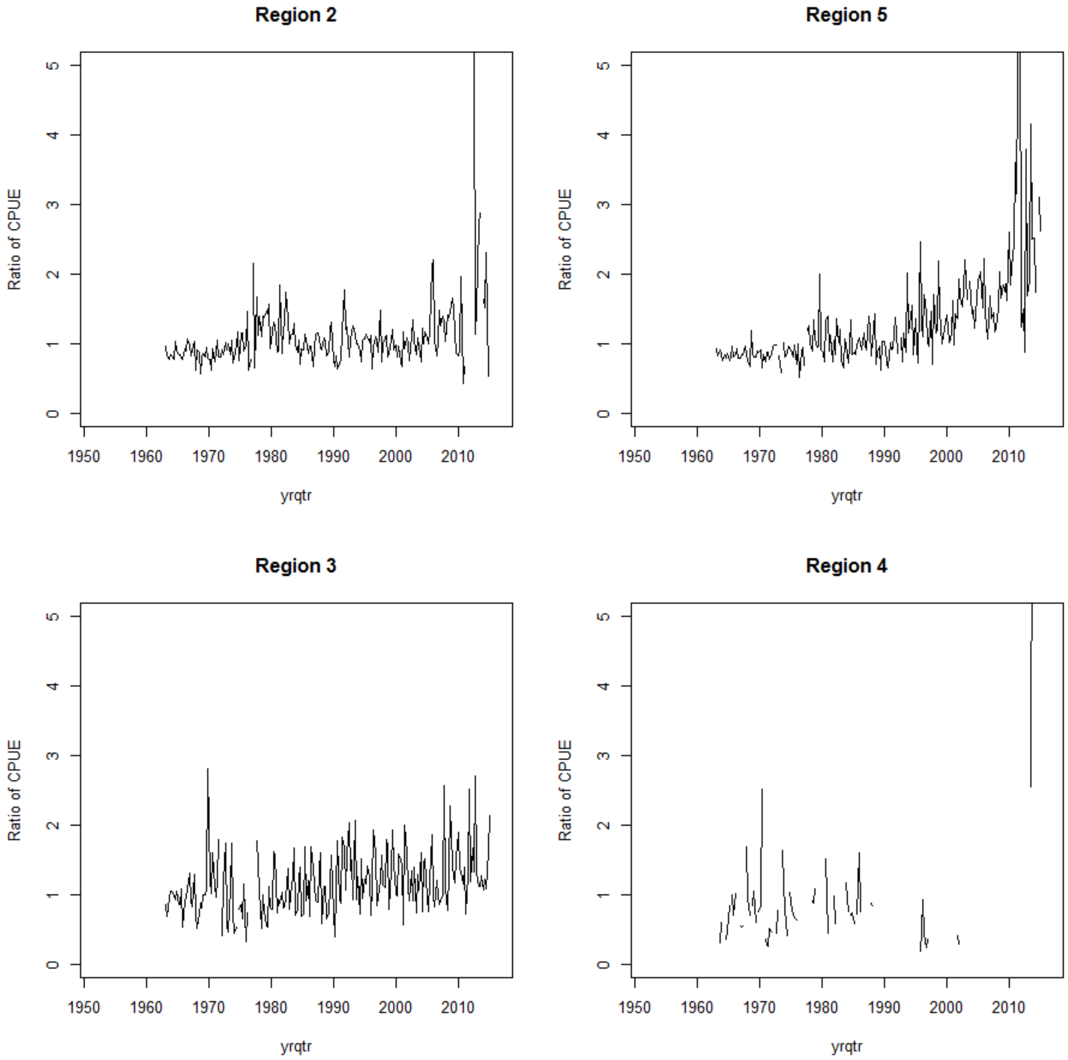


*Figure 67: Ratio of the joint indices described in this paper divided by the Japanese indices developed in 2015 and used in the 2015 yellowfin stock assessment.*

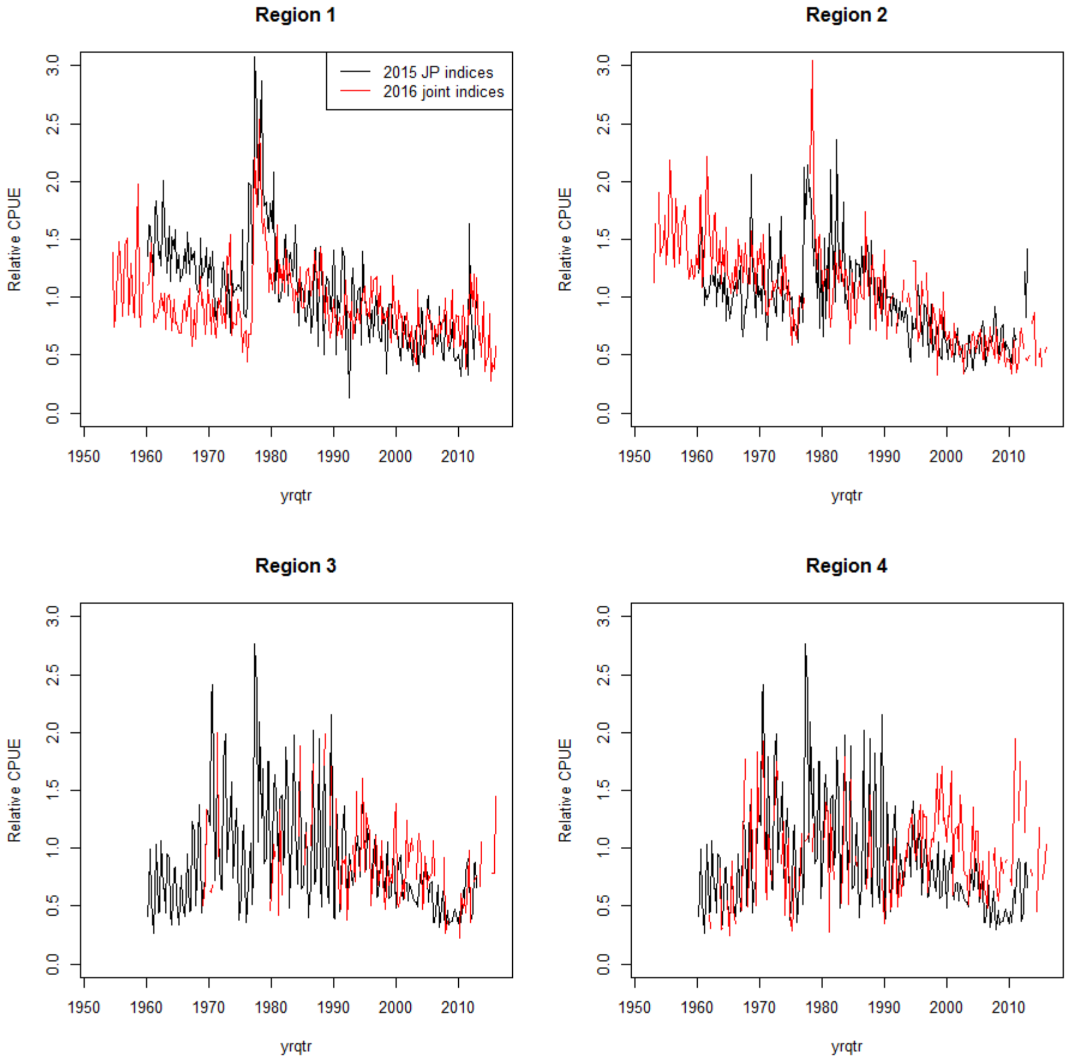


*Figure 68: Comparison of the joint indices described in this paper (red) with the Japanese indices developed in 2013 and used in the 2013 bigeye stock assessment (black).*

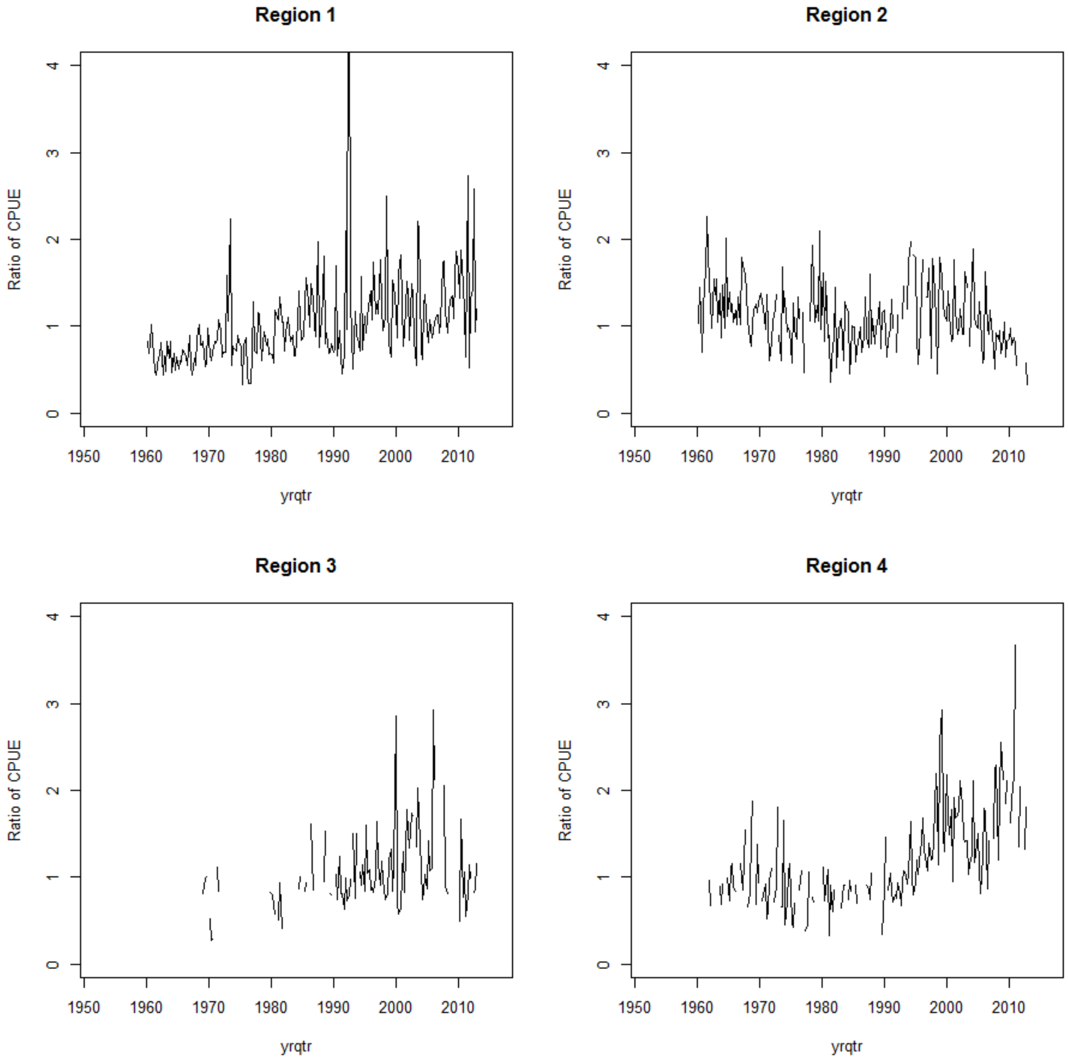


*Figure 69: Ratio of the joint indices described in this paper divided by the Japanese indices developed in 2013 and used in the 2013 bigeye stock assessment.*