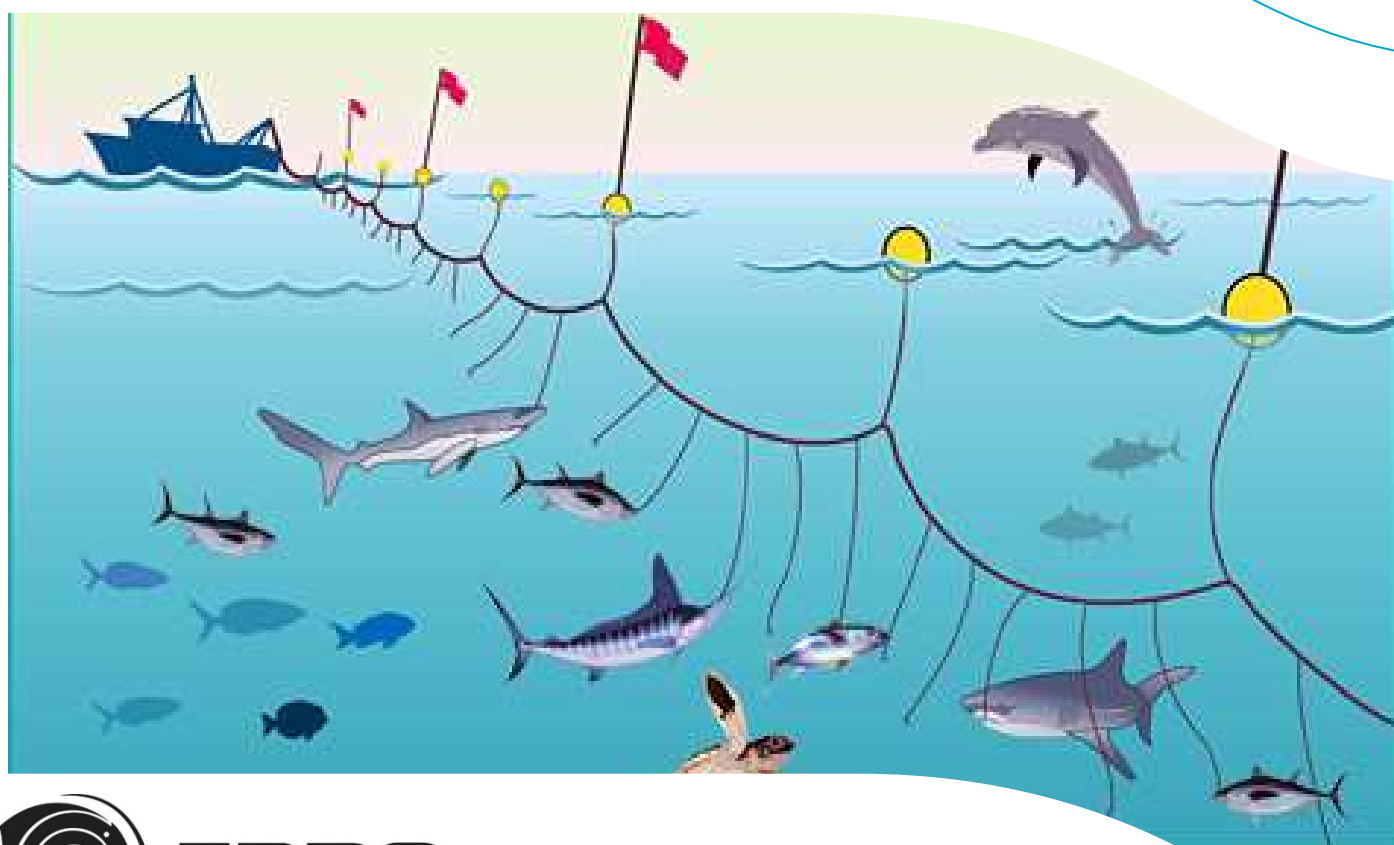


Developing innovative approaches to improve CPUE standardisation for Australia's multispecies pelagic longline fisheries

Robert Campbell, Shijie Zhou, Simon Hoyle, Rich Hillary, Malcolm Haddon and Steve Auld

May 2017

FRDC Project No. 2014-021



FRDC

FISHERIES RESEARCH &
DEVELOPMENT CORPORATION

Developing innovative approaches to improve CPUE standardisation for Australia's multispecies pelagic longline fisheries

Robert Campbell¹
Shijie Zhou²
Simon Hoyle³
Rich Hillary⁴
Malcolm Haddon⁴
Steve Auld⁵

1. CSIRO Oceans and Atmosphere, Private Bag No. 1, Aspendale, VIC 3195, Australia
2. CSIRO Oceans and Atmosphere, GPO Box 2583, Brisbane, QLD 4001, Australia
3. Hoyle Consulting, 14 Champion Terrace, Nelson, 7011, New Zealand
4. CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, TAS 7001, Australia
5. Department of Agriculture and Water Resources, GPO Box 858, Canberra ACT, Australia.

May 2017

FRDC Project No 2014-021



FRDC

FISHERIES RESEARCH &
DEVELOPMENT CORPORATION

© 2017 Fisheries Research and Development Corporation.

All rights reserved.

National Library of Australia Cataloguing-in-Publication entry:

Creator: Campbell, Robert Alan, 1957- author.
Title: Developing innovative approaches to improve CPUE standardisation for Australia's multispecies pelagic longline fisheries
ISBN: 9781486308521 (ebook)
Subjects: Fisheries--Australia.
Fisheries--Equipment and supplies--Australia.
Fisheries--Catch effort--Australia.
Fishery resources--Management--Australia.
Fishery management--Australia.
Other Creators Zhou, Shijie, author.
/Contributors: Hoyle, Simon D., author
. Hillary, Rich, author.
Haddon, Malcolm, author.
Auld, Steve, author.
Issuing body: Fisheries Research and Development Corporation (Australia).

Ownership of Intellectual property rights

Unless otherwise noted, copyright (and any other intellectual property rights, if any) in this publication is owned by the Fisheries Research and Development Corporation and CSIRO Oceans and Atmosphere.

This publication (and any information sourced from it) should be attributed to:

Campbell, R.A., S. Zhou, S.D. Hoyle, R. Hillary, M. Haddon and S. Hall. (2017) *Developing innovative approaches to improve CPUE standardisation for Australia's multispecies pelagic longline fisheries*. Final report for project 2014-021 to the Fisheries Research Development Corporation, Canberra, Australia.

Creative Commons licence

All material in this publication is licensed under a Creative Commons Attribution 3.0 Australia Licence, save for content supplied by third parties, logos and the Commonwealth Coat of Arms.



Creative Commons Attribution 3.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. A summary of the licence terms is available from creativecommons.org/licenses/by/3.0/au/deed.en. The full licence terms are available from creativecommons.org/licenses/by/3.0/au/legalcode.

Inquiries regarding the licence and any use of this document should be sent to: frdc@frdc.com.au

Disclaimer

The authors do not warrant that the information in this document is free from errors or omissions. The authors do not accept any form of liability, be it contractual, tortious, or otherwise, for the contents of this document or for any consequences arising from its use or any reliance placed upon it. The information, opinions and advice contained in this document may not relate, or be relevant, to a readers particular circumstances. Opinions expressed by the authors are the individual opinions expressed by those persons and are not necessarily those of the publisher, research provider or the FRDC.

The Fisheries Research and Development Corporation plans, invests in and manages fisheries research and development throughout Australia. It is a statutory authority within the portfolio of the federal Minister for Agriculture, Fisheries and Forestry, jointly funded by the Australian Government and the fishing industry.

Researcher Contact Details

Name: Dr Robert Campbell
Address: CSIRO, Private Bag No. 1
Aspendale VIC 3195
Phone: 03 9239 4681
Fax: 03 9239 4444
Email: Robert.Campbell@csiro.au

FRDC Contact Details

Address: 25 Geils Court
Deakin ACT 2600
Phone: 02 6285 0400
Fax: 02 6285 0499
Email: frdc@frdc.com.au
Web: www.frdc.com.au

In submitting this report, the researcher has agreed to FRDC publishing this material in its edited form.

CONTENTS

List of Tables	vi
List of Figures	viii
List of Acronyms	xiv
1. Acknowledgements	1
2. Executive Summary	2
3. Introduction	7
3.1 The Eastern Tuna and Billfish Fishery	8
3.2 The Multispecies problem	11
4. Objectives	16
5. Methods	17
5.1 Overview (and Report Structure)	17
5.2 Review of current methods to standardise CPUE	18
5.3 Current method used for standardising CPUE in the ETBF	23
5.4 Clustering catch as a proxy for targeting	30
5.5 Bayesian spatial and spatial-temporal models using INLA	39
5.6 Development of multispecies catch and effort simulators	47
5.7 Models selected for comparative analysis	69
6. Results	78
6.1 Review of factors influencing CPUE	78
6.2 Targeting and gear setting practices in the ETBF	79
6.3 Species coexistence	96
6.4 Comparative performance of standardisation models	106
6.5 Further analyses	119
6.6 Implications for the ETBF harvest strategy	129
6.7 Further implications for the ETBF harvest strategy in a stock assessment context	137
7. Discussion and Conclusions	148
8. Implications	158
9. Recommendations	159
10. Extension and Adoption	161
11. References and Further Reading	163
Appendix A: Annual longline catch by species in the ETBF	170
Appendix B: Spatial and temporal distributions of CPUE in the ETBF	187
Appendix C: Clustering by deployed fishing gear	206
Appendix D: Spatial Dynamics Factor Analysis Model	212
Appendix E: List of project staff	219

List of Tables

Table 3.1:	The time-series of effort, catch, CPUE and value of catch for a hypothetical fishery described in the text.	13
Table 5.1:	Listing of explanatory variables, and category definitions, fitted to GLM used to standardise CPUE in the ETBF. The number of observations within each category for the swordfish analysis undertaken in 2016 is also shown.	26-27
Table 5.2:	Catch composition of the four clusters based on the cluster analysis for each month. The last column titled Target indicates the final targeting type (1-7) based on the final cluster analysis of the 48 clusters shown in the Table. (NOPS = Number of fishing operations.)	34
Table 5.3:	Composition of cluster types based on (a) monthly analyses combined into 7 clusters, and (b) single analysis of all data into 7 clusters. Note, the clusters in each sub-table are ordered by decreasing percentage of Yellowfin Tuna. (NOPS = Number of fishing operations.)	38
Table 5.4:	Comparison of relative error between predicted abundance index and true index across 6 models and 5 species.	45
Table 5.5:	Comparison of absolute relative error between predicted abundance index and true index across 6 models and 5 species.	46
Table 5.6:	Comparison of absolute relative error between predicted abundance index and observed index across 6 models and 5 species.	47
Table 5.7:	Mean Annual Error (<i>MAE</i>) for each GLM fitted to the four versions of the simulated data sets. For each data-set and species, the GLM with the smallest <i>MAE</i> is highlighted in light grey.	61
Table 6.1:	Pearson correlation coefficient between the <i>MAE</i> for each error type for the (a) 1D-Trial and (b) 1R-Trial analyses.	106
Table 6.2:	Species Ranked Scores (<i>SRS</i>) for each species and model and the total Model Ranked Score (<i>MRS</i>) for the 1D-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same <i>SRS</i> .	110
Table 6.3:	Species Ranked Scores (<i>SRS</i>) for each species and model and the total Model Ranked Score (<i>MRS</i>) for the 1R-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same <i>SRS</i> .	112
Table 6.4:	Pearson correlation, <i>R</i> , between the model ranks (1-24) for the 1-D and 1-R analyses.	114
Table 6.5:	Species Ranked Scores (<i>SRS</i>) for each species and model and the total Model Ranked Score (<i>MRS</i>) for the 2R-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same <i>SRS</i> .	115
Table 6.6:	Species Ranked Scores (<i>SRS</i>) for each species and model and the total Model Ranked Score (<i>MRS</i>) for the 3D-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same <i>SRS</i> .	118
Table 6.7:	Species Ranked Scores (<i>SRS</i>) for each model for the D9-Trial analyses. Models having equal rank scores are shaded.	125
Table 6.8:	Mean rank (1=Best, 6=Worst) of the performance of the A-1 model fitted to the 100 simulated ETBF data-sets using the R-catches and the six types of <i>Area</i> effects described in the text.	128
Table 6.9:	Over-all ranking and Model Rank Score (<i>MRS</i>) for each model based on ranking the mean RBCC-Error across all species for the two 1-Trial analyses. Best=1, Worst=24.	134
Table A.1:	Complete listing of the logbook reported catch of all species reported by longline vessels operating in the ETBF. (Note, the last column indicates whether the species is reported in main document.)	181-183

Table A.2: Listing by logbook of the reported catch of all species by longline vessels operating in the ETBF. (Note, the last column indicates whether the species is reported in main document.) 184-186

List of Figures

Figure 3.1: Area of the Eastern Tuna and Billfish Fishery (<i>source: Australian Fisheries Management Authority</i>).	9
Figure 3.2: (a) Number of longline vessels operating in the ETBF each year and the associated effort (number of hooks deployed), and (b) Distribution of number of sets per trip (2007 to 2013).	9
Figure 3.3: Example of completed AL06 logbook used in the ETBF indicating the catch and effort information collected for each longline set deployed.	10
Figure 3.4: The time-series of fishery indicators for a hypothetical fishery described in the text.	14
Figure 5.1: Total number of 1-degree squares fished each year in the ETBF and the corresponding number of squares in which each of the main target species were caught.	29
Figure 5.2: Pseudo-F criteria for identifying number of clusters. Results shown for each month.	31
Figure 5.3: Pictorial representation of monthly clusters based on plots of the first two canonical variables. (NB, the lines of points are assumed to be representative of the zero catches in the data.)	32-33
Figure 5.4: Dendograms of cluster associations. Results are shown for the following four months: (a) January, (b) February, (c) November and (d) December.	33
Figure 5.5: Clustering criteria used to identify the appropriate final number of clusters based on clustering the 48 monthly cluster types identified in the first stage of the analysis.	35
Figure 5.6: (a) Dendrogram of cluster associations based on analysis of the 48 monthly clusters, and (b) representation of the seven cluster types based on plots of the first two canonical variables for the 48 monthly clusters.	35
Figure 5.7: Distribution of the 48 monthly cluster types identified in stage 1 of the analysis across the seven cluster types identified in stage two of the analysis.	36
Figure 5.8: Distribution of cluster types (percentage of sets) across each (a) year and (b) month	36
Figure 5.9: Distribution of gear-settings across the seven cluster type identified in the 2-stage monthly analysis.	37
Figure 5.10: Catch composition by seven clusters based on (a) the analysis by month, and (b) the single analysis of all sets.	38
Figure 5.11: Building mesh for Gaussian Markov random field model in ETBF. The mesh is based on unique coordinates in the simulated data for all year (between 2000 and 2014). The numbers in red are areas used in GLM.	43
Figure 5.12: Comparison between GLM Model (1), the full spatio-temporal Model (5) and the simple spatio-temporal Model (6).	46
Figure 5.13: Hypothetical result of a GLM illustrating the relationship between the number of hooks-per-float and relative CPUE of both yellowfin and Bigeye Tuna.	48
Figure 5.14: Observed hook-depth profiles for eight different hooks-per-float longline configurations. Fitted LOWESS curves are also shown for two different smoothing parameters.	51
Figure 5.15: Hook-depth profiles used in the multispecies data simulator.	52
Figure 5.16: Mean observed depth profiles by month for YFT, BET, SBT and SWO based on data collected from archival tags. A fifth hybrid profile used in the multispecies data simulator is also shown.	52
Figure 5.17: Hourly mean observed depth profiles for Yellowfin Tuna during January and August.	53
Figure 5.18: An example of the density surface for Swordfish during June 2014 across the 139 1x1-degree squares incorporated into the multispecies data simulator.	54

Figure 5.19: Distribution of sets used in the simulation across (a) Year, b) Month, (c) Start-time, (d) Hooks-per-float, and (e) 1x1-degree squares.	54
Figure 5.20: Comparison of the distribution of simulated catches, both with (Random) and without (Determin.) the random component added, with the distribution of observed catches.	56
Figure 5.21: Comparison of the modelled annual abundance and the nominal CPUE based on the simulated catch and effort data (where both have been scaled so that the mean is one over the time-series).	56
Figure 5.22: Percentage of sets deployed each year stratified by (a) Start-time, (b) Hooks-per-float, (c) Longitude, (d) Latitude and (e) Month.	57
Figure 5.23: Distribution of the 139 one-degree squares included in the simulated data sets across the seven <i>Area</i> effects used in the associated GLM analyses.	58
Figure 5.24: The distribution of the data records and fishing effort across these seven areas together with the distribution of the D-catches and associated catch rates for one realization of the simulated data sets.	59
Figure 5.25: Annual time-series of (a) the log-relative error and (b) the relative abundance index for each species for the NEG-3 model fitted to each of the four data-sets	62
Figure 5.26: Annual time-series of (a) the log-relative error and (b) the relative abundance index for each species for each of the four GLMs fitted to the M-M data-set.	64
Figure 5.27: (a) Distribution of the annual abundance index for each year based on 100 random analyses with the mean indicated by the solid black line (and assumed true index shown by the red line). The mean of the nominal CPUE for each simulated data set is shown by the dashed grey line. (b) Distribution of the error for each year with the mean indicated by the solid black line.	65
Figure 5.28: Relative catchability by species and fishing tactic ($qtac_{species,tactic}$), as assumed in the simulation model.	68
Figure 6.1: Index of availability of the five primary target species in the ETBF as measured by nominal CPUE.	80
Figure 6.2: Percent of longline sets deployed in the ETBF each year using different categories of gear-settings. Information based on ETBF logbooks. (NR denotes Not Recorded, Mack=mackerel, Pilch=pilchard, D=dead bait, A=live bait, M=mixed dead and live bait).	82
Figure 6.3: Percent of longline sets deployed in the ETBF each year using different categories of gear-settings. Information based on ETBF logbooks. (UNK denotes Unknown).	83
Figure 6.4: Average setting (and standard deviation) of individual gear-types during each year based on observer data collected from longline vessels operating in the ETBF (solid blue line). The number of observer voyages from which data was available is also shown (green columns).	84
Figure 6.5: Time-series of branch-line lengths deployed on five different vessels operating in the ETBF. Each dot represents an individual fishing trips and the length recorded by an observer for that trip.	85
Figure 6.6: Annual distribution of hook-types deployed on observed longline vessel operating in the ETBF.	86
Figure 6.7: Histogram of the number of sets per observed trip.	88
Figure 6.8: Percentage of trips where the setting for each gear type was similar for all FOPS during a trip. The solid line displays the total number of trips for which data on the given gear type was available for all sets while the dashed line displays the number of trips for which the gear-settings were similar for all sets within a trip. Trips are stratified by the number of sets deployed per trip.	89
Figure 6.9: Percentage of trips where the setting for each gear type was similar (i.e. the same category level was used) for all FOPS during a trip. The solid line displays the total number of trips for which data on the given gear type was available for all sets while	

the dashed line displays the number of trips for which the gear-settings were similar for all sets within a trip. Trips are stratified by the number of sets deployed per trip. For those trips deploying the same gear type for all sets the percentage of trips using each gear-type category is also displayed.	90
Figure 6.10: Distribution of various gear-setting stratified by primary target species recorded by observer– ordinal variables.	92
Figure 6.11: Distribution of various gear-setting stratified by primary target species recorded by observer – categorical variables.	93
Figure 6.12: Percent of observed fishing operations by primary target species (a) by month and (b) aggregated by month across years.	94
Figure 6.13: Catch composition of all observed sets by (a) month and (b) primary target species. (YFT=Yellowfin Tuna, BET=Bigeye Tuna, ALB=Albacore Tuna, SWO=Broadbill Swordfish, STM=Striped Marlin, SBT=Southern Bluefin Tuna, LEC=Escolar/Oilfish, ALX=Long-nosed Lancetfish, DOL=Dolphinfish, POA=Ray's Bream/Promfrets, BSH=Blue Shark, SKJ=Skipjack Tuna, WAH=Wahoo, MAK=Mako Shark, GES=Snake Mackerel, ALO= Short-nosed Lancetfish, MOP=Sunfish, SBS=Short-bill Spearfish, BAM=Black Marlin, BUM=Blue Marlin, OPA=Opah, RAY=Stingrays, TIG=Tiger Shark, BRO=Bronze Whaler, DSK=Dusky Shark, OCS=Oceanic Whitetip Shark, CSH=Crocodile Shark, SKS=Silky Shark and SPN=Hammerhead Sharks).	95
Figure 6.14: Cumulative percent of sets catching less than or equal to the indicated number of fish stratified by primary target species.	96
Figure 6.15: Pearson correlation between quarters of the spatial distributions of CPUE for each species.	97
Figure 6.16: Pearson correlation, by quarter, between the spatial distributions of CPUE for a given specie and all other species.	98-101
Figure 6.17: Comparison of results of cluster analysis for summer and winter data.	103
Figure. 6.18:(a) Composition of the mean monthly catch recorded in ETBF logbooks, and (b) Percentage of each species recorded in the ETBF logbook comprising the mean monthly catch.	104
Figure 6.19: Catch comparison of clusters based on analyses using different number of species (B) comprising the total catch and a subset of different species (A) upon which the clusters were defined. Individual results are labelled (AxB).	105
Figure 6.20: (a) Error-1 results for the Group A and B models fitted to the ETBF simulated dataset. Results are shown for each species and for three sets of Trials described in the text.	107
Figure 6.20: (b) Error-2 results for the Group A and B models fitted to the ETBF simulated dataset. Results are shown for each species and for three sets of Trials described in the text.	108
Figure 6.21: Minimum and maximum MAEs for each the four Error-types and species for the 1D-Trials and 1R-Trials.	109
Figure 6.22: Comparison of the Species Ranked Scores (SRS) for each species and model for the 1D-Trial analyses. The mean and standard deviation of the SRS across all five species is also shown.	110
Figure 6.23: Comparison of the rank of models which include the two gear effects in the fitted GLM with those models which do not include these effects. The comparison is shown for both the 1-D and 1-R analyses for (a) all species and (b) the four species other than ALB.	111
Figure 6.24: Comparison of the Species Ranked Scores (SRS) for each species and model for the 1R-Trial analyses. The mean SRS across all five species is also shown.	112
Figure 6.25: Comparison of the model ranks (1-24) between the 1-D and 1-R analyses by species.	113

Figure 6.26: Comparison of the mean Error-1 and Error-2 results across all species for models fitted to the D-catches and the R-catches. The comparison is also shown for both sets of analyses which include the two gear effects in the fitted GLM with those models which do not include these effects.	114
Figure 6.27: Comparison of the Species Ranked Scores (<i>SRS</i>) for each species and model for the 2R-Trial analyses. The mean <i>SRS</i> across all ten species is also shown.	115
Figure 6.28: Comparison of the overall rank of the eleven models used in both the 1R-Trial (ETBF) and 2R-Trial (DIST) analyses.	116
Figure 6.29: Error-1 and Error-2 results for the Group A (blue) and C models (green) fitted to the streamlined (N3) ETBF simulated data-set. Results are shown for each species. Note, y-axis values differ between panels.	117
Figure 6.30: Comparison of the Species Ranked Scores (<i>SRS</i>) for each species and model for the 3D-Trial analyses. The mean <i>SRS</i> across all five species is also shown.	118
Figure 6.31: Comparison of the four index-based errors after fitting the following four variates of the A-1 GLM to the 100 sets of R-catches for the ETBF simulated data set: (i) a two-stage delta-(Bin-NegBin) model with either the two <i>Year*Qtr</i> and <i>Qtr*Area</i> 2-way interactions (Delta-2) or the single 3-way <i>Year*Qtr*Area</i> interaction (Delta-3) and (ii) a single-stage-NegBin-only model with either the two <i>Year*Qtr</i> and <i>Qtr*Area</i> 2-way interactions (NoDelta-2) or the single 3-way <i>Year*Qtr*Area</i> interaction (NoDelta-3).	119
Figure 6.32: Gear types associated with fishing operations included in the summer and winter YFT and ALB clusters identified in Section 6.3 (c.f. Figure 6.17.)	121
Figure 6.33: (a) Comparison of the <i>MAE</i> associated with the fitting the A-3 model to the ETBF data-set where the linear predictor includes; (i) no gear effects, i.e. cluster effects only, (ii) both gear and cluster effects, or (iii) the gear-by-cluster interactions.	121
Figure 6.33: (b) Comparison of the <i>MAE</i> associated with the fitting the A-5 model to the ETBF data-set where the linear predictor includes; (i) no gear effects, i.e. cluster effects only, (ii) both gear and cluster effects, or (iii) the gear-by-cluster interactions.	121
Figure 6.34: Pearson correlation between the annual distributions of nominal CPUE between years at: (a) the 1x1-degree level, and (b) the GLM-area level.	123
Figure 6.35: Annual distributions of nominal ALB CPUE for the simulation ETBF data-set. The colours indicate the following range of CPU: (i) yellow, <2.0, (ii) orange, <4.0, (iii) green, <6.0, (iv) light blue, <8.0, (v) dark blue, <10.0, (vi) red, <12.0, and (vii) purple, >12.	124
Figure 6.36: Comparison of (a) the standardised CPUE indices and (b) the index errors based on the A-1a model for all years and the A-3 model for the last 9 years. The nominal CPUE and assumed abundance indices are also shown.	125
Figure 6.37: Comparison of the overall model rank for the five Group-A models based in the D9-Trial analyses with and without the two gear effects included in the linear predictor.	126
Figure 6.38: The Mean Annual Error (<i>MAE</i>) for the A-1 model fitted to the 100 simulated ETBF data-sets using the R-catches and the six types of <i>Area</i> effects described in the text. Note: NB-2 refers to the NEG-2 model, NB-3 refers to the NEG-3 model while ERR-1 refers to Type-1 Error and ERR_2 refers to the Type-2 Error.	128
Figure 6.39: Conceptual example of how the slope-to-target parameter used in the primary control rule is derived.	130
Figure 6.40: For the four simulation trials undertaken based on the ETBF data, comparison of the Mean Annual Error in the RBCC by species and model.	131
Figure 6.41: For each species and model, comparison of the Mean Annual Error in the RBCC across the three simulation trials undertaken based on the ETBF data.	132

Figure 6.42: Distribution of the mean RBCC-Error (rounded to the nearest percent) across the 24 models for the two 1D-Trial and 1R-Trial analyses. Results are shown for both the mean across all five species and across the four species excluding DOL.	133
Figure 6.43: For each species, comparison of the mean and standard deviation of the distribution of the nominal values of the RBCC-error across the 100 simulations.	134
Figure 6.44: For each species (a) example comparison of the assumed abundance index, standardised CPUE index and three LOWESS smoothed indices for a single analysis and (b) the distribution of the errors in the RBCCs calculated using either the standardised CPUE or the LOWESS smoothed indices across 100 simulations.	135
Figure 6.45: For each species plot of the standard deviation of the mean of the four RBCC-errors shown in Figure 6.44a versus the CV of the last 5 data points of the standardised CPUE.	136
Figure 6.46: Comparison of the mean and standard-deviation of the distribution of the RBCC-error across the 100 simulations discussed in the text.	137
Figure 6.47: Boxplot summary of the Swordfish CPUE for all size classes (small, prime, large) and for each of the six series (true, nominal, the two Gamma and two Negative-Binomial models).	138
Figure 6.48: Relative errors (median and 80% CI) in the TACC predicted using the whole data series (15 years) for each of the 100 simulations. The percentage error is expressed relative to the true TACC for each of the other candidate series.	139
Figure 6.49: Boxplot summaries of the AAV statistics for each of the six CPUE series, for a five year quasi-retrospective, and across all 100 simulations.	140
Figure 6.50: CPUE indices (mean standardised) from 1997 to 2014 for each of the five main target species in the ETBF.	144
Figure 6.51: Predicted log-scale CPUE (left) for the Bigeye and Yellowfin Tuna data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the right the predicted relative abundance, again in terms of the MLE and approximate 95% CI.	145
Figure 6.52: Predicted log-scale CPUE (top left) for the Swordfish and Striped Marlin data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the top right is the predicted relative abundance, again in terms of the MLE and approximate 95% CI. On the bottom left is the MLE and approximate 95% CI of the inferred F_y terms, given the effort time-series and the species-shared catchability term.	146
Figure 6.53: Predicted log-scale CPUE (left) for the Albacore data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the right is the predicted relative abundance, again in terms of the MLE and approximate 95% CI.	147
Figure 7.1: Comparison of the overall rank achieved by each model (1=Best, 24=Worst) when used to estimate either the abundance index over all years or the RBCC. The analyses are based on the average errors calculated across 100 ETBF data-sets where the catches were sampled from a random negative binomial distribution.	156
Figure A.1: Annual logbook coverage (as a percentage of sets) in the ETBF. Note: ELINE refers to an electronic logbook.	170
Figure A.2: Logbook recorded annual catch (number of fish retained and discarded) of tuna species in the ETBF.	173
Figure A.3: Logbook recorded annual catch (number of fish retained and discarded) of billfish species in the ETBF.	174
Figure A.4: Logbook recorded annual catch (number of fish retained and discarded) of shark species in the ETBF.	175-177
Figure A.5: Logbook recorded annual catch (number of fish retained and discarded) of byproduct species in the ETBF.	178

Figure A.6: Logbook recorded annual catch (number of fish retained and discarded) of bycatch species in the ETBF.	179
Figure A.7: Logbook recorded annual catch (number of fish retained and discarded) of skate and ray species in the ETBF.	180
Figure B.1: Spatial distribution (1x1-degree) of aggregate longline effort in the ETBF over the years 2000-2013 for each quarter of the year.	188
Figure B.2: (a-q): Spatial distribution (1x1-degree) of nominal CPUE by species in the ETBF over the years 2000-2013 for each quarter of the year.	189-205
Figure C.1: Clustering criteria used to identify the appropriate final number of clusters based on clustering the 84 monthly cluster types identified in the first stage of the gear-based analysis.	207
Figure C.2: Pictorial representation of the nine cluster types based on plots of the first two canonical variables for the 84 monthly gear-based clusters.	207
Figure C.3: Distribution of the 84 monthly cluster types identified in stage 1 of the gear-based analysis across the nine cluster types identified in the second cluster analysis.	208
Figure C.4: Distribution of nine cluster types (percentage of sets) based on the gear-based cluster analysis across (a) month, and (b) year.	209
Figure C.5: Distribution of gear-settings across the nine gear-based cluster types identified in the 2-stage monthly analysis.	210
Figure C.6: Catch composition of the nine clusters based on (a) the analysis by month, and (b) the single analysis of all sets.	211
Figure D.1: Comparisons by species of annual indices between true, nominal, and SDFA model results. The SDFA-R10 model indices are the means of the 10 model runs using random data, while the SDFA-D model indicates the index for the single run using the deterministic catches.	214
Figure D.2: Comparisons among the Group-A and SDFA models by species, for the four error types. All results based on fitting each model to the 10 realisations of the ETBF dataset using the random catches.	215
Figure D.3: Comparison of the Species Ranked Scores (SRS) for the Group-A and SDFA models by species, when fitted to the 10 realisations of the ETBF dataset using the random catches.	215

List of Acronyms

AAV	Average Annual Variation
ABARES	Australian Bureau of Resource Economic Sciences
AIC	Akaike Information Criterion
AFMA	Australian Fisheries Management Authority
ALB	Albacore Tuna
ARE	Absolute Relative Error
BAM	Black Marlin
BBL	Broadbill Swordfish
BET	Bigeye Tuna
BIC	Bayesian Information Criterion or Schwarz Criterion
BSH	Blue Shark
BUM	Blue Marlin
CCC	Cubic Cluster Criterion
CCSBT	Committee for the Conservation of Southern Bluefin Tuna
CPUE	Catch-per-Unit-Effort
CV	Coefficient of Variation
DPC	Direct Principal Component
DOF	Degrees of Freedom
DOL	Dolphin Fish
ETBF	Eastern Tuna and Billfish Fishery
FOP	Fishing Operation
FRDC	Fisheries Research and Development Corporation
GMRF	Gaussian Markov Random Fields
HBS	Habitat-Based Standardisation
HPF	Hooks-per-Float
HS	Harvest Strategy
GAM	Generalised Additive Model
GAM -2	GLM with Gamma error with two-way interactions
GAM -3	GLM with Gamma error with three-way interaction
GLM	Generalised Linear Model
GLMM	Generalised Linear Mixed Model
ICES	International Council for the Exploration of the Sea
INLA	Integrated Nested Laplace Approximation

IOTC	Indian Ocean Tuna Commission
ITQ	Individual Transferrable Quota
LOWESS	Locally Weighted Scatterplot Smoother
LOOCV	Leave-One-Out Cross-Validation
MAE	Mean Annual Error
MAK	Mako Shark
MARE	Mean Absolute Relative Error
MCMC	Markov Chain Monte Carlo
MEY	Maximum Economic Yield
MRS	Model Ranked Score
MSE	Management Strategy Evaluation
MSY	Maximum Sustainable Yield
MS-MSY	Multispecies Maximum Sustainable Yield
NBT	Northern Bluefin Tuna
NEG-2	GLM with Negative Binomial error with two-way interactions
NEG-3	GLM with Negative Binomial error with three-way interaction
NOAA	National Oceans and Atmospheric Administration
OIL	Oilfish
OPA	Opah (Moonfish)
PCA	Principal Components Analysis
POA	Promfrets / Ray's Bream
RAG	Resource Assessment Group
RBCC	Recommended Biological Commercial Catch
RFMO	Regional Fishery Management Organisation
RMSE	Root Mean Squared Error
SAF	Sailfish
SBS	Short-Billed Spearfish
SBT	Southern Bluefin Tuna
SDFA	Spatial Dynamics Factor Analysis
SESSF	South-east Scalefish and Shark Fishery
SKJ	Skipjack Tuna
SPC	Secretariat of the Pacific Community
SPDE	Stochastic Partial Differential Equation
SRS	Species Ranked Score

STIME	Start-Time of set
STM	Striped Marlin
SWO	Broadbill Swordfish
TACC	Total Allowable Commercial Catch
TDR	Temperature-Depth Recorder
TMB	Template Model Builder
TTRAG	Tropical Tuna Resource Assessment Group
WAH	Wahoo
WCPFC	Western and Central Pacific Fisheries Commission
WCPO	Western Central Pacific Ocean
WTBF	Western Tuna and Billfish Fishery
YFT	Yellowfin Tuna

1. Acknowledgments

This work was supported by funding from the Fisheries Research Development Corporation. The authors would like to acknowledge the support of James Thorson (Northwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Seattle, USA) for the provision of the source code for the Spatial Dynamic Factor Analysis (SDFA) model developed by Thorson *et al.* (2016) and his assistance in running this code. Thanks are also extended to Ross Marriott (University of Western Australia, Perth) for his thoughtful and constructive comments on an earlier draft of this report and Christopher Izzo (FRDC) for additional editorial comments, both of which helped improve the content and clarity of this final report.

2. Executive Summary

This project was undertaken by a collaboration of senior fishery scientists at CSIRO and from New Zealand, together with a former fisheries manager now with the Commonwealth Department of Agriculture and Water Resources in Canberra, on the development of methods to construct indices of stock abundance trends from commercial catch-per-unit-effort (CPUE) in multispecies pelagic longline fisheries. Such indices are crucial inputs into stock assessments undertaken around the world and play a vital role in achieving the sustainable management of global fisheries. The project work was undertaken during 2015 and 2016, using the multispecies longline fishery for tuna and billfish on the east coast of Australia (the Eastern Tuna and Billfish Fishery) as the example case study. As indices of stock abundance constructed from CPUE data are the central inputs into the harvest strategy used in this fishery to inform the determination of annual Total Allowable Commercial Catch (TACC) limits, there was a need to identify the accuracy of current methods and develop new methods to construct more reliable indices of stock abundance. In this regard, the analyses undertaken during the project and presented here were designed to address specific issues related to this fishery. However, it is also hoped that the general results of this project will have broader applicability to other multispecies species, both domestically and internationally.

Background

Australia's two tropical tuna fisheries (the Eastern Tuna and Billfish Fishery, ETBF, and the Western Tuna and Billfish Fishery, WTBF) are both multispecies fisheries which target a range of large pelagic fish. However, it is often cited that a major constraint for assessing multispecies fisheries is a lack of reliable abundance indices that are a pre-requisite for the accompanying stock assessments, and this has flow-on impacts on identifying appropriate management measures (e.g. such as TACCs obtained from harvest strategies). Unlike single species fisheries where all effort is directed at the target species, in multispecies fisheries the effort is directed at a range of species. Consequently, the fishing effort needs to be standardised so that the 'effective' effort directed at any single species of interest can be ascertained. If this is not undertaken correctly then the resulting index of resource abundance is likely to be biased and unreliable. Although there are methods available that are currently used to standardise effort in multispecies fisheries, these methods need to be evaluated and where necessary new techniques need to be developed so that the resulting indices of resource abundance based on standardising CPUE can be made more reliable.

Objectives

The project had the following five objectives:

- (1) to identify the factors likely to influence CPUE in multispecies pelagic longline fisheries,
- (2) review methods which may be used for standardising CPUE in multispecies fisheries,
- (3) develop and compare the most appropriate methods for standardising CPUE for multispecies pelagic longline fisheries, (
- 4) use simulated catch and effort data to test the potential of each method to adequately account for the influence of factors influencing CPUE and accurately reflect the underlying resource abundance, and
- (5) investigate the sensitivity of the outcomes of the ETBF harvest strategy on the adoption of the candidate methods for standardising CPUE within the ETBF.

Methods

Targeting Practices in the ETBF

To adequately standardise CPUE one needs an understanding of the relationship between how the fishing gear is deployed and the subsequent catch. Factors influencing this relationship include both the availability of fish to the fishing gear and the catchability of the fishing gear. Availability will be influenced by the environmental and oceanographic conditions prevailing at the time of the fishing operation, while catchability will be influenced by the types of fishing gear used and the manner in which it is deployed. In a multispecies fishery there is particular interest in understanding the differences in fishing tactics utilised when targeting different species, as this information can be used to identify the 'effective' effort directed at any single species. For the ETBF the relationships between fishing practices and catch, together with changes in the seasonal availability and distribution of co-related species (i.e. species caught within the same set) in the fishery, was undertaken using the catch and effort data recorded in logbooks and by on-vessel observers.

Development of Standardisation Models

After a review of existing methods proposed and used in standardising multispecies CPUE, a number of methods were selected for further development. While most models are based on the Generalised Linear Model (GLM) framework developed for single species analyses, they included a number of extensions aimed at standardising the fishing effort to account for the differential targeting of species within a multispecies fishery. One set of extensions utilises a cluster analysis approach to group the fishing operations for a fishery into groups of different targeting or fishing strategies based on the species composition of the catch. The utility of clustering across different levels of data aggregation (i.e. at the set or trip level), as well as undertaking separate analyses at different temporal periods (e.g. monthly) to account for seasonal changes in species availability and assemblages, are also investigated. A second set of models uses a Principal Components Analysis approach to group fishing operations, while a third, and new, set of Bayesian spatial-temporal models utilising the relatively new tool Integrated Nested Laplace Approximation (INLA) were developed. In total, the performance of twenty-four different model variations were investigated.

Development of the Simulation Framework

Two simulators, which are useful for comparing the performances of different methods at estimating known quantities, were developed for testing the comparative performance of the standardising models. The first simulator utilized an empirical approach, based on a framework generally known as the habitat-based-standardisation method, and made direct use of observations of longline hook depths and fish depth profiles obtained from archival tags deployed on fish caught and tagged within the ETBF. The spatial-temporal distribution of fishing effort across the fishery was taken to be the same as that observed in the ETBF during the years 2000 to 2014, while the spatial-temporal density surfaces for each of the five species included in the simulator were modelled on the monthly distributions of nominal CPUE observed in the ETBF over this period. Finally, in order to simulate the stochastic characteristic of any catch, the simulated catch was selected from a negative binomial distribution and the probability of success for each species selected so that the distribution of simulated catches for each species was similar to that in the distribution of observed catches. For the second simulator, catch and effort data was modelled using a more generic, flexible and individual-based approach and was designed to generate catch rate data that reflected species abundance, targeting practises, and (unlike the ETBF simulator) individual vessel efficiency, to capture the

fact that different vessels have their own characteristic catchability and fishing behaviour. There is also spatial and seasonal variation in catch rates, independently among species.

Results

Targeting Practices in the ETBF

The information recorded in the ETBF logbooks, and by observers, indicated that there can be considerable variability in the manner that longline gears are deployed, both among years and within years (and trips) on individual vessels. A substantive change was noted in 2006 when vessels began deploying more than 25 hooks-per-float (HPF). Commensurate with this change were shifts in the bait usage (more Australian Sardine –commonly called as pilchard), light-stick usage (fewer) and set start-time (earlier in the day). The introduction of this new ‘deep-longline’ technique was due to a significant change in fishing strategy to incorporate the direct targeting of Albacore Tuna in this fishery. Fortunately, each of these gear settings are recorded in the ETBF logbook, and so these changes can be taken into account when standardising CPUE for changes in targeting and associated fishing strategies. On the other hand, there have also been changes in other gear settings which are not recorded in the logbook and therefore cannot be accounted for in the CPUE standardisation, though the results of previously published research can provide some guidance.

Analysis of the gear setting practices within a fishing trip indicate variable degrees of consistency in the deployment of each gear type. For example, similar hooks-per-float settings are used for 60-80% of all sets during a trip, while generally similar start-times are used for less than 20% of sets. Observer data also indicates that the recorded primary target species is not always the same for all sets deployed during a trip and suggests that multiple targeting strategies can be utilised within individual trips. Examination of the particular gear settings associated with the target species recorded by the observers, also indicated that while a range of gear settings are utilised when targeting particular species, there are some combinations of gears that are more commonly used. Observer data also indicates that the proportion of fishing operations targeting particular species changes during the year, and suggests that the seasonal availability of the principal target species changes throughout the year, and the fishers are able to change their targeting practices to avail themselves of these changes in relative abundance. Changes in the temporal availability of species is likely linked with the movement of fish, associated with changes in the oceanographic conditions within the ETBF. Catch data was used to investigate the seasonal persistence of species associations and the results suggest that if associations between species do exist, due possibly to co-habitation within defined habitats, then these associations show generally weak persistence throughout the year (or across years).

Finally, the catch composition of observed sets also indicates that the proportional catch of a given species is generally highest when that species is recorded as the primary targeted species. This suggests that the vessel (skipper) has some ability to target and catch a desired species. Nevertheless, it is also clear that each of the other principal species are also usually caught, indicating that it is not possible to just target and catch a single species. Indeed, the target species sometimes is not the dominant catch. Cluster analyses of the ETBF catch data also indicates strong seasonal changes in species composition of the major clusters identified, which is likely to be related to seasonal differences in the co-occurrence of species reported above. Furthermore, while the fishing operations associated with some clusters display distinct differences in the gear configurations, most clusters were characterised by a broad mix of gear configurations, indicating that the relationship between the composition of the catch and the configuration of the gear is not strong. This suggests that ‘targeting’, and the consequent

composition of the catch, is likely to depend on more factors than just the configuration of the fishing gear, and exploring the nature of the relationship between the catch and these other factors (e.g. availability effects associated with the spatial location such as water temperatures, sea-mounts, eddy features) is encouraged.

Comparative Performance of Standardisation Models

The main results from these trials were as follows:

- The relative performance of each model varied considerably between the analyses conducted on the different species included in the two simulated data-sets. As a consequence, there was no one best performing model across all species.
- The overall best performing model fitted to the deterministic catches of the ETBF-simulated data-set was the model where the clustering was undertaken at the set level, while the overall best performing model fitted to the catches randomly sampled from a negative-binominal distribution was the model where the clustering was undertaken at the trip level. The higher variability in species composition at the set level with the use of the randomly sampled catches may be leading to a higher misallocation of sets using different fishing strategies and aggregating the data across trips may help to reduce this variability, and therefore the misallocation of sets.
- Models which included a two-stage delta-GLM approach performed considerably better than models which only used a single staged approach.
- Models which included the two gear-effects included in the simulated data also generally out-performed models which did not include these effects. This indicates that the inclusion of gear effects in the standardizing models has greater explanatory power than the inclusion of derived effects, such as those based on catch-composition derived clusters.
- The current GLM used in the ETBF to standardise CPUE is performing reasonably well, no doubt due to the fact that this model incorporates several of the features (e.g. two-stage analysis, inclusion of gear effects) that were found to perform well in the simulations undertaken.

Despite the above results, a consistent observation from all the analyses conducted was the variability in the performance of each model across the species included in the simulated data-sets. The reasons for variability remain uncertain, but further investigations found that taking account of substantive changes in the spatial distribution of catch rates can improve the performance of models used to standardise CPUE. Whether such patterns explain the variability in the performance of the models across all species remains unclear, though this result does demonstrate that there may be a wide range of factors specific to each species that influence the fit of any model to the data, and as such the estimation of the annual abundance index. No doubt further research is required to identify and improve our understanding of the factors used by fishers to 'target' the deployed effort and which control the composition of the catch for individual fishing operations.

Implications for the ETBF Harvest Strategy

The performance of the harvest strategy used in the ETBF in determining a 'correct' Recommended Biological Commercial Catch (RBCC) was tested using the same simulation framework as that used for testing the standardising models. As with the previous Index-based results, the size of the error in the RBCC was found to vary across the different species and models, such that no single model performed best across all species. The mean error was generally less than 4% (and often less than 2%), but for one of the species analysed the error

was higher at around 6-7%. Results also indicate that the direction of the error was not consistent, being conservative and under-estimating the true RBCC for some species, while over-estimating the true RBCC for other species. The best performing model was the delta-GLM cluster-by-trip model closely followed by the base delta-GLM model currently being used in the ETBF.

In some situations the performance of models to standardise CPUE was also found to be dependent on the time-series of data being analysed. This implies that analysts should give some consideration to the question of over what time-period of data should an analysis be conducted taking into consideration the end-purpose for which the results are to be used (e.g. abundance index for use in a stock assessment or estimation of abundance trend over last five years for use in a harvest control rule). This decision should be guided by investigation of both temporal changes in the characteristics of the data to be analysed that may influence (or bias) subsequent results, and the sensitivity of the constructed abundance index to changes in the time-series of data included in the analyses.

Implications for relevant stakeholders

Within Australia the standardised CPUE is a central input to the assessments and harvest strategies for the ETBF as well as other fisheries, including the multispecies South-East Scalefish and Shark Fishery (SESSF). The outcomes of this project will provide guidance on improving the methods used to standardise CPUE in these fisheries which should have follow on benefits to ensuring (i) more reliable and accurate stock abundance indices, (ii) improved inputs and to the harvest control rules dependent on standardised CPUE, (iii) improved outputs of harvest strategies, in particular the appropriateness of identified TACCs, and (iv) improvements to our ability to assess the resource status of non-target species as required to achieve the management objective of ecologically sustainable fisheries.

The outcomes of this project will also benefit fish stock assessments associated with the pelagic fisheries within the Western Central Pacific Ocean and Indian Ocean, to which Australia's domestic tuna fisheries are connected.

Recommendations

Based on the outcomes of this project, a number of recommendations have been made including (i) that the results of this project should be taken into consideration by fisheries scientists undertaking analyses of catch and effort data and for selecting appropriate methods for standardizing CPUE for stock assessment purposes, and (ii) that the recording in logbooks of information on the characteristics of the fishing gears deployed by fishers at the set level should be encouraged, and where this information is available it should be incorporated into the models used to standardise CPUE. Some recommendations are also made for further research.

Keywords

Catch-per-unit-effort, standardised CPUE, Generalised Linear Models, INLA, abundance indices, harvest strategies, pelagic longline, tunas and billfish, Eastern Tuna and Billfish Fishery

3. Introduction

For most of the past century fisheries scientists have tried to use catch and effort data collected from a fishery to infer trends in population abundance (Beverton and Holt 1957; Kimura 1981, 1988; Smith 2007) and the construction of relative indices of stock abundance from commercial catch-per-unit-effort (CPUE) data continues to be a routine and important aspect of many stock assessments worldwide. In particular, CPUE is usually assumed to be proportional to abundance and therefore changes in CPUE between years are assumed to be informative about changes in population abundance. To this end, CPUE based-indices are often included in the stock assessment as a relative index of abundance. However, many jurisdictions do not analyse fishery catch rates due to concerns that the relationship between nominal CPUE and abundance may be weak because CPUE may also reflect changes of catchability as well as population abundance, i.e. these data confound changes in fishing behaviour (adjustments in fishing location or fishing gear operation) with trends in abundance (Harley *et al.* 2006). As a consequence, many methods have been developed over the years to help “standardise” catch rates so that such changes in fisher behaviour (as distinct from changes in fish populations) that influence CPUE are removed from the index, so that changes in the standardised CPUE better represents changes in fish populations (see Maunder and Punt 2004 for a review; Maunder *et al.* 2006b; Bishop *et al.* 2008).

The usefulness of CPUE as an index of resource abundance depends on many assumptions. Central to these assumptions is that a unit of effort is uniformly effective across all areas, seasons and in all environmental conditions. For single species fisheries, like that for Southern Bluefin Tuna, where fishers do not generally change their fishing strategies this assumption will generally hold. However, within a multispecies fishery fishers may alter their fishing strategy in order to target different species. Such changes, such as changes in gear configuration, allow the fisher to switch the effectiveness of the fishing operation from one species to another. Consequently, the effectiveness with which the unit measure of effort catches different species is altered by the choice of fishing strategy.

The need to standardise catch rates is therefore particularly apparent in multispecies fisheries as research suggests that fishers can affect the assortment of species caught in a multispecies fishery by modifying the location, timing, and gear characteristics of their fishing activity (Sanchirico *et al.* 2006; Abbott *et al.* 2015). Furthermore, these operational choices, usually driven by commercial considerations such as profit, quota availability and market demand, influence how the catch is distributed across space, time, and species. So unlike single species fisheries where all effort is directed at the target species, in multispecies fisheries the effort is directed at a range of species. Consequently, the fishing effort needs to be adjusted so that the “effective” effort directed at any specific species of interest can be ascertained. If this is not undertaken appropriately then the resulting index of resource abundance is likely to be biased and unreliable. An example illustrating this problem is provided in the next Section. A major constraint for assessing multispecies fisheries, therefore, is a lack of reliable abundance indices that are in many instances a pre-requisite for the accompanying stock assessments and any applicable harvest strategy and has flow-on impacts to the results obtained (such as the determination of TACCs).

Although there are methods available that are currently used to standardise effort in multispecies fisheries (see Section 5.2 for a review), the precision and reliability of these

methods currently remains uncertain and it is generally believed that new techniques need to be developed to overcome these problems.

Australia's two tropical tuna fisheries (the Eastern Tuna and Billfish Fishery, ETBF, and the Western Tuna and Billfish Fishery, WTBF) are both multispecies fisheries that target a range of large pelagic fish. A fuller description of the ETBF is provided in Section 3.1. Currently the assessment of the resource status for these two fisheries, together with the related harvest strategy used to provide advice on TACCs for the principal target species, is based on the standardisation of catch and effort (CPUE) data collected from these fisheries. As effective fisheries management is dependent on good scientific advice having the least bias, improved management of these fisheries would be achieved with the identification and/or development of better methods to analyse the related CPUE data and construct more reliable indices of abundance. This would also help improve the management of other multispecies fisheries that rely on the use of CPUE-based abundance indices (e.g. the South Eastern Scalefish and Shark Fishery, SSSF).

In achieving the above outcome this project directly addresses the following FRDC strategic theme: (4) Ecologically sustainable development, which is concerned with the use and management of aquatic resources. In particular, by developing improved methods to standardise CPUE and construct more reliable indices of abundance, this project addresses the Theme 4 priority by assisting end-users to *"develop practical tools that implement ecosystem-based fisheries management and incorporate understandings of the cumulative impacts of fishing into fisheries management plans"*.

The project developed from discussions with the Tropical Tuna Resource Assessment Group (TTRAG) to construct more reliable indices of resource abundance as these indices are used in the harvest strategy used to manage the related fisheries. It also builds on the initial review undertaken of changes in fishing and targeting practices in the ETBF after the introduction of Individual Transferrable Quotas (ITQs) in 2011 (see Preece *et al.* 2013).

3.1 The Eastern Tuna and Billfish Fishery

The Eastern Tuna and Billfish Fishery (ETBF) targets highly migratory tuna and billfish species in Australian waters and on the high seas off the east coast of Australia. The area of the ETBF (Figure 3.1) includes:

- waters within the Australian Fishing Zone (that is, from the outer limit of state waters to the edge of the Australian exclusive economic zone), extending from Cape York around the eastern and southern coast of Australia to the South Australian/Victorian border;
- Commonwealth waters around Norfolk Island, excluding waters adjacent to Norfolk Island within the 'Norfolk Island box' (see Figure 3.1)
- the High Seas area of the Pacific Ocean.

The species caught in the ETBF are considered to be part of larger stocks extending across all or sub-regions of the Western Central Pacific Ocean (WCPO) and which are managed by the Western and Central Pacific Fisheries Commission (WCPFC). The stock status of these species is based on regional stock assessments which are conducted for the WCPFC.

There are two sectors in the ETBF, the pelagic longline sector and the minor line sector (hand lining, trolling and rod and reel fishing). The longline sector makes up the majority of the

Figure 3.1 Area of the Eastern Tuna and Billfish Fishery (source: Australian Fisheries Management Authority).

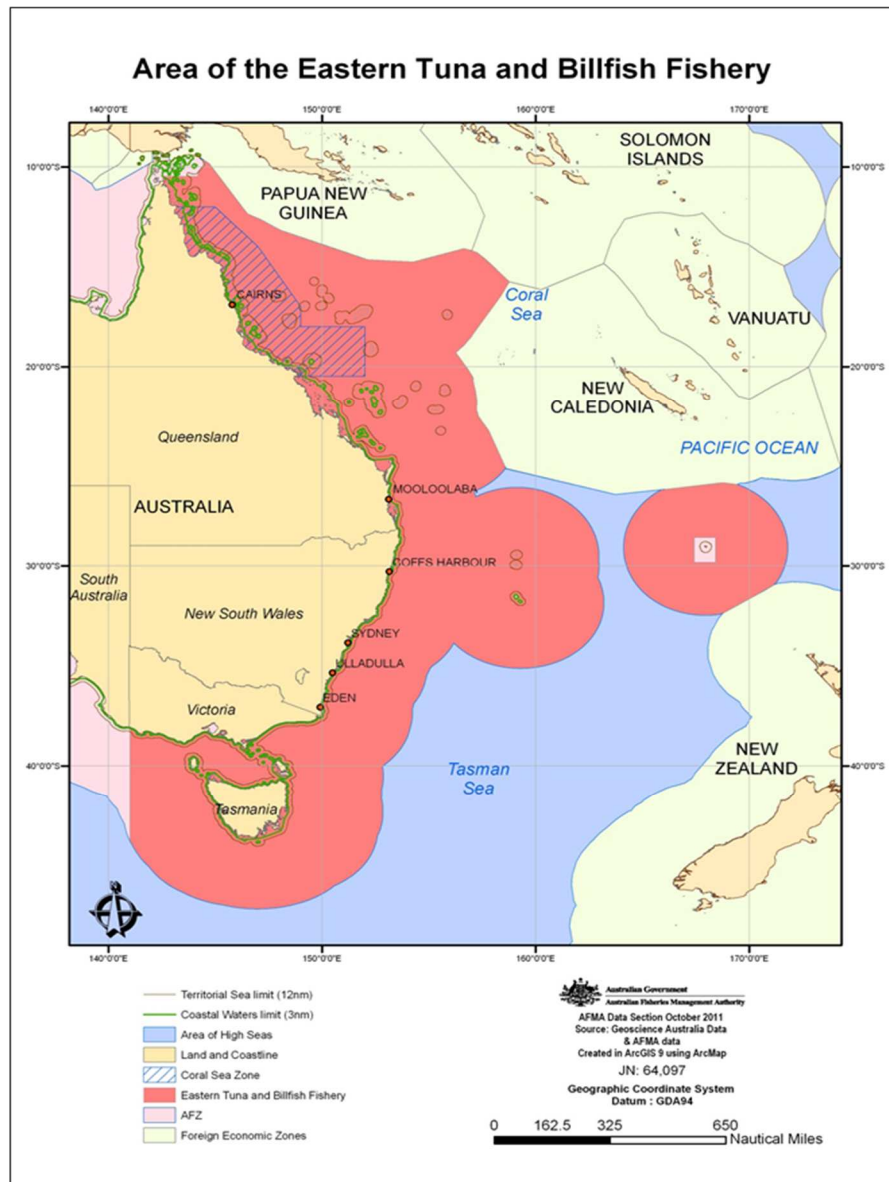


Figure 3.2 (a) Number of longline vessels operating in the ETBF each year and the associated effort (number of hooks deployed), and (b) Distribution of number of sets per trip (2007 to 2013).

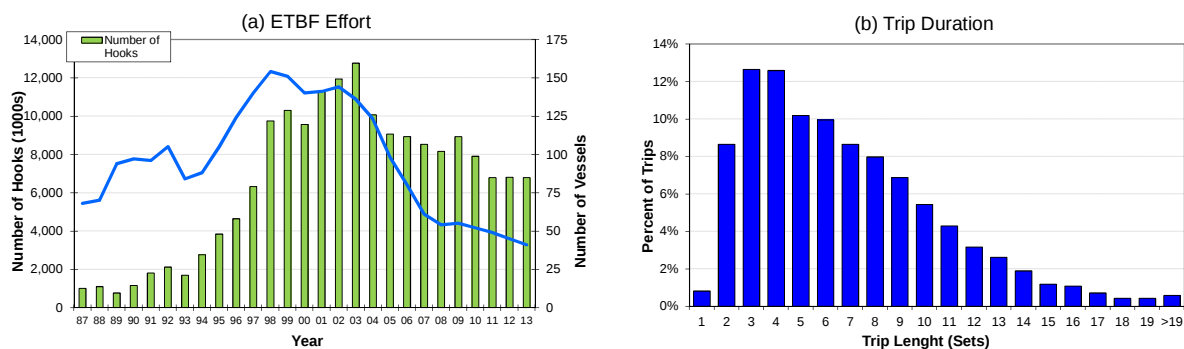


Figure 3.3: Example of completed AL06 logbook used in the ETBF indicating the catch and effort information collected for each longline set deployed.

Australian Fisheries Management Authority.
Box 7051
Canberra Mail Centre ACT 2610

Australian Pelagic Longline Daily Fishing Log – AL06

NOTE: DO NOT USE A SINGLE PAGE FOR MORE THAN ONE TRIP

Original Copy – Send to AFMA

Boat Name Cormorant		Dist. Symbol LFB963		Log No.	Page No.
Port Departed SYDNEY		Date Departed 25 / 6 / 07		NON-FISHING PERIOD I did not work between 19 / 6 / 07 and 24 / 6 / 07	
Port Returned ULLADALLA		Date Returned 27 / 6 / 07		Non-Fishing Codes (PLEASE CIRCLE) ☒ 1 – Bad Weather 2 – In Port 3 – Broken Down 4 – Steaming 6 – Searching 5 – Other Fishery (SPECIFY) 10 – Refit	

SHOT INFORMATION		Shot 1	Date	Shot 2	Date	Shot 3	Date
Target species		Yellowfin, Bigeye		Yellowfin, Bigeye			
Start set time (24h)		0300		0230			
Start set	Lat. (dd mm)	33	35	36	31		
Position	Long. (ddd mm)	151	42	151	55		
End set time (24h)		0610		0515			
End set	Lat. (dd mm)	35	19	36	25		
Position	Long. (ddd mm)	151	40	151	40		
Start Haul time (24h)		1500		1300			
Start Haul	Lat. (dd mm)	35	20	36	20		
Position	Long. (ddd mm)	151	41	151	42		
End Haul time (24h)		2200		1900			
End Haul	Lat. (dd mm)	33	36	36	30		
Position	Long. (ddd mm)	151	40	151	56		
Vessel shooting speed (kn)		7		7			
Mainline length/Hooks		30 nm/km 1000 hooks		25 nm/km 700 hooks			
Line shooter used (CIRCLE)		Yes ☒ No ☐		Yes ☒ No ☐			
Seabird mitigation measures used (CIRCLE) (see template)		TORI CHUTE ☒ THAW CAPS ☒ PSBL DYED ☒ NSET ☒ LWEL ☒ NAPP ☒		TORI CHUTE ☒ THAW CAPS ☒ PSBL DYED ☒ NSET ☒ LWEL ☒ NAPP ☒			
Targeted depth (in metres)		30 min 100 max		30 min 100 max			
No. hooks between bubbles		6		6			
No. of lightsticks used		500		300			
Bait type(s)/ source(s)/life status/weight(s) used for shot		SQO ☒ B ☒ S ☐ L ☒ D ☐ 50 Kg MAY ☒ B ☒ S ☐ L ☒ D ☐ 50 Kg		SQO ☒ B ☒ S ☐ L ☒ D ☐ 35 Kg PIL ☒ B ☒ S ☐ L ☒ D ☐ 45 Kg			

CATCH DETAILS	No. Fish Kept	Est. Processed Wt Kept (kg)	Form Code	No. Fish Discarded	DISCARD/RETAINED CODE	No. Fish Kept	Est. Processed Wt Kept (kg)	Form Code	No. Fish Discarded	DISCARD/RETAINED CODE	No. Fish Kept	Est. Processed Wt Kept (kg)	Form Code	No. Fish Discarded	DISCARD/RETAINED CODE
Yellowfin Tuna	11	350	GG	3	US	14	480	GG	1	DM					
Bigeye Tuna	4	150	GG			6	160	GG	4	TL					
Albacore Tuna	7	50	W			4	40	W							
Southern Bluefin Tuna															
Broadbill Swordfish						2	90	TR							
Striped Marlin	1	35	TR												
Shortbilled Spearfish															
Ray's Bream						3	10	GG							
Moonfish															
Rudderfish															
Oilfish/Escolar															
Dolphinfish															
Wahoo															
Lancetfish				4	UM										
Short Finned Mako Shark	1	80	TR						1	US					
Bronze Whaler Shark															
Dusky Whaler Shark															
Blue Shark									4	UM					
Oceanic Whitetip Shark															
Thresher shark									1	UM					
Yellowfin	4	60	GG		SD										
Bigeye	1	20	GG		SD										
Bigeye	3	25	W												

No Take Species	Species	Number Released		Number Released		Number Released	
		Alive	Dead	Alive	Dead	Alive	Dead
	Blue Marlin		1	1			
	Black Marlin						

Did you have an Observer on Board (circle) **(No)** / Yes Observer Trip ID

Please provide an estimate of the time taken to complete this form: **10** mins

Did you have an interaction with a Listed Marine or Threatened Species? (circle) Yes / **(No)**
Further details of all Listed Marine and Threatened Species interactions must be recorded on the Listed Marine and Threatened Species Form at the back of the logbook.

Concession holder or authorised agent - I certify that the information provided on this form is a true and accurate record.
Printed Name: **Tim Gardener**

Comments: **5 fish damaged by sharks in first shot but fish still retained**

Signature: **7 Gardener** Date: **27 / 6 / 07**

NOTE • If tagged fish / animals or banded birds are captured, please complete tag form at back of book and return to AFMA.

fishery, with 39 longline vessels and 7 minor line vessels active in 2015 though both the number of longline vessels and associated effort has varied over time (c.f. Figure 3.2a). During this year 5,324 longline sets were deployed setting a total of 8.25 million hooks. The number of hooks per set generally ranges between 300 and 3000 and the mean during 2015 was 1550. Catch by species (both number of fish and processed weight of retained fish together with the number of discarded fish) and effort are reported on a per set basis. A number of other details related to how the longline gear is deployed (e.g. set-time, location, bait-type, number of hooks between floats, number of light-sticks used) are also recorded on the logbook (c.f. Figure 3.3). In general, a single longline set is deployed each day during a fishing trip, though within a region off northern Queensland several short sets (< 500 hooks) can sometimes be deployed. Trips vary in length, with the number of sets per trip generally varying between one and 20 (c.f. Figure 3.2b).

The ETBF is managed by the Australian Fisheries Management Authority (AFMA) under the 2011 Management Plan. The five primary target species, Yellowfin Tuna (*Thunnus albacares*), Bigeye Tuna (*Thunnus obesus*), Albacore Tuna (*Thunnus alalunga*), Broadbill Swordfish (*Xiphias gladius*) and Striped Marlin (*Kajikia audax*), are each managed under output controls in the form of individual transferable quotas by setting a Total Allowable Commercial Catch (TACC) for each quota year (which starts on 1 March each year). TACCs were first introduced in the 2011/12 quota year and the ETBF harvest strategy (Davies *et al.* 2008) was utilised to determine a Recommended Biological Commercial Catch (RBCC) for each species to help inform the AFMA Commission in determining the respective TACCs. Since the 2014/15 quota year, however, the ETBF harvest strategy has been used only for the two billfish species (Campbell 2016b), while the TACCs for the three principal tuna species are set after taking into consideration current fishery indicators of stock status both within the WCPO and the ETBF (Campbell 2016c). This change was introduced since the majority of the catch taken for the three tropical tunas within the principal 'region of interest' to the ETBF is taken by fleets other than the ETBF (Campbell 2016d). This has the consequence that the successful management of these resources cannot be undertaken by Australia alone but will require a regional management approach.

While the majority of effort in the fishery is directed at the five primary species, a wide range of other secondary species (species of fish that may be taken in the fishery and retained) are reported in the fishery, including Mahi Mahi (Dolphinfish, *Coryphaena hippurus*), Wahoo (*Acanthocybium solandri*), Opah (Moonfish, *Lamprdae guttatus*) and various Oilfishes. However, apart from the catch of Southern Bluefin Tuna (*Thunnus maccoyii*, which is caught in the southern part of the ETBF and is managed by the Commission for the Conservation of Southern Bluefish Tuna) the catch of species other than the five quota species generally makes up less than 10% of the annual retained catch.

3.2 The Multispecies Problem

3.2.1 Basic Equations

The relation between catch rates (CPUE) and stock abundance is based on the catch equation which, as a first order approximation, relates the number of fish in the catch, C , fishing effort, E , and average fish population density, D , on the fishing grounds:

$$C = qED \quad (3.2.1)$$

where q is a fixed constant of proportionality known as the catchability coefficient and is related to the efficiency of the fishing gear. From this equation:

$$CPUE = \frac{C}{E} = qD = \frac{qN}{A} \quad (3.2.2)$$

where N is the number of fish on the fishing grounds and A is the spatial area of the fishing grounds. It follows that changes in CPUE are due either to changes in the stock density (or number of fish on the fishing grounds) or to changes in the catchability coefficient. If the changes in q can be accounted for, then the remaining changes in CPUE can be related to those in stock density. This is the basic idea underlying what is known as the standardisation of catch rates.

However, the concept of abundance needs some elaboration. Of particular importance is the related concept of availability. The following definitions were proposed by Marr (1951):

Abundance is the absolute number of individuals in a population. Availability is the degree (a percentage) to which a population is accessible to the efforts of a fishery. Apparent abundance is the abundance as affected by availability, or the absolute number of fish accessible to the fishery.

From these definitions, if B represents the true abundance and N measures the apparent abundance, then

$$N = aB \quad (3.2.3)$$

where a represents the availability or proportion of the total stock available to the fishery. Substituting into Eqn. (3.2.2) and rearranging gives:

$$B = \frac{N}{a} = \frac{A \cdot CPUE}{aq} \quad (3.2.4a)$$

and

$$CPUE = \frac{aqB}{A} \quad (3.2.4b)$$

From this equation it is seen that the relationship between CPUE and the true abundance of fish within a given spatial region is influenced by both the availability of the fish to the fishing gear (a) and the efficiency of the fishing gear (q).

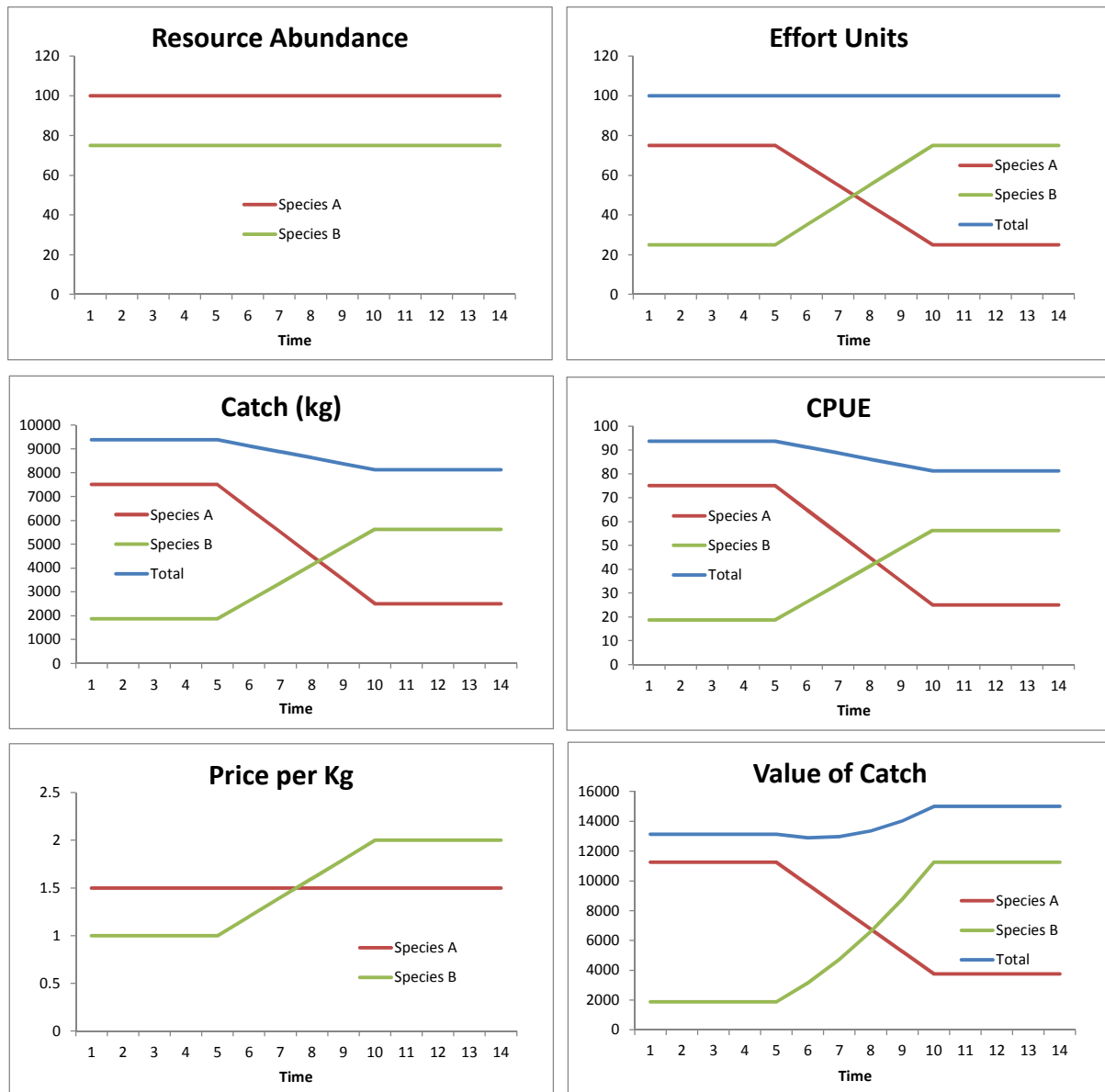
3.2.2 Example

The discussion above concerning the need to account for the factors which influence both the availability of fish to the fishing gear and the catchability of the fishing gear is just as relevant for standardizing CPUE in a single species fishery as for standardizing CPUE in a multispecies fishery. However, an additional problem arises in a multispecies fishery when amongst the suite of principle species caught there is a change in the specific species targeted.

For example, in the EBTF there are five principle catch species, though on many individual fishing operations only one of these species may be the main target species with the others considered by-product species. Furthermore, there are usually distinct differences in the manner in which the longline is deployed when targeting these different species. When there is a change in target between any suites of species, there is also an associated change in the 'effective' effort directed at these species. Such a shift in species targeting has been described by Carruthers *et al.* (2011) who noted how the blast freezer technologies adopted in the 1950s and 1960s improved the viability of long distance Japanese sashimi fisheries. Consequently,

Table 3.1: The time-series of effort, catch ,CPUE and value of catch for a hypothetical fishery described in the text.

	Density Units		Effort Units			Catch=q*Effort*Density			CPUE = Catch/Effort			Price/Unit of Catch		Value of Catch		
	q=1															
year	Bio(SpA)	Bio(SpB)	Effort	%SpA	%SPB	cat(SpA)	cat(SpB)	cat(tot))	cpe(SpA)	cpe(SpB)	cpe(tot)	\$(SpA)	\$(SpB)	Val(SpA)	Val(SpB)	Val(tot)
1	100	75	100	75	25	7500	1875	9375	75	18.75	93.75	1.5	1	11250	1875	13125
2	100	75	100	75	25	7500	1875	9375	75	18.75	93.75	1.5	1	11250	1875	13125
3	100	75	100	75	25	7500	1875	9375	75	18.75	93.75	1.5	1	11250	1875	13125
4	100	75	100	75	25	7500	1875	9375	75	18.75	93.75	1.5	1	11250	1875	13125
5	100	75	100	75	25	7500	1875	9375	75	18.75	93.75	1.5	1	11250	1875	13125
6	100	75	100	65	35	6500	2625	9125	65	26.25	91.25	1.5	1.2	9750	3150	12900
7	100	75	100	55	45	5500	3375	8875	55	33.75	88.75	1.5	1.4	8250	4725	12975
8	100	75	100	45	55	4500	4125	8625	45	41.25	86.25	1.5	1.6	6750	6600	13350
9	100	75	100	35	65	3500	4875	8375	35	48.75	83.75	1.5	1.8	5250	8775	14025
10	100	75	100	25	75	2500	5625	8125	25	56.25	81.25	1.5	2	3750	11250	15000
11	100	75	100	25	75	2500	5625	8125	25	56.25	81.25	1.5	2	3750	11250	15000
12	100	75	100	25	75	2500	5625	8125	25	56.25	81.25	1.5	2	3750	11250	15000
13	100	75	100	25	75	2500	5625	8125	25	56.25	81.25	1.5	2	3750	11250	15000
14	100	75	100	25	75	2500	5625	8125	25	56.25	81.25	1.5	2	3750	11250	15000

Figure 3.4: The time-series of fishery indicators for a hypothetical fishery described in the text.


the Japanese longline tuna fleets of the Atlantic and Pacific expanded to higher latitudes to target species of higher market price (such as Bigeye and Bluefin Tuna) and away from tropical waters inhabited by species such as Yellowfin and Skipjack Tuna.

A simple example helps illustrate this situation. Consider the case where there are two species being fished by a fishery. The abundance of each species remains constant over time and the total effort deployed in the fishery is also constant. At the start of the fishery the price paid per kilogram is greater for species A but at some period the price per kilogram paid for species B increases to be greater than species A. This change results in a decrease in the percentage of effort targeted on species A and an increase in the percentage of effort targeted on species B (in this instance due to changes in the manner that the fishing gear is deployed).

Using the equations described in the previous Section, a time-series of effort, catch, CPUE and value of the catch in such a hypothetical fishery is provided in Table 3.1 and displayed in Figure 3.4. While the total catch decreases during this period of change the total value of the catch

increases, resulting in a more profitable fishery (given that total effort remains unchanged). However, if one uses nominal CPUE as a proxy for resource abundance over time, then one would conclude that the abundance of species A has declined and the abundance of species B has increased. This is an incorrect inference due to the fact that the changes in the relative catchability of the fishing gear for the two species have not been taken into account.

To analyse this situation further, let E_A be the effort targeted at species A, E_B be the effort targeted at species B, $q_{A,A}$ be the catchability of species A when targeting species A and let $q_{A,B}$ be the catchability of species A when targeting species B. Then from Eqn. (3.2.1) the total catch of species A is given by:

$$C_A = (q_{A,A}E_A + q_{A,B}E_B)D_A \quad (3.2.5)$$

from which we obtain:

$$D_A = \frac{C_A}{q_{A,A}E_A} \cdot \frac{1}{\left(1 + \frac{q_{A,B}E_B}{q_{A,A}E_A}\right)} \quad (3.2.6)$$

In general, the estimate of the density of species A will be biased by the second factor shown in Eqn. (3.2.6), however, there will be situations where this does not pose a problem. For example, where $q_{A,B} \approx 0$ (i.e. species A is not caught to any extent when targeting species B) then the estimate of the density of species A will be relatively unbiased. Also, when the ratio E_B/E_A is constant over time then there is no bias in the relative measure of D_A over this period as the bias will be constant. However it is unlikely that this ratio will be invariant over long periods of time because many of the factors influencing the behaviour and preferences of fishers may change.

In the above example, it is assumed that the fisher has some ability to target the different species in the fishery as the need or desire arises, i.e. target fishing on a species is a deliberate act that can be predicted by fishers in advance and controlled. Nevertheless, in many multispecies fisheries there are technological interactions where fishing effort directed towards one species will normally result in a mixed catch of fish that may include many other species, i.e. it is not possible to target and catch only one species to the exclusion of others. While fishers can usually 'target' effort to some degree through fishing different areas and depths, seasons, times of day and by modifying gear, it is the degree to which fishers can target that is the issue. In such situations, the allocation of effort to particular species in a multispecies fishery (i.e. targeting) is not trivial (Klaer and Smith 2012).

4. Objectives

The project had the following five objectives:

1. Identify the factors likely to influence CPUE in multispecies pelagic longline fisheries and review the data requirements and data availability so that these factors can be used for standardising CPUE in these fisheries.
2. Review all methods (both those currently used and any other novel methods) which may be used for standardising CPUE in multispecies fisheries.
3. Based on experiences in other relevant research, and the outcomes of objectives 1 and 2 identify, develop and compare the most appropriate methods for standardising CPUE for pelagic longline fisheries.
4. Use simulated catch and effort data to test the potential of each method to adequately account for the influence of factors influencing CPUE and accurately reflect the underlying resource abundance.
5. Investigate the sensitivity of the outcomes of the ETBF harvest strategy on the adoption of the candidate methods for standardising CPUE within the ETBF.

5. Methods

5.1 Overview (and Report Structure)

An overview of the main approaches used to address each of the five project objectives is outline here. Where indicated additional details are outlined in the remainder of this chapter.

1) To adequately standardise CPUE one needs an understanding of the relationship between how the fishing gear is deployed and the subsequent catch. Factors influencing this relationship include both the availability of fish to the fishing gear and the catchability of the fishing gear. Availability will be influenced by the environmental and oceanographic conditions prevailing at the time of the fishing operation, while catchability will be influenced by the types of fishing gear used and the manner in which it is deployed. In a multispecies fishery there is particular interest in understanding the differences in fishing tactics utilised when targeting different species, as this information can be used to identify the 'effective' effort directed at any single species. For the ETBF the relationships between fishing practices and catch, together with changes in the seasonal availability and distribution of co-related species (i.e. species caught within the same set) in the fishery, was undertaken using the catch and effort data recorded in logbooks and by on-vessel observers. Results of these investigations are outlined in Sections 6.1 to 6.4.

2) A literature review was conducted of the merits of currently available methods for standardising CPUE appropriate for multispecies pelagic longline fisheries, with an emphasis on those methods not currently used within the ETBF assessments. The data requirements, statistical assumptions, and pros and cons for implementing each were identified. For example, habitat-based methods require specific information on species habitat preferences and the depths of the fishing gears for which direct information within the ETBF is limited in both space and time. Results of these review are outlined in Section 5.2.

3) Taking account of both the review of methods and the investigations into the relationships between fishing practices and catch in the ETBF outlined in (1) and (2) above, a number of novel methods for standardising CPUE in the ETBF were developed. These included refinements to existing methods, and where possible the development of new methods. Refinements of existing methods included the incorporation of additional explanatory factors in GLMs to better model the range of targeting strategies used in the ETBF, and the identification of more appropriate temporal and spatial stratification of the data (i.e. *Area* effects). Consideration was also given to segmenting the analyses on a seasonal basis (i.e. monthly or b-monthly) to take account of seasonal changes in the availability of co-related species. Finally, a novel geo-statistical model was also developed. Following a summary of the current method used to standardise CPUE in the ETBF (provided in Section 5.3), details of the alternative methods developed are outlined in Sections 5.4 and 5.5.

4) To evaluate the performance of the methods identified in (3) above, two simulation models were developed for generating set-by-set catch and effort data within a longline fishery context. These models were developed to incorporate a range of hypotheses for simulating how CPUE can be influenced by the operational factors included in the simulation, and were used to generate candidate catch and effort data to test and compare the outcomes and utility of candidate methods for standardising CPUE. In particular, to ascertain potential biases inherent in the use of each method, the resulting time-series of standardised CPUE were compared with

the known stock trends used to generate the simulated data. The details of the simulation models developed are outlined in Section 5.6, while a listing of the standardising models used in the comparative analyses is provided in Section 5.7. The results of comparing the performance of each candidate method for standardising CPUE are provided in Section 6.4 and some additional comparison are outlined in Section 6.5.

5) Using both the historically available catch and effort data for the ETBF and the simulated data for this fishery, annual indices of abundance for each of the main target species were calculated using candidate methods for standardising CPUE. The results were then input into the ETBF harvest strategy to determine the sensitivity of the calculated RBCCs to the range of standardising methods used. Results are provided in Section 6.6. Further implications for the ETBF harvest strategy in a stock assessment context is provided in Section 6.7.

5.2 Review of methods to standardise CPUE in multispecies fisheries

Due to the importance of CPUE in many stock assessments and the assumption that CPUE is proportional to abundance, it is important that any other factors that may influence CPUE are accounted for and removed from the index. The process of reducing the influence of these factors on CPUE is commonly referred to as standardizing the CPUE. While a short review of methods used to standardise CPUE is provided in this Section, the emphasis is on those methods particularly suited to multispecies fisheries.

In an attempt to standardize CPUE time series there has been a steady publication of papers on statistical techniques which may be considered appropriate for such purposes (see review by Maunder and Punt 2004; Tascheri *et al.* 2010; Lynch *et al.* 2012). Early approaches adjusted nominal effort to account for the differences in relative vessel efficiency (Beverton and Parrish 1956; Gulland 1956; Robson 1966), while more recently the advent of high speed computing and the use of more advanced statistical methods has allowed the inclusion of more factors in the standardisation process and has helped to overcome some of the more obvious limitations of the earlier methods. These techniques cover a range of methods including Generalised Linear Model (GLMs), Generalised Additive Models (GAMs) and Generalised Linear Mixed Models (GLMMs) and Regression Trees. There has also been considerable discussion on which of the many error distributions to choose from (e.g. Normal, Gamma, Poisson, Negative Binomial and Tweedie distributions) may be the most appropriate to use with these methods.

Generalised Linear Models (GLM; e.g. Allen and Punsly 1984; McCullagh and Nelder 1989) are the most common method used to standardize CPUE. The CPUE is predicted as a linear combination of explanatory variables. Variables can be either categorical or continuous. Often, continuous variables are grouped into intervals and included as categorical variables. This is done, for example, to provide indicators for intrinsically nonlinear relationships, to reduce problems encountered with large numbers of zero observation strata (see also delta lognormal method below), or to create strata which reflect combinations of continuous variables with certain characteristics, that taken together serve as classifications that have no meaning on an ordinal scale. Higher order terms and intrinsically linear terms can be included as continuous variables in GLM models. For example, if the relationship is assumed to be domed shaped, the CPUE could be related to the square of the explanatory variable. Interaction terms can also be added to the model to allow for interactions among explanatory variables when appropriate.

The main objective of the analysis is to estimate a year effect. The year effect (included in the GLM as a categorical variable) is used to represent the annual relative levels of abundance and is used as the relative index of abundance to include in the stock assessment. Interactions with the year effect invalidates this interpretation of the year effect as an index of abundance and for this reason, most analyses do not consider interaction terms for the year effect. However, when there is high inter-annual variability in the spatial or temporal distribution of a resource between years the inclusion of year-interaction terms may be appropriate and in such situations the annual index of abundance needs to be constructed from the parameter estimates from the GLM (see Campbell 2015).

Many fishery data sets (especially when the analysis is undertaken at the operational level, i.e. not aggregated) have a large number of unsuccessful units of effort (i.e. sets or strata with positive effort and zero catch), and this can cause bias in the analysis. Historically, standard GLM analyses based on a log-transformation of the data required a non-zero CPUE and it was common practice to combine strata to eliminate zero catch observations or to add a constant to the data, so that CPUE is always greater than zero. Both of these approaches have disadvantages. When strata are combined, it is possible that important information contained in explanatory variables on levels not related to the combined strata may be compromised. This may reduce the performance of the GLM or require the development of alternate strata for certain explanatory variables in order to conduct the analysis. In the second approach, adding a constant may cause some bias in the estimated year effect. The delta-lognormal method (Pennington 1983, 1996; Lo *et al.* 1992) was developed to overcome these problems in a GLM framework. This method models the zero catches separately and then models the positive catches using a GLM. The model for the zeros and the GLM are then combined to generate an index of abundance.

GLMs are convenient because they have a long history, they are well understood, and they have accepted methods to choose factors, or variables, in a model. Unfortunately, they are limited in their functional form to linear relationships. While these relationships can be made more complex, such as by adding higher order terms or by adding interaction terms, in many situations nonlinear relationships may better describe the relationships between CPUE and explanatory variables. A number of techniques have been explored in this regard including:

- General Additive Models (GAMs) - see Bigelow *et al.* (1999) for an example of use of GAMs in the Pacific to standardize CPUE of Broadbill Swordfish (*Xiphias gladius*) and Blue Shark (*Prionace glauca*),
- Neural Networks, which uses the data to estimate the structure of the non-linear relationship between CPUE and the explanatory variables – see Maunder and Hinton (2006), and
- Regression Trees, which like neural networks are suited to detecting and extracting important and complex interactions of the explanatory variables - see Watters and Deriso (2000).

While nonlinear models are more general in their functional forms when compared to GLMs, they do not in general use analytical reasoning to define the functional form of the relationship between the explanatory variables and CPUE. For example, parameter estimates from fitted GAMs are less readily interpretable than those from fitted GLMs. An alternative to these statistical approaches to standardising CPUE is to use our knowledge of the distribution of the target species and the fishing gears more directly. The approach, known as the Habitat-Based-Standardisation (HBS) developed by Hinton and Nakano (1996), presents a method of standardising CPUE which combines information on the spatial and depth distributions of the

target species (using information on habitat preference and mapping of this habitat provided by oceanographic models) with information on the depths fished by longline hooks. The basic premise is that if a hook is fished in an environment that is preferred by the species, then it has a higher probability of capturing that species. This is particularly important, for example, when standardizing effort of longline gear targeting tuna, because the depth of the gear has increased over time as fishermen targeted Bigeye Tuna, which are generally found at deeper depths in the water column.

This method was first applied to Pacific Blue Marlin (*Makaira nigricans*) before being applied to Bigeye Tuna in the Western Central Pacific Ocean (WCPO) by Hampton *et al.* (1998). The method has subsequently been further developed within a statistical framework (and is known as the stat-HBS method, Maunders *et al.* 2006a) and for a period following 2000 was routinely applied to both Bigeye Tuna and Yellowfin Tuna within the context of the stock assessments undertaken for these species within the WCPO (see Langley *et al.* 2005). The HBS methods require detailed information on the depths fished by hooks together with the distribution of the habitat of the target species, and poor performance of the habitat model has been attributed to (i) problems in estimating hook depth, (ii) fine-scale variations in environmental conditions, and (iii) incomplete knowledge of habitat preferences (Ward and Meyers 2006; Lynch *et al.* 2012).

While most of these methods have been developed for single-species analyses, a number of papers have attempted to develop approaches which take into account features of multispecies targeting. These methods can be broadly divided into the following two categories:

- 1) Sub-setting;
- 2) Covariates for other species.
- 3) Spatial-temporal models.

An important consideration for the standardization of multispecies CPUE data is that the choice of fishing tactic allocates effort toward a particular target species or species complex and away from other species. The term 'fishing tactic' is defined as a sequence of choices of fine-scale fishing strategies (e.g., gear, time of day, bearing, small-scale movements) made by the skipper during a fishing trip. A short review of each category is given here.

5.2.1 Sub-setting

The main idea associated with these methods is to select from amongst all the catch and effort records for a fishery a sub-set of records which are deemed informative about the species under consideration.

Several variations on this approach have been illustrated in the fisheries literature:

1. Categorising and allocating shots / trips based on species compositions, (e.g. Biseau 1998; Klaer and Smith 2012). In the latter, the rule used assumed that fishers target according to the value of the species in the catch rather than weight, and that targeting is informed by prior knowledge of where and when certain species may be caught. In particular, in order to assign a target to all individual trawls with a catch the following rule was used: The target species was deemed to be the species with the greatest portion of the total catch value in a 0.51 subdivision 50 m depth stratum month during the same time of the day.
2. Identifying 'indicative' vessels based on vessel characteristics and catch history. For example, Punt *et al.* (2000) applied this to School and Gummy Shark, selecting vessels

for inclusion in the analyses based on the number of years with a catch, median annual total catch, and median annual catch for the species of interest.

3. When fishing location is included in the catch and effort records, it should be possible to restrict the analysis to those data for only those locations known to be habitat associated with the species of interest. However, when location is not included it may be possible to allocate shots based on the species compositions from fishing trips, to infer whether the fishing occurred in habitat appropriate for use in CPUE calculations for the species of interest. The idea is that the species composition (excluding the species under assessment) from a fishing trip provides information that can be used to make predictions as to whether the fishing trip included at least some effort expended in the target species' habitat. Stephens and MacCall (2004) applied this method to groundfish caught by recreational fishermen on the west coast of the United States. They used a logistic regression that uses the presence or absence of other common species to estimate the probability that the target species would be encountered. Selection of a critical value allows the catch and effort data to be divided into the records in target and non-target habitat.
4. Using auxiliary survey data to determine the percentage of a species within the logbook recorded catch using GLMs from which one can define a cut-off for identifying the appropriate sets for use in any subsequent single species CPUE analyses. Venables and Dichmont (2004) developed this approach for two species of tiger prawns in the fishery off northern Australia.

While these methods are relatively easy to use and explain, and appear to work well when there are very different and well-defined metrics for categorising the data, there is often an *ad hoc* element (e.g. in the choice of cut-offs) in manner in which the data is subsetted. Those methods which make use of physical variables would appear to be preferred, as these can be used to infer distinct physical habitats. On the other hand, methods which are based on the use of species compositions means that the results may be susceptible to bias due to trends in other species and possibly due to changes in the species recorded over time.

5.2.2 Covariate Methods

The main idea associated with these methods is to categorise the catch and effort data using a number of covariates recorded on logbooks. Again, several variations on this approach have been proposed in the fisheries literature:

1. Use of the catch rates of alternative target or bycatch species as covariates to correct for the effort directed away from the target species or species under consideration (Glazer and Butterworth 2002; Maunder and Punt 2004; Su *et al.* 2008). Although the catch rates of alternative species does not hold direct information about the magnitude of the catch, it is arguably of concern that the information contained in the predictor variables derived from these covariates is not entirely independent from the response CPUE and may have unpredictable impacts on the standardized CPUE trends. The following example illustrates this problem.

Let C_t and C_o be the catch of the target and another species (respectively) and let E_t and E_o be the effort directed at each species such that total effort $E = E_t + E_o$. The catch rate of the target species can be then be expressed as follows:

$$CPUE_t = \frac{C_t}{E} = \frac{C_t}{E_t} * \frac{E_t}{E_t + E_o} \approx I_t * \left(1 - \frac{E_o}{E_t} \right) \quad \text{if } E_o \ll E_t$$

where $I_t = C_t/E_t$ is the abundance index of the target species. Similarly,

$$CPUE_o = \frac{C_o}{E} = \frac{C_o}{E_o} * \frac{E_o}{E_t + E_o} \approx I_o * \left(\frac{E_o}{E_t} \right) \quad \text{if } E_o \ll E_t$$

Therefore, solving for E_o/E_t and substituting into the first expression gives:

$$CPUE_t \approx I_t * \left(1 - \frac{CPUE_o}{I_o} \right)$$

As $CPUE_o/I_o = E_o/E \ll 1$ and using the approximation $\ln(1+x) \approx x$ when x is small, then:

$$\ln(CPUE_t) \approx \ln(I_t) - \frac{CPUE_o}{I_o}$$

So, if we use a linear model for log catch rate of the species of interest with the catch rate of other species, as an explanatory variable, the associated implicit assumptions are: (i) the true abundance of the other species (as measured by the related index, I_o) has no temporal trend, and (ii) the proportion of the total effort targeted at the other species is very small.

For this reason, it is important that the alternative species should not co-occur with the target species. For example, if two species were to co-occur in the catches and would be fished down simultaneously, the use of the catch rate of the one species as a negative predictor of the CPUE of the other may result in an erroneous removal of the underlying year-effect for the species of interest. An additional challenge in situations where a large number of species are caught by the fishery is the objective selection of species-specific catch rates to be included as covariates in the standardization model.

2. Applying the assumption that distinct fishing strategies will result in distinctive suites of species in the catch, one can use well-known clustering techniques to categorise catch records into groups with similar catch compositions to identify those records which pertain to supposedly distinct targeting practices. The identified clusters are assumed to be representative of fishing tactics, which may be treated as categorical variables in the standardization model to adjust for differences in catchability associated with each cluster. This method was applied by He *et al.* (1997) to the Hawaii-longline fishery. Cluster analysis is useful in segregating dissimilar types of fishing effort particularly when aspects of fishing strategies remain unknown. After applying the cluster analysis, any CPUE index could be further improved using GLM techniques which incorporate additional sources of variability (e.g. gear configuration, area, time, and environment). He *et al.* (1997) recommended including data from several categories of set types if the CPUE trends for that species are similar when calculated for each type. When the CPUE time series pattern is uncorrelated among set types, it is not recommended that a combined data set be used to describe the dynamics of that resource unless a GLM analysis accounts for those differences. Clustering sets based on species proportions in the catch could potentially cause biases in cluster-based CPUE indices if, for example, fishing intensified and the abundance of a particular target species declined as the number of sets targeting that species increased but with a low catch might not be included in the cluster for that species. As a result a real decline in CPUE for sets targeting that species might be dampened or obscured. However, some

preliminary simulation testing by He *et al.* (1997) did not find this to be a major problem.

3. The 'Direct Principal Component' (DPC) procedure recently proposed by Winker *et al.* (2013, 2014) uses continuous principal component scores (PCs), derived from a Principal Component Analysis (PCA) of the catch composition data, as nonlinear predictor variables in a General Additive Model to adjust for the effect of temporal variations in fishing tactics. The DPC procedure is based on the common assumption that information on the direction and extent of targeted effort can be found in the species composition of the catch. PCA represents one of the most commonly used approaches to describe patterns of variations in multivariate data-sets. The idea is that meaningful sources of variation are retained in the first few PC-axes (nontrivial PCs). One of the difficulties is to select the relevant number of nontrivial PCs, which in this case represent meaningful separations of fishing tactics. Nevertheless, on initial testing and application the DPC standardization procedure has been shown to be a reliable method for removing the effects of targeting on multispecies CPUE. An obstacle to its immediate widespread application could be the onerous data requirement, in the form of detailed records of catches for all or the majority of species at the trip or shot level.

5.2.3 Spatial-temporal Methods

Recently Thorson *et al.* (2016) have suggested a new statistical approach for constructing abundance indices in a multispecies fisheries. Noting that fishers can affect the assortment of species caught in a multispecies fishery by modifying the location, timing and gear characteristics of their fishery activity, they have proposed a method for simultaneously estimating fishing tactics and relative fish abundance when standardizing fishery dependent catch rate data. The proposed 'spatial dynamic factor analysis' (SDFA) model estimates spatial and temporal variation in abundance for multiple species caught in a multispecies fishery by decomposing covariation in multispecies catch rates into components representing spatial variation and fishing behaviour. This is achieved by decomposing the catch equation, $C = qED$, for each fishing operation into components representing spatiotemporal variation in fish density (D), small-scale tactics that allocate fishing effort (E), and measurement variables affecting catchability (q), as well as residual variation. This decomposition implies that a spatiotemporal model of fish density can account for spatial variation in density (and therefore control for changes in the spatial allocation of fishing effort) while filtering out covariation in model residuals (as caused by fine-scale fisher tactics e.g. daily timing of fishing activity). Despite the promise of this new method at present, spatially stratified models remain more common than spatiotemporal models when analysing fishery CPUE data.

The above summarises the main methods published in the literature for standardising multispecies CPUE data. In the following Sections we outline the main methods which were selected for further investigation and development, but first we outline the current method used to standardise CPUE within the ETBF.

5.3 Current method used for standardising CPUE in the ETBF

Standardised CPUEs are principal inputs to the ETBF harvest strategy (used to determine RBCCs) and the set of fishery indicators (including outputs from the regional stock assessments, catch levels in the south-west Pacific, and catch, CPUE and size-based indicators for the ETBF) used for the three tropical tunas, both of which are used to determine TACCs.

The method used for standardising CPUE is fully described in Campbell (2016a) and is summarised here. This method will be used as a base-line against which the outcomes of alternative methods for standardising CPUE will be assessed.

5.3.1 Standardisation Method

The current method used to standard the catch rates in the ETBF is based on Generalised Linear Models (GLM) and is similar to approaches used in many fisheries around the world (Campbell 2004, 2015). Due to the inflated number of zero catch observations it is also considered best practice to standardise the CPUE data following a two stage process: one stage being concerned with the pattern of occurrence of positive catches, and the other stage with the mean size of the positive catch rates. The GLM approach assumes that both the probability of a positive catch and the size of a positive catch rate can be modelled as linear combinations of the factors. Once this is done, the means from the two distributions can be combined to give an overall mean abundance index.

A small example helps illustrate this approach. Consider a season for which there are n catch rate observations, C_i . The average catch rate can be expressed as follows:

$$\mu = \frac{1}{n} \sum_{i=1}^n C_i = \frac{1}{n_s + n_F} \sum_{i=1}^{n_s} C_i = \frac{n_s}{n_s + n_F} \frac{1}{n_s} \sum_{i=1}^{n_s} C_i = p_s \mu_s$$

where n_s is the number of positive or successful catch rates obtained ($C_i > 0$), n_F is the number of zero or failed catches ($C_i = 0$), p_s is the proportion of positive catches and μ_s is the average of the positive catch rates. This result shows that the overall mean catch rate can be expressed as the combination of the parameters from the distributions used to model the probability of a successful catch, and that used to model the non-zero catch rates. This approach was used in the estimation of egg production based on plankton surveys (Pennington 1983; Pennington and Berrien 1984) and for estimating indices of fish abundance based on aerial spotter surveys (Lo *et al.* 1992) and has since become widely adopted (see Maunder and Punt 2004).

Stage 1: Prob(positive catch)

The Binominal distribution is used to model the probability of a non-zero catch where we model each observation as either a success ($C_i > 0$) or a failure ($C_i = 0$), with the probability of either expressed as follows:

$$\Pr(C_i > 0) = p_s \quad \text{and} \quad \Pr(C_i = 0) = 1 - p_s$$

Associated with each observation is a vector of covariates or explanatory variables X_j thought likely to influence the probability of a positive catch. For the ETBF analysis this includes a combination of both gear effects (e.g. start-time of set, bait-type, number of hooks-per-float, mainline-length) and environmental effects (e.g. sea-temperature, mixed-layer-depth, southern-oscillation index); a full list is given in Campbell (2016). Furthermore, we assume that the dependence of p_s occurs through a linear combination $\eta = \sum \beta_j X_j$ of the explanatory variables. In order to ensure that $0 \leq p_s \leq 1$ we use the logit link function that takes the following form:

$$\eta = \log \left(\frac{p_s}{1 - p_s} \right)$$

The inverse of this relation gives the probability of a positive sighting as a function of the explanatory variables:

$$p_s = \frac{e^\eta}{1+e^\eta} = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots)}$$

The following model is then fitted to the data:

$$\text{MODEL } p_s = \text{intercept} + f(\text{year}, \text{qtr}, \text{area}) + \sum \beta_j X_j + \text{Logit}\tilde{H} \\ / \text{dist=binomial link=logit}$$

where the following forms of the function $f()$ are fitted as two separate models:

$$\text{Model 1: } f(\text{year}, \text{qtr}, \text{area}) = \text{year} * \text{qtr} + \text{qtr} * \text{area}$$

$$\text{Model 2 } f(\text{year}, \text{qtr}, \text{area}) = \text{year} * \text{qtr} * \text{area}$$

and * represents an interaction between the variables shown. As the probability of catching a fish is likely to be related to the number of hooks (H) deployed, the variable $\text{Logit}\tilde{H} = \log(h/(1-h))$ where $h=H/M$ and $M > \text{maximum}(H)$ was also fitted as an additional fitted effect (Campbell 2015). After fitting the above model to the data the standardised probability for a positive catch, p_s , was then calculated for each spatio-temporal strata (year, quarter and area) against a standard set of model factors.

Stage 2: Mean Size of Positive Catch Rate

A separate model is fitted to the distribution of positive catch rates, μ_s . For this purpose a log-Gamma model is adopted, such that the μ_s was assumed to have a gamma distribution with a log link to the vector of covariates or explanatory variables X_j . The data fitted to the model are limited to those observations having a positive catch. As before, the following model is fitted to the data:

$$\text{MODEL } \mu_s = \text{intercept} + f(\text{year}, \text{qtr}, \text{area}) + \sum \beta_j X_j / \text{dist=gamma link=log}$$

where again the two functional forms of $f()$ are as described previously. A standardised mean positive catch rate, μ_s , is then calculated for each spatio-temporal strata (year, quarter and area) against a standard set of model factors.

An alternative model using the discrete negative binomial distribution, a log link and a log(effort) offset can be fitted to the catch. This distribution also provides a more general form of the assumed variance function ($\mu + k\mu^2$).

Fitted Variables

A range of variables collected from the ETBF are used in the GLMs as explanatory variables. These variables, listed in Table 5.1, are divided into the following four groups:

- 1) Statistical effects – these attempt to account for differences in availability of the fish due to differences in the spatial and temporal distribution of the resource and changes in the size of the resource each year. Variables include Year, Quarter and Area.
- 2) Fishing Practice Effects – these attempt to account for differences in the effectiveness of the set due to differences in the manner that the longline was deployed. Variables include Start-Time, Bait-Type, Use of Light-sticks, Length of Mainline, Distance between Floats, Number of Hooks-between-Floats, and Number of Hooks-per-Kilometer of deployed mainline.

Table 5.1. Listing of explanatory variables, and category definitions, fitted to GLM used to standardise CPUE in the ETBF. The number of observations within each category for the Swordfish analysis undertaken in 2016 is also shown

No.	Standardising Variable	Category Level	Category Definition	Number of Observations
1. Statistical Effects				
1	Year	1 to 19	1997 to 2015	3,519 to 11,815
2	Quarter of Year	1 2 3 4	Jan-Mar Apr-Jun Jul-Sep Oct-Dec	30,119 36,209 38,634 32,407
3	Region fished, Area	1 2 3 4 5 6 7	Species specific Example: Refer to Figure 5.22	17,334 45,022 11,347 24,150 20,715 11,081 7,720
2. Fishing Strategy Effects				
4	Start Time of Set	1 2 3 4 5 6	before 4am 4am to 8am 8am to noon noon to 4pm 4pm to 8pm 8pm to midnight	14,318 17,028 18,191 18,141 51,031 18,660
5	Bait Type Used	1 2 3 4 5 6 7 8 9	squid, dead yellowtail scad, alive pilchard, dead other, dead other, alive mixed species, dead mixed species, alive mixed species, mixed life-status all other categories	70,594 12,478 14,000 2,309 1,324 19,107 1,594 13,648 2,315
6	Hooks-per-Float	1 2 3 4 5 6 7 8 9 10	HPB ≤ 5 HPB=6 HPB=7 HPB=8 HPB=9 HPB between 10 and 11 HPB between 12 and 14 HPB between 15 and 19 HPB between 20 and 29 HPB between 30 and 40	4,579 13,233 8,848 37,511 6,311 39,300 13,186 5,883 3,324 5,194
7	Percentage of Hooks with Light-sticks	1 2 3 4 5 6 7	0% 1 to 19 % 20 to 39 % 40 to 59 % 60 to 79 % 80 to 99 % 100%	34,931 10,640 17,327 37,997 6,275 7,898 22,301
8	Length of Mainline	1 2 3 4 5 6 7	<20km between 20km and 30km between 30km and 40km between 40km and 50km between 50km and 60km between 60km and 70km >70km	4,730 13,436 20,746 28,248 35,577 21,038 13,594
9	Distance between Floats	1 2 3 4 5 6	<300m between 300m and 400m between 400m and 500m between 500m and 600m between 600m and 750m >750m	19,628 36,999 36,895 24,913 10,159 8,775
10	Number of Hooks per Kilometer of Deployed Mainline	1 2 3 4 5 6	<15 between 15 and 20 between 20 and 25 between 25 and 30 between 30 and 35 >35	8,584 43,829 40,849 19,778 15,317 9,012

Table 5.1 (cont'd). Listing of explanatory variables, and category definitions, fitted to GLM used to standardise CPUE in the ETBF. The number of observations within each category for the Swordfish analysis undertaken in 2016 is also shown

3. Environmental/Oceanographic Effects				
11	Sea-Surface Temperature - Weekly mean	1	Normalised SST<-1.0	20,953
		2	Normalised SST between -1.0 and -0.43	25,774
		3	Normalised SST between -0.43 and 0.0	24,109
		4	Normalised SST between 0.0 and 0.43	23,445
		5	Normalised SST between 0.43 and 1	22,226
		6	Normalised SST >1.0	20,862
12	Southern-Oscillation Index	1 to 6	As for Sea-Surface Temperature	20,325 to 25,908
13	Mixed-Layer Depth	1 to 6	As for Sea-Surface Temperature	13,148 to 32,991
14	Wind Speed	1 to 6	As for Sea-Surface Temperature	19,326 to 28,325
15	Bathymetry	1	<1000m	6,832
		2	between 1000 and 1500m	16,500
		3	between 1500 and 2000m	21,929
		4	between 2000 and 3000m	29,486
		5	between 3000 and 4000m	27,357
		6	>4000m	35,265
16	Moon Phase	Continuous	Fraction of visible moon	
4. Cooperative/Competitive Effects				
17	Number of other Vessels in same 1-Degree Square and Day	1	>6 other vessels	4,317
		2	6 other vessels	2,590
		3	5 other vessels	4,126
		4	4 other vessels	6,743
		5	3 other vessels	10,958
		6	2 other vessels	18,157
		7	1 other vessel	32,293
		8	no other vessels	58,185
18	Number of other Vessels in same 1-Degree Square and Month	1	more than 20 other vessels	7,422
		2	18-20 other vessels	4,365
		3	15-17 other vessels	7,470
		4	12-14 other vessels	10,711
		5	9-11 other vessels	15,738
		6	6-8 other vessels	25,600
		7	3-5 other vessels	31,363
		8	less than 3 other vessels	34,700

- 3) Environmental/Oceanographic Effects – like the statistical effects listed above, these effects attempt to account for differences in the availability of the fish due to behavioral responses to local changes in ocean conditions and changes in their diurnal behavior. Variables include daily moon phase (expressed as the fraction of the whole moon visible), weekly mean sea-surface temperature, monthly mean Southern-Oscillation Index, mixed-layer-depth, wind-speed, and current strength, with both east-west and north-south components, current speed and current direction.
- 4) Vessel Cooperative/Competitive Effects – these effects attempt to account for the influence of vessels cooperating or competing within a similar area of the fishery. Variables include the number of vessels within same 1-degree square/day) and the number of vessels within the same 1-degree square/month.

To allow flexibility in the nature of the fitted relationship and to avoid spurious results which can be associated with a lack of data at the tails of fitted splines (for example, if one uses GAMs), most variables were fitted as categorical variables with a given range of values for each variable being associated with a discrete category (e.g. the start times were categorized into six 4-hourly intervals of time, c.f. Table 5.1). Only moon-phase was fitted as a continuous variable. Some of the environmental / oceanographic variables were normalized using the mean

and standard deviation of the values across all data included in the analysis [i.e. $z=(x-\mu)/\sigma$], then categorized into one of the six categories depending on whether the value of the normalized variable $|z|$ was less than or greater than $z=-1.0$, -0.43 , 0 , 0.43 or 1.0 . (Note, for the two current variables the transformation used was limited to $z=x/\sigma$ in order to preserve the directional relationship between z and the corresponding current component.)

5.3.2 Abundance Index

The results from fitting the above two models are used to calculate the standardized index, S , in each year, quarter and area strata:

$$S(\text{year}, \text{qtr}, \text{area}) = S_{y,q,a} = p_s(\text{year}, \text{qtr}, \text{area}) * \mu_s(\text{year}, \text{qtr}, \text{area})$$

The expected value of the standardised values of both p_s and μ_s in each year, quarter and area can be found by selecting a standardising reference level for each of the fitted explanatory effects. In practice that categorical level for which the related parameter β_j is zero is selected (for SAS this corresponds to the last level of each fitted effect). An annual index of abundance, $I(\text{year})$, is then determined by first calculating the area-weighted sum of the standardized index across all NA areas and then taking the average across all $NQ=4$ quarters as follows:

$$I(\text{year}) = \sum_{\text{area}=1}^{NA} \left[\frac{\text{Size}_{\text{area}}}{NQ} \sum_{\text{qtr}=1}^{NQ} p_s(\text{year}, \text{qtr}, \text{area}) * \mu_s(\text{year}, \text{qtr}, \text{area}) \right]$$

where $\text{Size}_{\text{area}}$ is the spatial size of the individual areas (as measured by the number of 1-degree squares in each area).

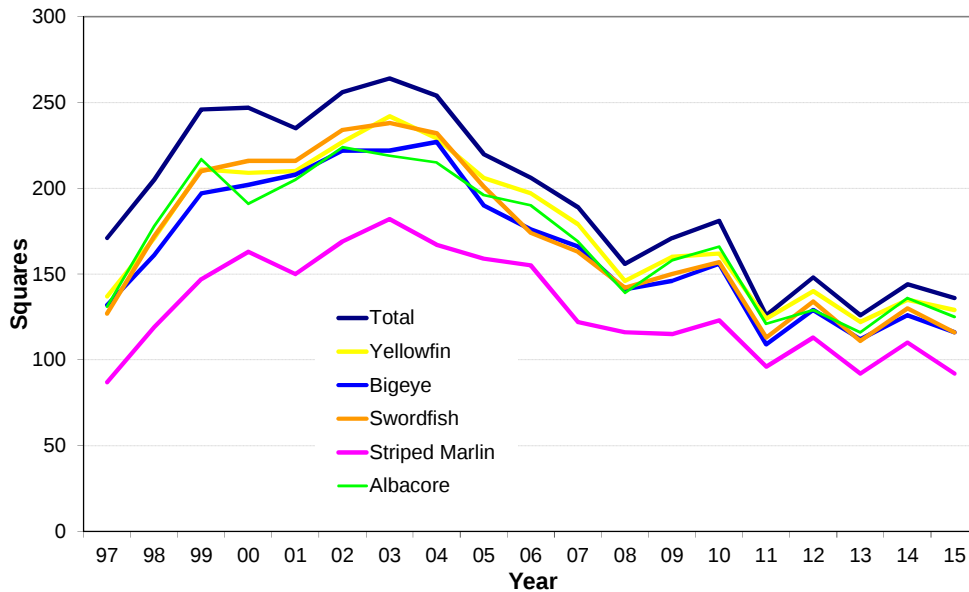
Given the above equation there is a need to be able to calculate a standardized CPUE for each combination of the year, quarter and area strata included in the standardizing model. While this is usually achieved for models without interaction terms, when interactions are included the model may not provide an estimate of the standardized CPUE for all strata. In those instances where strata remain unobserved, then there will be a need to impute a value of the standardized CPUE for these strata. While various methods have been suggested for imputing such values (Walters 2003; Campbell 2004; Carruthers *et al.* 2011; Campbell 2015), where this occurred the standardised CPUE value for each missing stratum was taken from the value estimated for this stratum based on fitting the simpler standardizing Model 1.

Finally, the annual index for all years is scaled so that the mean of the annual index over the entire time-series was equal to 1.

5.3.3 Selection of Core-Catch Area for each species

Ideally, one would like to construct an annual abundance index based on the total size of the resource available to the fishery. This can be defined as the resource to be found within the total area fished by the fishery. However, the changing spatial extent of the ETBF (c.f. Figure 5.1) creates a number of problems for the calculation of annual abundance indices. For example, since 1997 the spatial extent of the fishery has ranged between 126 (in 2011 and 2013) and 264 (in 2003) one-degree squares. Also, different sets of squares are fished each year, and of the total of 387 distinct squares fished in the ETBF since 1997, less than one-fifth (68) have been fished in all years between 1997 and 2015. Furthermore, in the past some of the highest catch rates of Swordfish have been achieved in the off-shore areas of the ETBF east of 160°E. However, in recent years there has been little or no fishing in this region. As such, it is not possible to estimate the values of p_s and μ_s in these areas and include them in the annual abundance index defined above.

Figure 5.1: Total number of 1-degree squares fished each year in the ETBF and the corresponding number of squares in which each of the main target species were caught.



So that changes in the index are due to changes in the abundance of available fish and not due to changes in the spatial extent, there is a need to define some core spatial region of the ETBF that remains constant over the assessment period and over which an abundance index for each species can be calculated each year and compared. Inferences about changes in abundance based on changes in the standardised CPUE will therefore be limited to the size of the resource available within this core region (and not the whole population of stock).

In order to identify a core region for each species over which the abundance index can be calculated, and taking into account the need for such a region to generally coincide with the areas of the fishery with continuous history of being fished, the following approach was followed:

- 1) The number of years that each 1-degree square of the ETBF had been fished over the period of interest (N-years) was calculated.
- 2) The percentage of the total catch in each year which was caught in those squares that had been fished in all N-years was calculated. If this percentage exceeded 90% in all years then the core area for this species was taken to be the union of all these 1-degree squares.
- 3) If the percentages calculated in the previous step were not all greater than 90% then the percentage of the total catch in each year caught in all squares fished for N-1 or more years was calculated. Again, if these percentages exceeded 90% in all years then the core area for this species was taken to be the union of all these 1-degree squares.
- 4) This step-by-step analysis was continued until the percentage of the total catch taken in all 1-degree squares which has been fished at least C-years (where $C \leq N$) exceeded 90% in each year. The core area was then taken to be the union of all these 1-degree squares.

The application of this approach to identifying the core area for each of the five principal target species in the ETBF is described in Campbell (2012).

5.3.4 Determination of GLM Areas

Having selected a core region for each species, this region is then sub-divided into a number of sub-regions, or areas, (usually 6 or 7) to serve as (and test for) Area-effects within the GLM. For each species these areas were selected as follows:

- 1) The nominal CPUE (defined as the sum of catches divided by the sum of effort) within each 1-degree square within the core region was calculated for each year (but only where the number of fishing operations was 5 or more). The mean of these nominal CPUEs was then calculated over all years and the distribution of these mean CPUEs for each 1-degree square was mapped.
- 2) The core region was then subdivided into 6 or 7 spatially coherent and cohesive areas by grouping together 1-degree squares having similar CPUE. These areas, which are used as categorical *Area* effects in the GLM, are taken as being representative (and therefore explanatory) of differences in the density of the species of interest across the core fishery, possibly due to differences in habitat preferences. While alternative approaches can be adopted for identifying *Area* effects, such as using the distribution of structural habitat features across the fishery, this approach makes more direct use of the spatial distribution of observed CPUE to infer patterns in the spatial distribution of the resource. It also overcomes the limitations in just constructing *Area* effects simply based on blocks of latitude and longitude, as is often done.

Again, the application of this approach to identifying these sub-regions for each of the five principal target species in the ETBF is described in Campbell (2012).

5.4 Clustering catch as a proxy for targeting

It is often assumed that in a multispecies fishery that fishes have some ability to target individual species (or mix of species) through spatial (and temporal) targeting and by changing the manner in which the fishing gear is deployed. For example, there is some evidence based on archival tagging that different pelagic species (such as tunas and billfish) inhabit different depth strata, which may vary both diurnally and seasonally (Gunn *et al.* 2005; Evans 2010; Evans *et al.* 2011). As such, in a pelagic longline fishery a different mix of species can potentially be targeted by deploying the fishing gear at different specified depths. Evidence for such targeting in the ETBF will be explored in detail in the next chapter.

Unlike single species fisheries where all effort is directed at the target species, in multispecies fisheries the effort can potentially be targeted at different species (or a difference mix of species), i.e. the main species targeted by the fishing gear may change. Where this occurs, the fishing effort needs to be adjusted so that the "effective" effort directed at any specific species of interest can be ascertained. If this is not undertaken correctly then the resulting index of resource abundance is likely to be biased and unreliable. When there is a relationship between the targeting of different species and the manner in which the fishing gear is deployed, and where information is available for each fishing operation on the how the fishing gear is deployed, then it is desirable to include this information in the methods used to standardise the related CPUE. However, where such information is not available, other methods need to be used to segregate dissimilar types of fishing operations. It is possible that the composition of the catch will act as a proxy for targeting behaviour. For example, with multispecies trawl fisheries operations are often segregated according to the most abundant species in the catch. Alternately, cluster analysis can be used to categorise fishing operations by the similarity in

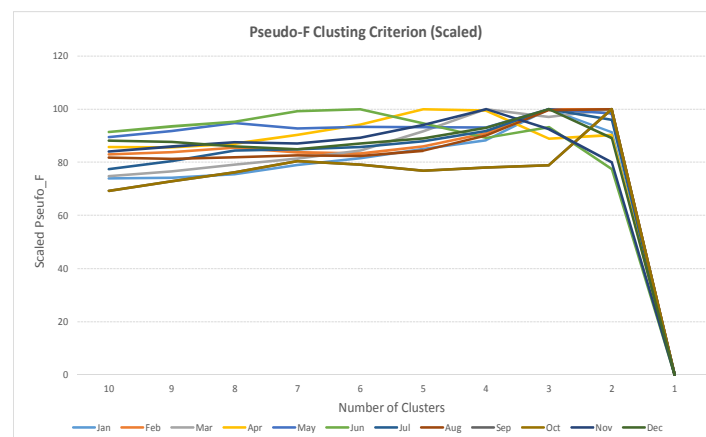
catch composition (Rogers and Pikitch 1992). In a similar manner, cluster analysis can be used to classify longline sets in relation to the species composition of the catches.

We follow the method of He *et al.* (1997), who used cluster analysis to group the entire set of fishing operations for the Hawaiian longline fishery into groups of different fishing strategies based on the species composition of the catch. We illustrate this approach using the data set consisting of catch and effort data for 111,600 individual longline sets in the ETBF between 2000 and 2013. Although information on up to twenty individual species is included in the catch for each set, in this initial exploratory analysis we limit the catch to the following five species: Yellowfin Tuna (YFT), Bigeye Tuna (BET), Albacore Tuna (ALB), Southern Bluefin Tuna (SBT) and Broadbill Swordfish (SWO). (Note, the catch was taken as a sum of retained and discarded catches for each of these species.).

Records were deleted if the total catch of these five species was zero leaving 108,650 records to be included in the analysis. Catch composition for each set was then calculated as the proportion of each species to the total catch of the set. Data were arcsin-square-root transformed before analysis to normalise their distribution (Snedecor and Cochran 1980). A hierarchical cluster analysis (known as the *Ward* method, Ward 1963) was then applied to the 108,650 records. The choice of the number of clusters was guided by the number of clusters identified using various statistical metrics (Cubic Clustering Criterion CCC, Pseudo-F), but to some extent was largely subjective. All analyses were undertaken using the SAS CLUSTER Procedure (SAS 1999).

Two sets of analyses were undertaken: (a) for each month separately, and (b) for all data combined. The monthly analyses explore the possibility that, due to changes in the prevailing oceanographic conditions, the associations between species available to the longline gear can change seasonally. The evidence for this is more fully explored in Section 6.3. Based on an inspection of the Pseudo-F clustering criterion for each month (c.f. Figure 5.2) the number of clusters for each month was set to be four.

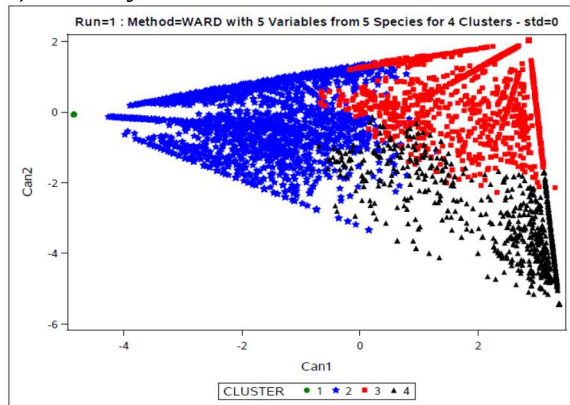
Figure 5.2: Pseudo-F criteria for identifying number of clusters. Results shown for each month.



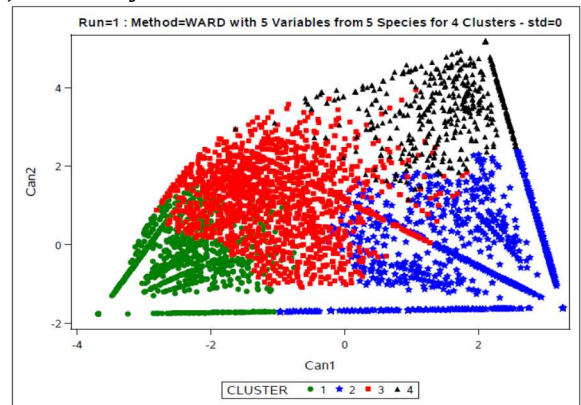
Pictorial representation of the four clusters for each month based on plots of the first two canonical variables are shown in Figure 5.3a-l, while illustrative dendrograms of cluster associations for the four months (January, February, November and December) are shown in Figure 5.4a-d. For most months these results clearly display a partitioning of the data in four semi-distinct clusters, indicating clear structure in the species composition of the catches. This structure may, or may not, be related to particular targeting strategies adopted by fishers.

Figure 5.3: Pictorial representation of monthly clusters based on plots of the first two canonical variables. (NB, the lines of points are assumed to be representative of the zero catches in the data.)

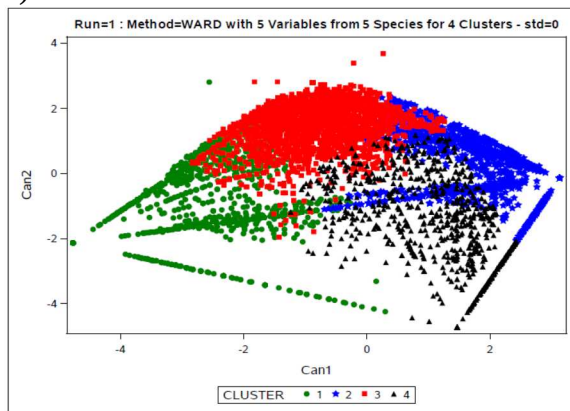
a) January



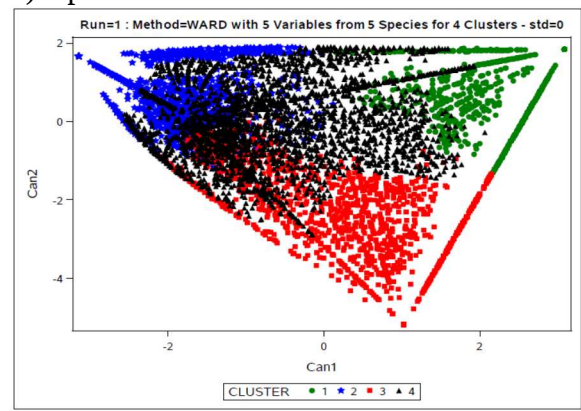
b) February



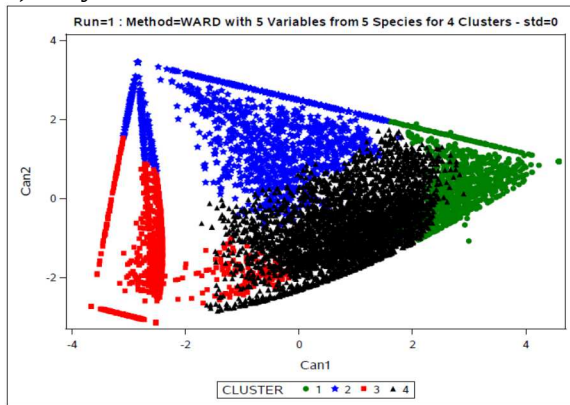
c) March



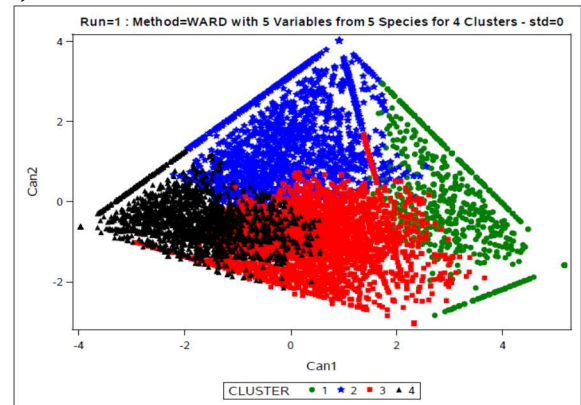
d) April



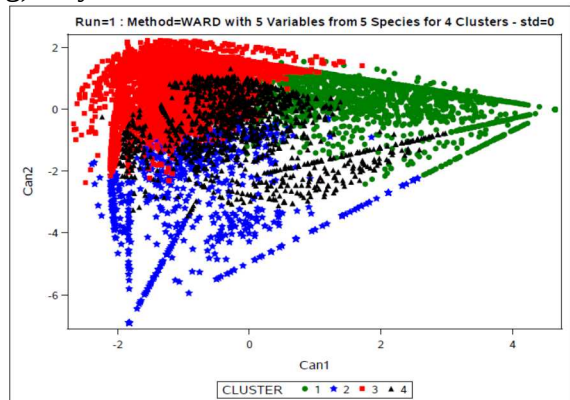
e) May



f) June



g) July



h) August

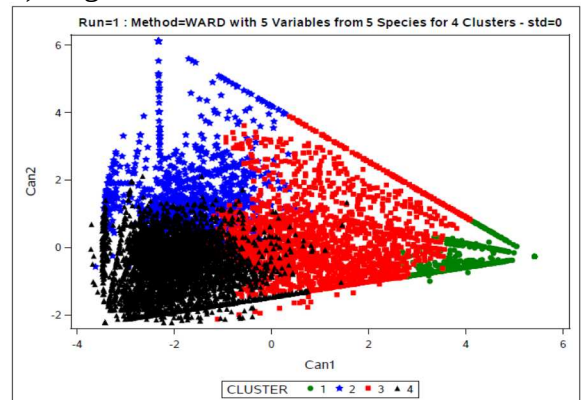
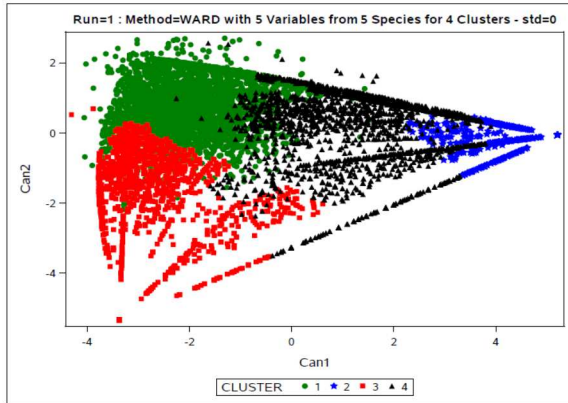
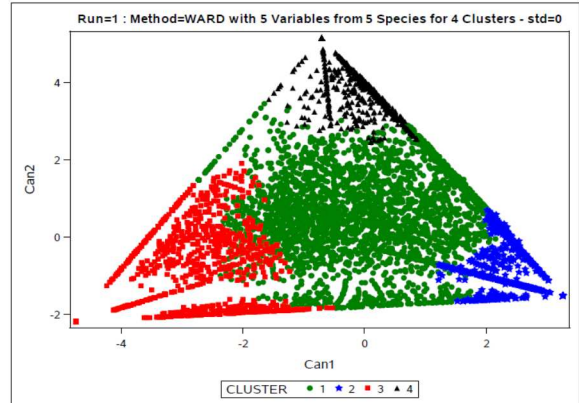


Figure 5.3: (cont'd). Pictorial representation of monthly clusters based on first two canonical axes.

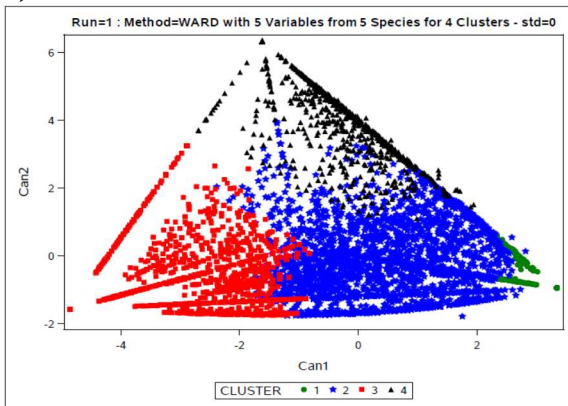
i) September



j) October



k) November



l) December

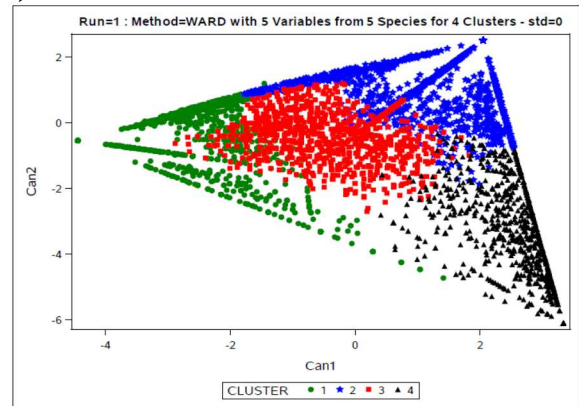
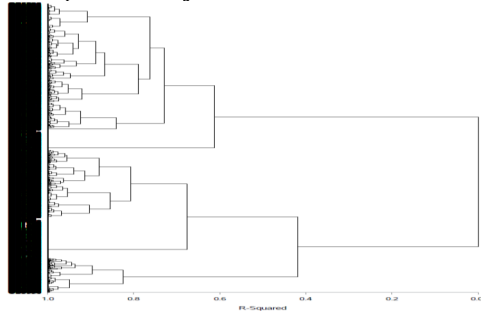
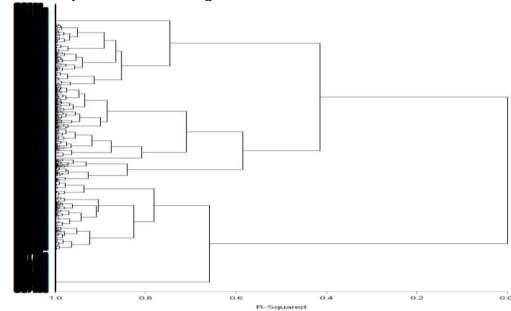


Figure 5.4: Dendrograms of cluster associations. Results are shown for the following four months: (a) January, (b) February, (c) November and (d) December.

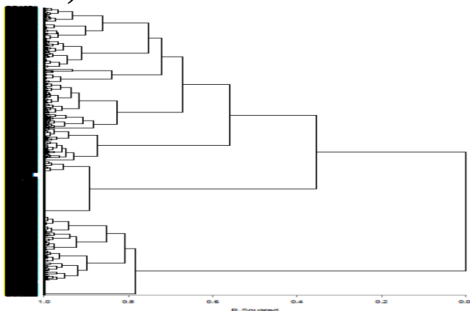
a) January



c) February



b) November



d) December

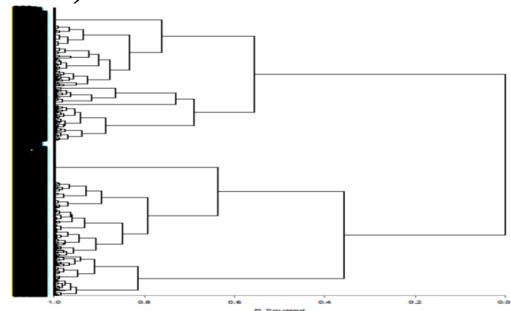


Table 5.2: Catch composition of the four clusters based on the cluster analysis for each month. The last column titled Target indicates the final targeting type (1-7) based on the final cluster analysis of the 48 clusters shown in the Table. (NOPS = Number of fishing operations.)

OBS	MONTH	CLUSTER	NOPS	YFT	BET	ALB	SBT	SWO	TARGET
1	1	1	557	0.00	0.00	0.00	0.00	100.00	2
2	1	2	3390	19.78	15.50	7.25	0.00	57.47	4
3	1	3	2917	82.09	6.21	6.73	0.00	4.97	1
4	1	4	924	16.21	6.63	72.23	0.43	4.50	3
5	2	1	2327	7.32	11.21	3.22	0.00	78.25	2
6	2	2	3159	83.22	3.22	5.66	0.00	7.90	1
7	2	3	2079	27.78	29.38	13.70	0.03	29.11	6
8	2	4	613	17.06	12.40	70.00	0.00	0.54	3
9	3	1	2036	11.20	7.85	7.19	0.01	73.75	2
10	3	2	3057	82.28	6.58	6.76	0.00	4.38	1
11	3	3	2790	23.02	36.72	12.90	0.15	27.21	6
12	3	4	1129	24.04	8.18	61.62	0.00	6.16	3
13	4	1	1910	88.41	1.69	6.81	0.02	3.07	1
14	4	2	1825	16.18	15.89	8.36	0.00	59.56	4
15	4	3	1846	16.92	8.67	65.33	0.03	9.05	3
16	4	4	4121	31.37	39.59	17.22	0.10	11.71	6
17	5	1	1401	7.90	8.72	80.96	0.20	2.22	3
18	5	2	2407	75.69	6.25	14.29	0.04	3.72	1
19	5	3	1843	24.29	42.13	0.93	0.03	32.62	6
20	5	4	4581	16.31	30.52	37.41	0.18	15.58	7
21	6	1	1044	30.24	53.01	2.36	0.21	14.18	6
22	6	2	2572	60.39	8.65	27.50	0.36	3.11	5
23	6	3	3483	10.58	24.03	38.90	3.62	22.87	7
24	6	4	2497	7.34	8.04	79.65	1.21	3.76	3
25	7	1	2445	72.00	4.15	20.64	0.38	2.82	1
26	7	2	976	16.40	12.27	18.29	1.09	51.96	4
27	7	3	4910	10.26	10.01	68.53	4.16	7.04	3
28	7	4	1737	23.81	44.92	24.58	0.11	6.58	6
29	8	1	1540	95.87	0.18	3.06	0.01	0.87	1
30	8	2	1407	14.23	33.43	27.86	0.13	24.34	7
31	8	3	2745	57.33	12.04	22.30	0.15	8.17	5
32	8	4	4347	13.14	6.92	64.94	2.62	12.38	3
33	9	1	3799	18.87	9.10	56.79	0.92	14.32	3
34	9	2	1503	97.84	0.77	0.72	0.00	0.67	1
35	9	3	1418	11.63	11.04	18.71	0.02	58.61	4
36	9	4	2923	60.43	13.27	16.94	0.03	9.33	5
37	10	1	4003	35.01	12.48	28.38	0.28	23.85	5
38	10	2	2219	91.14	5.88	2.59	0.01	0.38	1
39	10	3	1993	9.89	6.26	7.60	0.01	76.23	2
40	10	4	765	8.81	2.17	88.04	0.24	0.74	3
41	11	1	1599	98.43	0.88	0.69	0.00	0.00	1
42	11	2	3500	48.59	14.52	14.50	0.56	21.83	5
43	11	3	2294	9.25	5.69	7.60	0.00	77.46	2
44	11	4	938	19.07	2.18	75.71	0.60	2.44	3
45	12	1	1930	9.47	9.67	4.11	0.01	76.74	2
46	12	2	2850	81.78	5.81	3.71	0.03	8.68	1
47	12	3	1331	26.02	23.90	15.22	3.68	31.18	6
48	12	4	970	23.85	6.90	68.21	0.18	0.86	3

The catch composition of the four clusters based on the cluster analysis for each month is shown in Table 5.2. As it is likely that similar cluster types may occur across the different months (e.g. cluster-type 4 is similar for January and February), a second cluster analysis was undertaken where the 48 clusters shown in the Table 5.2 were clustered based on the species composition of each cluster. Again, the *Ward* method was used. Based on the clustering criteria shown in Figure 5.5, the final number of cluster types selected for this second stage of the analysis was taken to be seven. The dendrogram of cluster associations based on analysis of the 48 monthly clusters is shown in Figure 5.6a while the associated representation of the seven cluster types based on plots of the first two canonical variables is shown in Figure 5.6b. Again, these results appear to indicate structuring of the data into around seven cluster-types across all months.

Figure 5.5: Clustering criteria used to identify the appropriate final number of clusters based on clustering the 48 monthly cluster types identified in the first stage of the analysis.

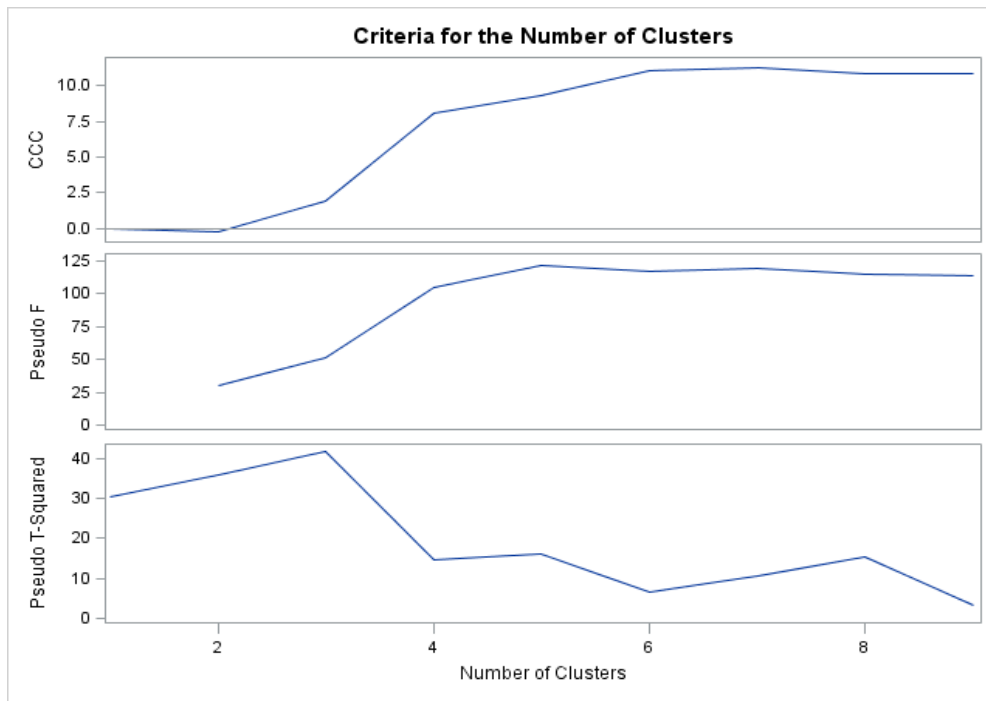


Figure 5.6: (a) Dendrogram of cluster associations based on analysis of the 48 monthly clusters, and (b) representation of the seven cluster types based on plots of the first two canonical variables for the 48 monthly clusters.

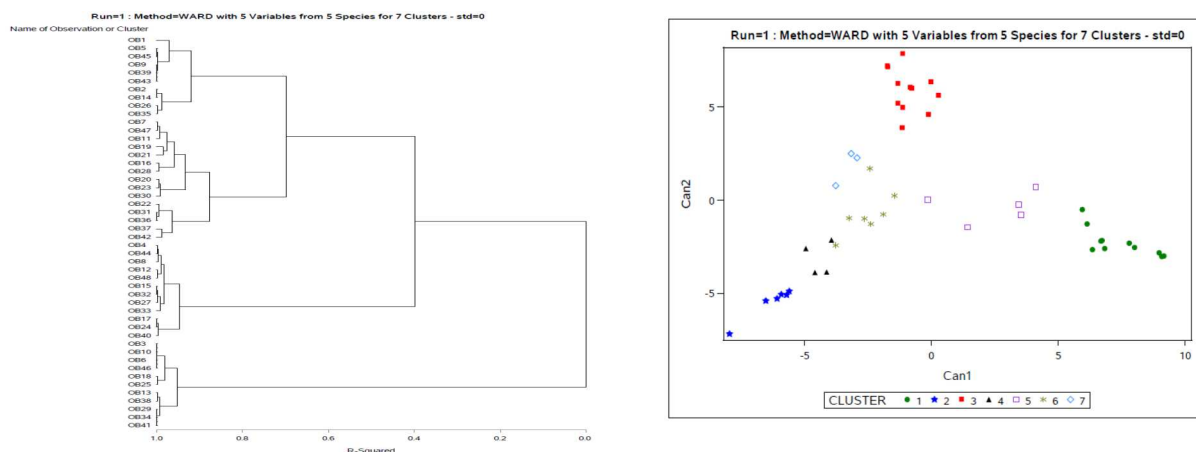


Figure 5.7: Distribution of the 48 monthly cluster types identified in stage 1 of the analysis across the seven cluster types identified in stage two of the analysis.

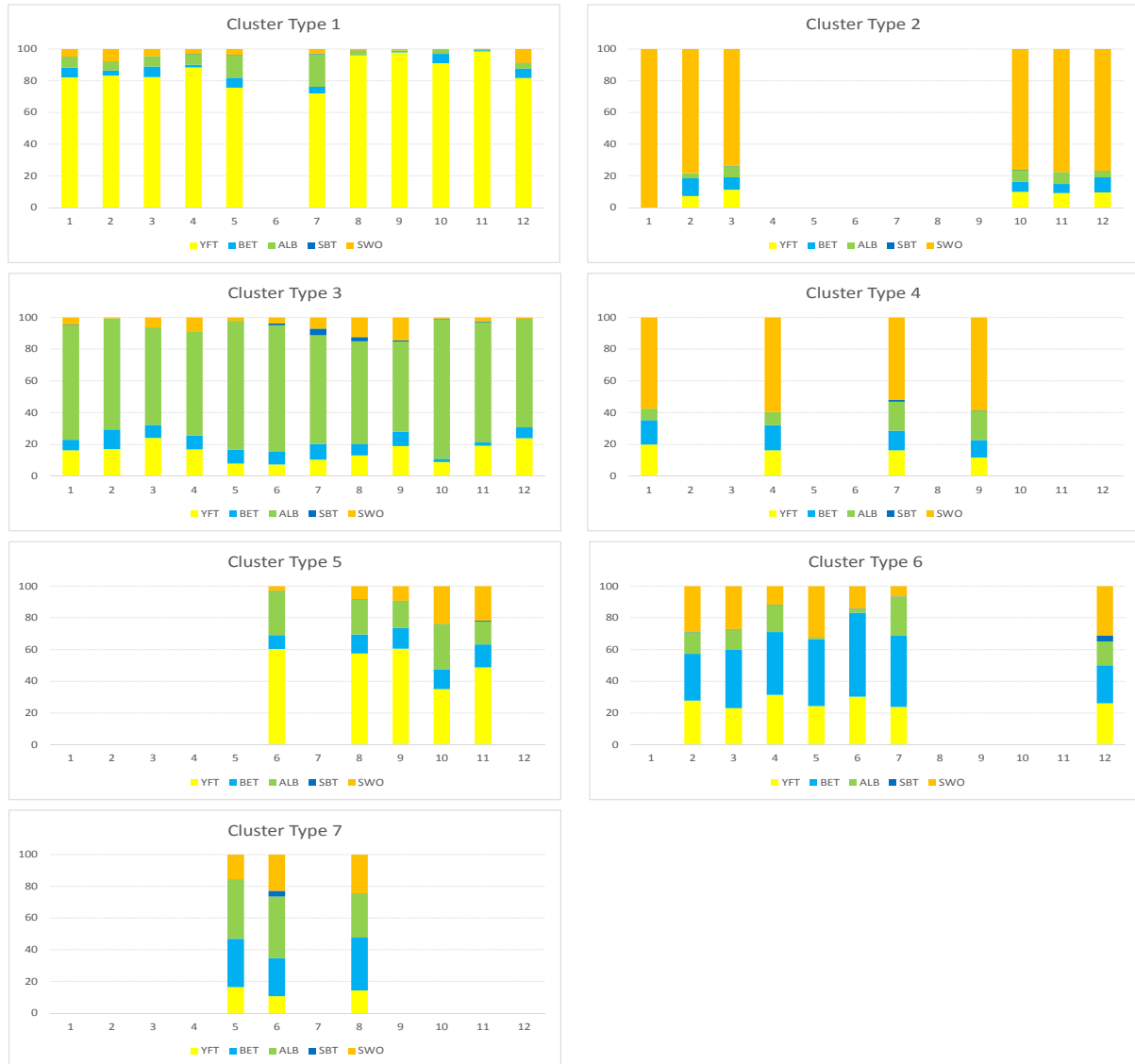


Figure 5.8: Distribution of cluster types (percentage of sets) across each (a) year and (b) month.

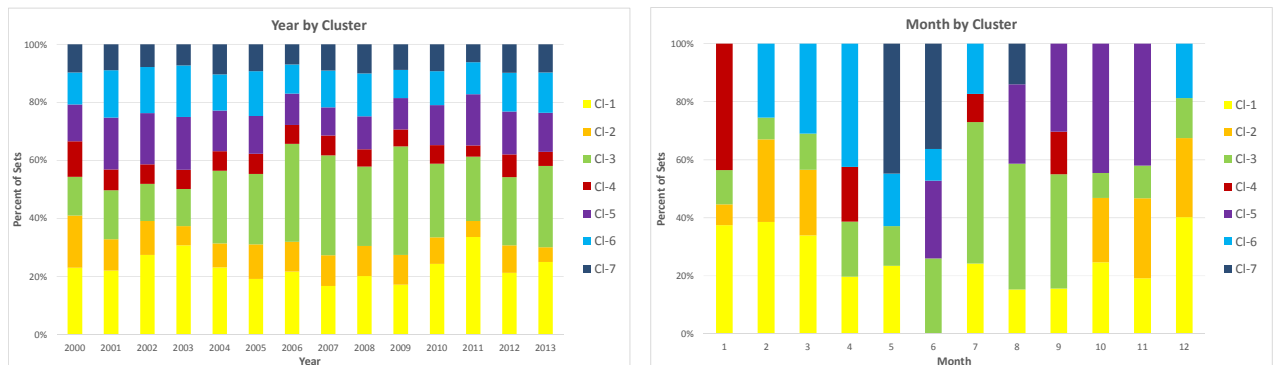
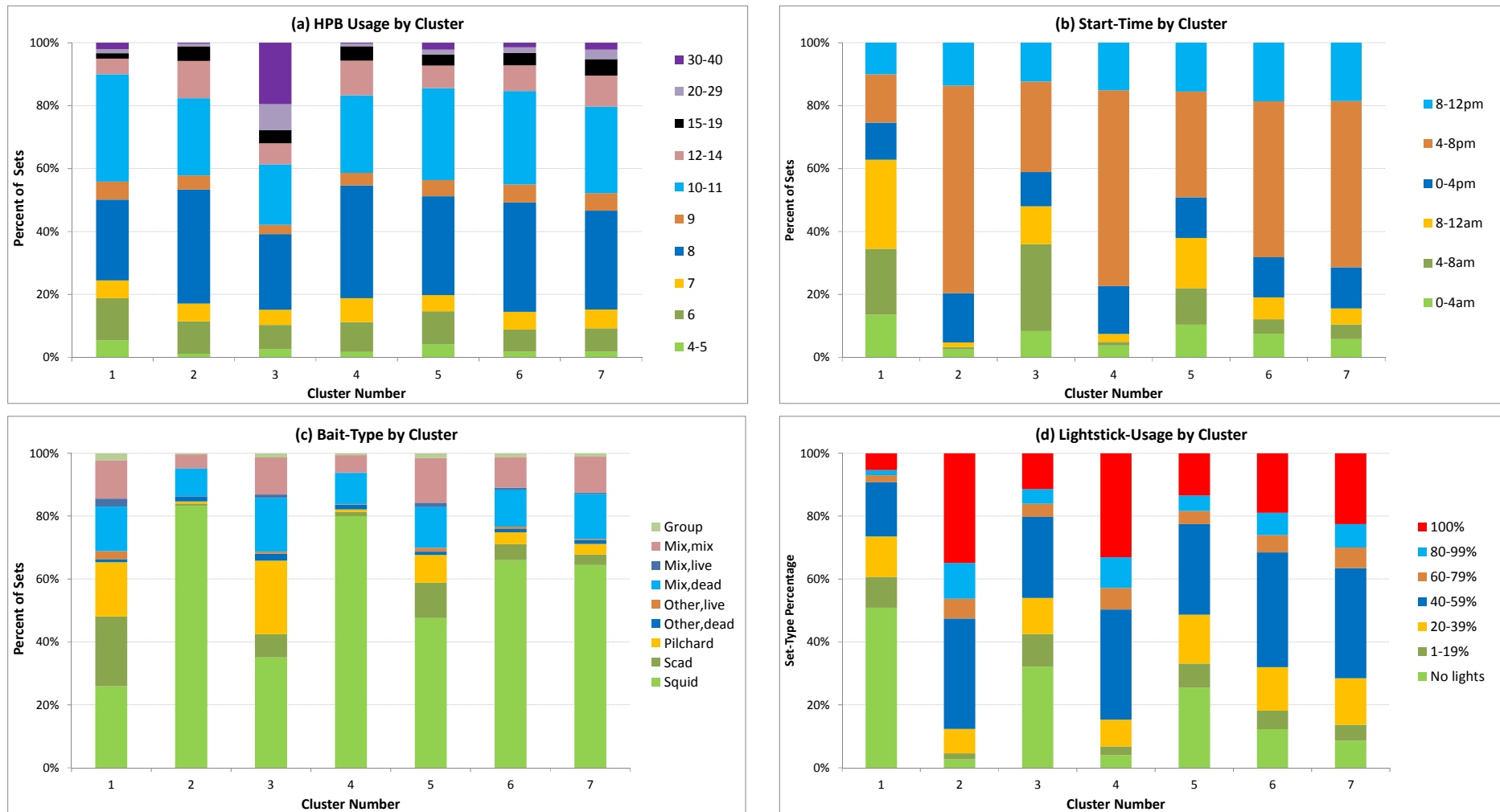


Figure 5.9: Distribution of gear-settings across the seven cluster type identified in the 2-stage monthly analysis.



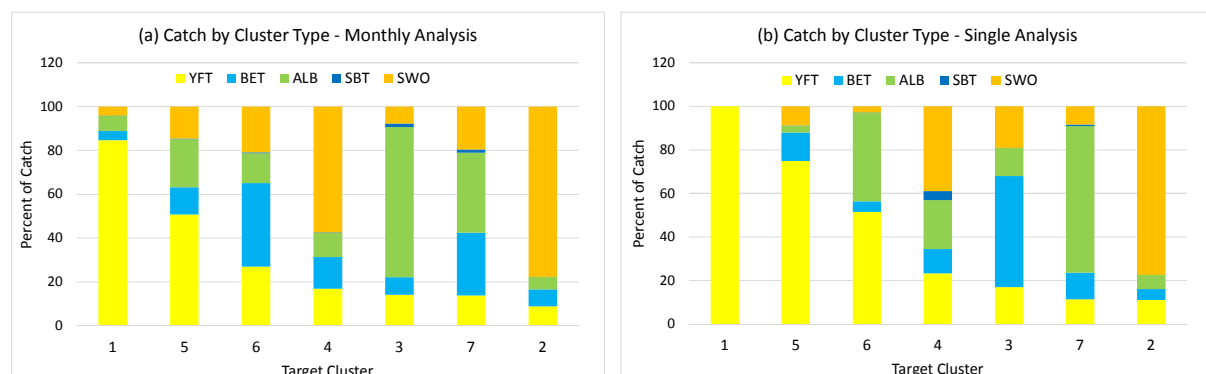
The distribution of the 48 monthly cluster types identified in stage 1 of the analysis across the seven cluster types identified in stage two of the analysis are shown in Figure 5.7. Based on the catch composition shown for each cluster type it is seen that Cluster-type 1 can be associated with targeting Yellowfin Tuna, Cluster 2 can be associated with targeting Swordfish, Cluster 3 can be associated with targeting Albacore Tuna and Cluster 6 can be associated with targeting Bigeye Tuna. The other three clusters are hybrids of these initial four. Interestingly, the months that the Swordfish cluster-type occurs corresponds to the months when this species is most available (c.f. Figure A.1, Appendix A), while a similar result is also seen for the Bigeye Tuna cluster-type.

Finally, the catch composition of the seven cluster-types identified in the 48 monthly clusters (shown in Table 5.3a, Figure 5.10a) can be compared with the catch composition of the seven cluster-types identified directly by a single analysis of all 108,650 records (shown in Table 5.3b, Figure 5.10b). A number of differences can be seen between these two results. First, while

Table 5.3: Composition of cluster types based on (a) monthly analyses combined into 7 clusters, and (b) single analysis of all data into 7 clusters. Note, the clusters in each sub-table are ordered by decreasing percentage of Yellowfin Tuna. (NOPS = Number of fishing operations.)

(a)	CLUSTER	NOPS	YFT	BET	ALB	SBT	SWO	Total
	1	25,606	84.68	4.27	7.00	0.05	4.00	100
	5	15,743	50.79	12.38	21.97	0.29	14.58	100
	6	14,945	27.00	38.11	13.56	0.42	20.91	100
	4	7,609	16.96	14.35	11.07	0.14	57.48	100
	3	24,139	14.09	8.10	68.40	1.66	7.75	100
	7	9,471	13.90	28.57	36.54	1.44	19.56	100
	2	11,137	8.89	7.75	5.63	0.01	77.73	100
		108,650						
(b)	CLUSTER	NOPS	YFT	BET	ALB	SBT	SWO	Total
	7	10,670	100.00	0.00	0.00	0.00	0.00	100
	6	17,913	74.91	13.06	3.13	0.05	8.91	100
	2	13,458	51.45	4.99	40.70	0.01	2.89	100
	1	11,802	23.29	11.14	22.50	4.16	38.96	100
	4	13,969	17.00	50.99	12.91	0.03	19.08	100
	3	25,636	11.34	12.26	67.34	0.63	8.49	100
	5	15,202	11.07	5.06	6.37	0.00	77.55	100
		108,650						

Figure 5.10: Catch composition by seven clusters based on (a) the analysis by month, and (b) the single analysis of all sets.



the catch compositions of certain cluster-types are similar for each result (e.g. cluster-type 2 for the monthly analysis and cluster-type 5 for the single analysis can both be associated with Swordfish dominated catches), there are substantial differences in the composition of other cluster-types. For example, for the dominant Yellowfin cluster-type based on the monthly analysis the average composition of the catch is 85% Yellowfin Tuna, while for the dominant Yellowfin cluster-type based on the combined analysis the catch is 100% Yellowfin Tuna. Second, the distribution of fishing operations across each of the seven cluster-types is also substantially different. While this difference is obviously associated with the dissimilarity in the associated catch compositions, even for those cluster-types that are similar between the two results (e.g. the two dominant Swordfish clusters) the number of sets associated with each is quite different (11,137 versus 15,202). As explained earlier, the monthly-based analysis was undertaken to take into account possible changes in species associations between months that may not be adequately accounted for by a single analysis across all months. While differences in the results of the two approaches have been noted, as to which of the two approaches may be better in identifying targeting behaviour will be explored in the simulation testing reported later in this report.

5.5 Bayesian spatial and spatial-temporal models with INLA

Pelagic longline fishing is a multispecies fishery where many species can be caught in one gear deployment. However, catch rate for each species can vary over time and space. Some species are very common in the catch, while other species can be rarely caught and the catch rate may be very low. Traditionally, catch-per-unit-effort (CPUE) standardisation is carried out species by species (Campbell 2004, 2015; Maunder and Punt 2004) so the level of uncertainty can vary significantly across species due to varying quantity and quality of data.

We believe that if several species are captured in the same gear deployments (shots), there must be some similarities among these species, e.g. distribution, behaviour, abundance. This raises the question as to whether there are advantages to model multispecies together in a single model? For example, it is well known that modelling multiple groups of subjects under a Bayesian hierarchical structure allows sharing and borrowing of information across these groups (Gelman and Pardoe 2006; Zhou *et al.* 2008).

Recently, Bayesian modelling has been used in single species CPUE standardisation (Cao *et al.* 2011; Zhang and Holmes 2009) and such an approach has several advantages. Through the specification of prior distributions, the Bayesian method allows the formal inclusion of information from previous studies, expert opinion, or similar studies in other occasions. From the Bayesian posterior distribution we can easily obtain the probability of a parameter in relation to a certain threshold. Furthermore, the Bayesian approach allows easy construction of a hierarchical structure on the data and parameters, and enables borrowing strength and sharing information across multiple groups of subjects. Although the Bayesian approach is very flexible, it has a major drawback -- slow computing speed because it generally uses the Markov Chain Monte Carlo (MCMC) technique. With typically many thousands of data points, MCMC can be extremely slow and has been rarely used in CPUE standardisation. For example, in this project we tested Bayesian multispecies modelling using WinBUGS. It took several hours to reach convergence for a fraction of the ETBF data. The time for whole data-set could be impracticable.

Fishery catch data are often characterised by a spatial and temporal structure that should be taken into account in the modelling process. Obviously, fish abundance and availability change continuously over space and time. It is ideal to incorporate temporal-spatial information into the model as continuously connected covariates. However, existing CPUE standardisation methods rarely incorporate spatial and temporal effect as continuous covariates, but merely include them as discrete categorical variables.

Recently, a new tool, Integrated Nested Laplace Approximation (INLA), has been developed for Bayesian inference (Rue *et al.* 2009). This powerful tool substitutes MCMC simulations with accurate, deterministic approximations to posterior marginal distributions. INLA possesses some main advantages over MCMC techniques. The most outstanding benefit is its computational strength. INLA can produce results many times faster than MCMC. The second advantage is that INLA allows greater automation of the inference process, particularly useful for spatial or temporal models (Beguín *et al.* 2012; Blangiardo *et al.* 2013; Muñoz *et al.* 2013).

In this Section we investigate the feasibility of using INLA for CPUE standardisation. We compared single species and multispecies CPUE modellings using two approaches: (i) a generalized linear model using maximum likelihood estimation (MLE), and (ii) a Bayesian estimation technique using INLA.

5.5.1 Model Descriptions

Base GLM model

We tested and compared alternative models to explore potential effect of covariate selections, model structure, and spatial and temporal configurations. For each species, the base model is a negative binomial generalized linear model (GLM) in the following form:

$$\eta_i = g(\mu_i) = \beta_0 + \sum_m \beta_m x_{mi}$$

where mean μ_i is the expected catch on set i and is linked to the linear predictor η_i , β_0 is intercept, β_m is coefficient for the explanatory variable x_m , which is considered a fixed effect. Specifically, we included the following explanatory variables:

$$\eta_i = \beta_0 + \beta_{YQ} Y_i Q_i + \beta_{AQ} A_i Q_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) \quad (\text{Model 1})$$

Where Y is year, Q is quarter, A is area (region), HPF is hook per float, T is the start-time of the set, and h is the number of hooks. All predictors, except h , are treated as categorical variables. Catch c_i is modelled with a negative-binomial distribution. Note that the spatial predictor A is simply a categorical variable without any spatial continuity and relationship among them, whether between neighbouring areas or distant areas.

Geostatistical models

Logbook records from a pelagic longline fishery are made up of data measured at known locations, either at defined areas (regions) or latitude and longitude coordinates. The observed catch depends on a range of unobserved spatial and temporal variables such as fish density, movement, availability, schooling behaviour, fishing gear efficiency, etc. Geostatistical models are ideal approaches for modelling fisheries data (e.g. Thorson *et al.* 2015). In spatial statistics, geostatistical models can incorporate spatial dependency by using spatially-structured random-effect models.

There are two major types of spatial data. The first one is the area data, where shot-by-shot catch is a random aggregate value over an area with defined boundaries within a management jurisdiction (e.g. ETBF). The area can be either irregular or regular shapes. The second type of data is point-referenced geostatistical data, where latitude and longitude coordinates are random locations and the spatial index of the locations can vary continuously in the fixed management jurisdiction. We focus on the point data because they are real locations, whereas dividing a management jurisdiction into areas can be subjective.

We explored four geostatistical models. The first two models (referred to as Model 2 and Model 3 hereafter as Model 1 is the GLM model described above) are spatial models that only consider spatial effects, but not continuous spatial-temporal interactions. In contrast, the other two models, referred to as Model 4 and Model 5 hereafter, are spatio-temporal models.

In the spatial models, spatial catch data $c(s)$ are the realizations of a spatial process characterized by a spatial index s , which varies continuously in the fixed domain D (= ETBF jurisdiction). The vector of catches is assumed to follow a multivariate normal distribution with mean $\boldsymbol{\mu} = [\mu(s_1), \dots, \mu(s_n)]$ and spatially structured covariance matrix $\boldsymbol{\Sigma}$. Such a multidimensional spatial process is called Gaussian Markov Random Fields (GMRF) (Blangiardo and Cameletti 2015). The GMRF model can be expressed as:

$$\eta_i = \beta_0 + \sum_m \beta_m x_{mi} + \sum_l f_l(y_{li}) + f_s(s) + \varepsilon_i$$

The first two terms are similar to a GLM, where x can be variables such as gear types and environmental variables. The $\{f_l(\cdot)\}$'s are nonlinear smooth effects of the further covariates y . These can take many different forms, including nonlinear effects of continuous covariates, time trends, seasonal effects, random intercepts or slopes, group specific random effects, etc. $f(s)$ is spatially structured effects. These two terms are the major difference between a GLM and Gaussian random field models. The final ε_i is unstructured random effects.

To compare with the base GLM Model (1), we use the following two GMRF spatial models

$$\eta_i = \beta_0 + \beta_{YQ} Y_i Q_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) + f_s(s) + \varepsilon_i \quad (\text{Model 2})$$

$$\eta_i = \beta_0 + \beta_{YQ} Y_i Q_i + \beta_{AQ} A_i Q_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) + f_s(s) + \varepsilon_i \quad (\text{Model 3})$$

Model 2 treats spatial effects as a GMRF. The difference between Models 2 and 3 is the fixed effect term $\beta_{AQ} A_i Q_i$, which is the same as in GLM Model 1. We consider that the latent fields represent fish distribution in the real-world, whether they are caught in a given fishing location or uncaught (unobserved). Thus, we are not building models solely for discretely observed data, but for approximations of entire processes defined on continuous spatial domains. The GMRF models are implemented in INLA. INLA handles continuous Gaussian random fields by stochastic partial differential equations (SPDEs), and writes down explicit links between the parameters of each SPDE and the elements of precision matrices for weights in a discrete basis function representation (Lindgren and Rue 2014). The commonly used SPDE for Gaussian Markov random field model is the stationary Matern function:

$$(\kappa^2 - \Delta)^{\alpha/2}(\tau x) = W$$

where x is the latent spatial location, Δ is the Laplacian, κ is the spatial scale parameter, α controls the smoothness of the realisations, τ controls the variance, and W is a Gaussian spatial white noise process. The Matern covariance is:

$$\text{Cov}[x_i, x_j] = \frac{\sigma^2}{\Gamma(v)2^{v-1}} (\kappa \|x_j - x_i\|)^v K_v(\|x_j - x_i\|)^v$$

where Γ is the gamma function, K_v is the modified Bessel function of the second kind, and v is the smoothness parameter ($v > 0$).

Spatio-temporal models

The second geostatistical model is a spatial-temporal model that takes into account of both continuous dependency of spatial pattern and temporal trend. The base GLM (Model 1) and spatial Models (2, 3) treat temporal predictors (year and quarter) as discrete categorical variables. It is likely that the latent field of fish distribution and movement exhibits temporal correlation from one time step to the next. Modelling the spatial pattern of fish distribution alone cannot explain their temporal variation, which is equally important for marine animals. The number of fish or biomass in a given location is contingent on the number of fish or biomass in previous time steps at that location and surrounding locations. Hence, the spatial models are extended to the spatio-temporal models that include a time dimension, with the catch data $c(s, t) = \{c(s, t), (s, t) \in D\}$, where t is a time point. The spatio-temporal covariance function is given by $\text{Cov}[c(s_i, t_m), c(s_j, t_n)] = C(c_{im}, c_{jn})$. Similar to the spatial model, the general representation of spatio-temporal model is:

$$\eta_i = \beta_0 + \sum_m \beta_m x_{mi} + \sum_l f_l(y_{li}) + f_{st}(s, t) + \varepsilon_i$$

To build temporal process into a GMRF model, we use a continuous time measurement at year-quarter scale. There are alternative presentations of space-time interactions. As an example, we adopt Kronecker product models where both space and time are treated continuously. The Matern model above is used in space and a first order autoregressive model AR(1) is used in time (i.e., the current state is assumed to be affected by the state one time step earlier).

$$\eta_i = \beta_0 + \beta_{YQ} Y_i Q_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) + f_s(s, t) + \varepsilon_i \quad (\text{Model 4})$$

$$\eta_i = \beta_0 + \beta_{YQ} Y_i Q_i + \beta_{AQ} A_i Q_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) + f_s(s, t) + \varepsilon_i \quad (\text{Model 5})$$

Compared to Models 2 and 3, the difference here is the GMRF term $f_s(s, t)$. Model 5 is considered a “full model” as it uses both discrete and continuous spatial and temporal covariates.

The Bayesian Models 2 to 5 all involve a discrete covariate Q (quarter) as fixed effect. Since Models 4 and 5 treat year-quarter as a continuous random variable, a simple spatio-temporal model can be

$$\eta_i = \beta_0 + \beta_Y Y_i + \beta_{HPF} HPF_i + \beta_T T_i + \beta_h \log(h_i) + f_s(s, t) + \varepsilon_i \quad (\text{Model 6})$$

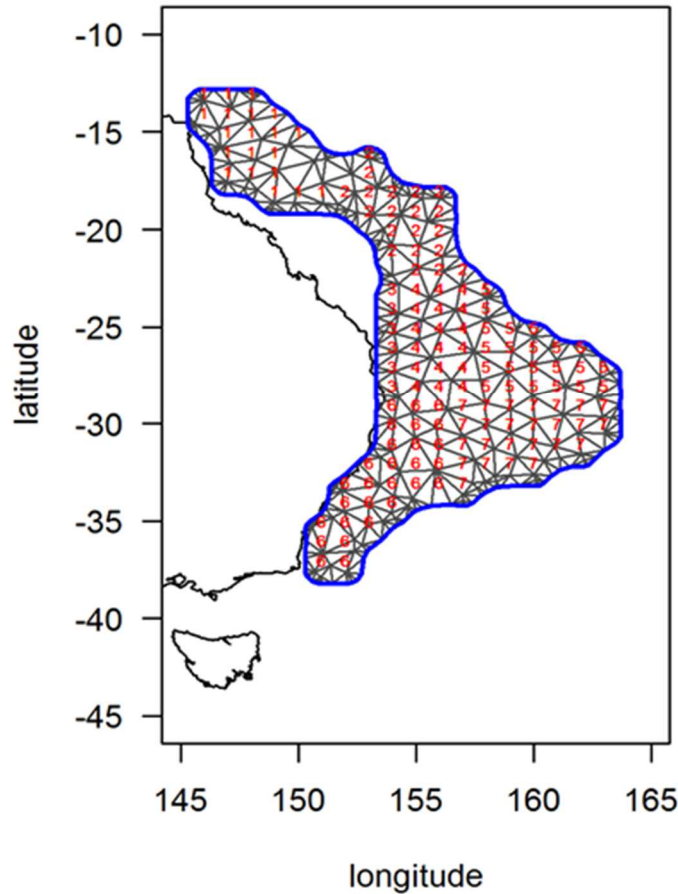
This model treats all spatial and temporal variables as random effects except year which is needed to derive annual abundance index.

5.5.2 Implementing Gaussian Markov random field model with INLA

Spatial model

Preparing this GMRF model requires much more effort than a GLM. The first step to model a latent GMRF is to build a spatial “mesh” based on latitude and longitude coordination from all

Figure 5.11: Building mesh for Gaussian Markov random field model in ETBF. The mesh is based on unique coordinates in the simulated data for all year (between 2000 and 2014). The numbers in red are areas used in GLM.



shots in the data (c.f. Figure 5.11). The triangulated mesh will provide a base for the GMRF models to build spatial representations. For example, in R codes (Version R 3.3.2) we define the ETBF boundary as:

```
>prdomain <- inla.nonconvex.hull(cbind(coord.uniq$lon, coord.uniq$lat), -0.03, -0.05,
resolution = c(100,100))
```

where *coord.uniq\$lon* and *coord.uniq\$lat* are unique longitude and latitude locations (respectively) where gear deployments have been observed in all years included in the assessment. The mesh is produced by:

```
>prmesh <- inla.mesh.2d(boundary=prdomain, max.edge=2)
```

We then compute the sparse weight matrices needed to map between the internal representation of weights for basis functions and the values of the resulting functions and fields by:

```
>A1 <- inla.spde.make.A(prmesh, loc=coord1)
```

```
>spde1 <- inla.spde2.matern(prmesh, alpha=2)
```

```
>mesh.index <- inla.spde.make.index(name="field",n.spde=spde1$n.spde)
```

The input data are prepared as:

```
> stk.dat <- inla.stack(data = list(catch = catch), A = list(A1,1), tag = "est", effects =  
list(c(mesh.index1, list(Intercept = 0)), list(year = dat$year, area = dat[,col.area[i]], qtr =  
dat$qtr, settime = dat$settime, hpf = dat$hpf, hooks = dat$hooks)))
```

For the purpose of CPUE standardisation, we need to estimate expected catch not only in actually fished locations but also in locations that were not fished in a particular year. Hence, prediction of expected catch in all locations is constructed:

```
>stk.pred = inla.stack(data = list(catch = NA), A = list(A.pred, 1), tag = 'pred', Effects =  
list(c(mesh.index1, list(Intercept = 0)), list(year = pred.dat$year, qtr = 1, area =  
pred.dat$area, settime = rep(1, npred), hpf = rep(1, npred), hooks = rep(mean(dat$hooks),  
npred)
```

Both input data and prediction are combined for model use:

```
>stk.all = inla.stack(stk.dat, stk.pred)
```

The GMRF Model (2) is specified as

```
> catch ~ factor(year) * factor(qtr) + factor(settime) + factor(hpf) + log(hooks) + f(field,  
model=spde1)
```

Spatio-temporal interaction model

To implement the spatio-temporal models in INLA, the formulation of Model 4 is coded as:

```
>catch ~  
1+Intercept+factor(year)*factor(qtr)+factor(settime)+factor(hpf)+log(hooks)+f(field,  
model=spde1, group=field.group, control.group=list(model='ar1'))
```

The model includes an explicit intercept, as for the spatial Models 2 and 3. The interaction between spatial field (named “*field*” here) and the temporal trend is represented by the term “*group = field.group*” and “*control.group=list(model='ar1')*”. The variable “*field.group*” is composed of temporal covariate at year-quarter scale (i.e. 60 continuous time steps for a time series of 15 years). The last term of *f()* indicates that at each time step the spatial locations are linked by the *spde* model object, while the temporal pattern is modelled as an AR(1) process.

For all these alternative models (GLM, spatial, and spatio-temporal models), we assume the catch per gear set (CPUE) follows a negative binomial distribution.

INLA is a full Bayesian approach and requires specifying priors for all parameters. For intercept and fixed effect variables, priors are controlled by functions *control.fixed()* and we use default vague prior Normal(0, 10⁶) for all parameters. For the random effect component in the GMRF model, prior is specified for the two parameters in the Matern function, τ and κ ; with the default Normal(0, 1) set in function *inla.spde2.matern()*.

5.5.3 Illustrative model evaluation and comparison

Data

The ETBF database contains a large number of shot-by-shot data. There are more than 100 thousand records for each species for the fourteen years 2000 to 2013. Multiplying by the number of species results in over a million records. In addition, the database has a range of covariates associated with each record. Even though INLA is deemed fast, running a

multispecies model that involves random effect terms with only a few covariates may take several hours. For the testing and exploring purposes, we used five species selected from the ETBF data-set with a wide range of catch records (i.e., YFT, BET, ALB, SWO, and STM) and randomly selected 10,000 records from 2000 to 2013 to reduce the computing time. Finally, as the seven *Area* effects used in the ETBF standardisation are unequal in size (c.f. Section 5.3.4), to derive an overall abundance index we used a weighted mean of the estimated catch in each location by the area size.

Results from Models 1 to 5 are compared with true abundance indices used to generate the simulated data (see next Section). Two measurements, mean relative error (MRE_M) and mean absolute relative error ($MARE_M$) for model M, are used

$$MRE_M = \frac{1}{n_Y n_A} \sum_Y \sum_A \frac{\hat{U}_{MYA} - U_{YA}}{U_{YA}}$$

$$MARE_M = \frac{1}{n_Y n_A} \sum_Y \sum_A \left| \frac{\hat{U}_{MYA} - U_{YA}}{U_{YA}} \right|$$

where n_Y and n_A are number of years and areas, $\hat{U}_{MYA}$ is estimated CPUE by model M in year Y area A, and U_{YA} is the true known CPUE in year Y area A.

5.5.4 Results

INLA produces a range of posterior estimates and we use the mean, including fixed and random effects. Spatial models that treat spatial information as random Gaussian fields do not outperform the generalized linear model that already includes spatial and temporal interactions (*Year*Quarter* and *Quarter*Area*). However, treating both spatial and temporal variables as continuous random effects in the spatio-temporal models (Model 4-6) clearly outperforms GLM and spatial models.

Table 5.4: Comparison of relative error between predicted abundance index and true index across 6 models and 5 species.

Model	ALB	BET	DOL	SWO	YFT	Abs Mean
1	0.003	-0.009	-0.027	-0.007	-0.007	0.011
2	0.007	-0.009	-0.032	-0.010	-0.008	0.013
3	0.000	-0.010	-0.031	-0.009	-0.008	0.012
4	0.011	-0.010	0.003	-0.001	-0.005	0.006
5	0.003	-0.011	0.002	-0.007	-0.004	0.005
6	0.012	-0.006	-0.004	0.003	-0.008	0.007

Compared with the true abundance index and based on the mean relative error MRE, the standardized indices from the spatio-temporal models (4, 5, 6) are less biased than GLM model (1) and spatial models (2, 3) (Table 5.4). On average, the full model 5 is least biased.

Compared with the assumed true abundance index, and based on the mean absolute relative error $MARE$, the standardized indices from the spatio-temporal Models (4, 5, 6) are more accurate than the GLM Model (1) and spatial Models (2, 3) (Figure 5.12, Table 5.5). On average, the full Model 5 is most accurate (the best for 3 out of 5 species), but the simple spatio-temporal Model 6 is similar (best for 2 out of 5 species). Overall, the simple spatio-temporal Model reduces GLM's $MARE$ by 20%.

Figure 5.12: Comparison between GLM Model (1), the full spatio-temporal Model (5) and the simple spatio-temporal Model (6).

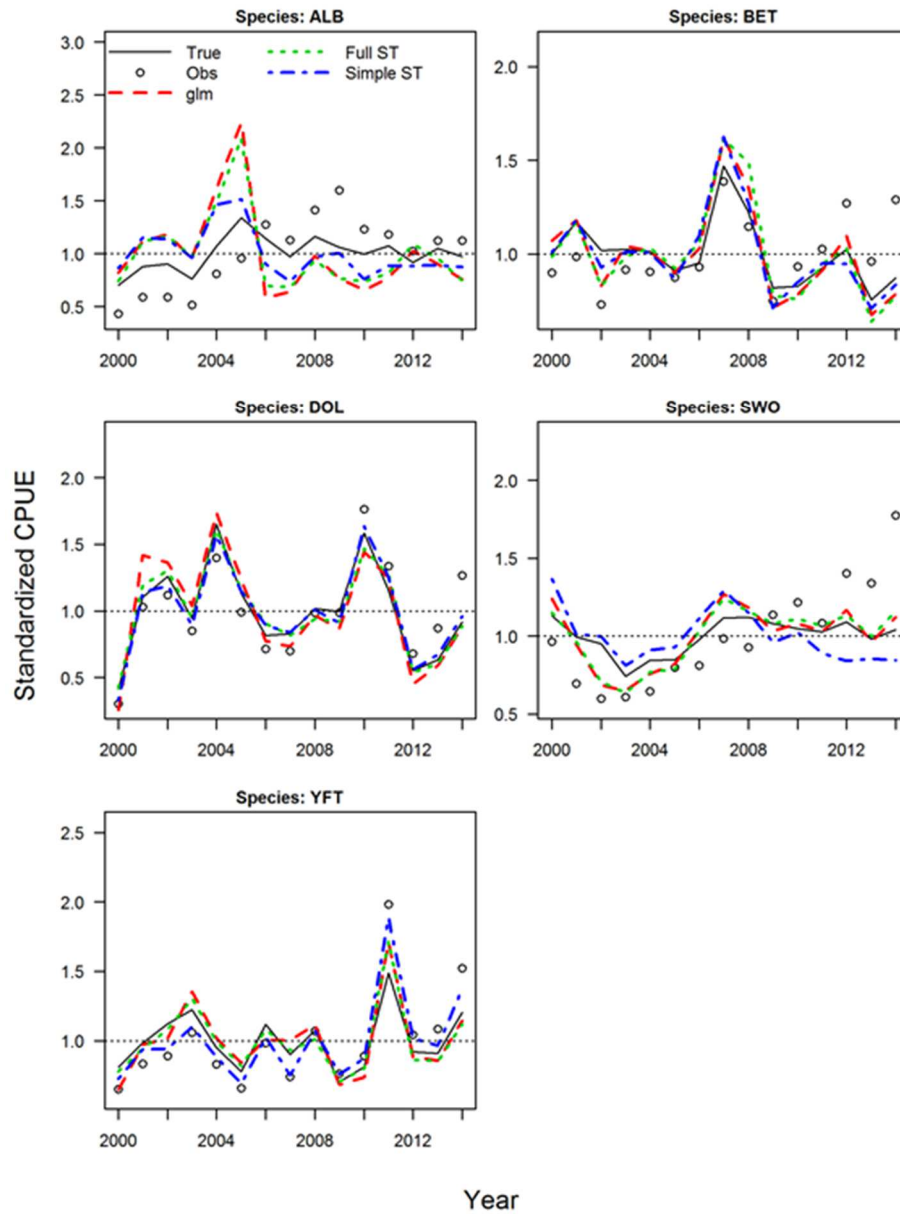


Table 5.5: Comparison of absolute relative error between predicted abundance index and true index across 6 models and 5 species.

Model	ALB	BET	DOL	SWO	YFT	mean
1	0.303	0.070	0.118	0.079	0.083	0.130
2	0.277	0.075	0.141	0.104	0.082	0.136
3	0.305	0.075	0.128	0.097	0.095	0.140
4	0.236	0.075	0.068	0.132	0.117	0.126
5	0.267	0.075	0.049	0.072	0.053	0.103
6	0.194	0.056	0.056	0.112	0.105	0.104

Table 5.6: Comparison of absolute relative error between predicted abundance index and observed index across 6 models and 5 species.

Model	ALB	BET	DOL	SWO	YFT	mean
1	0.630	0.184	0.182	0.317	0.159	0.295
2	0.601	0.191	0.192	0.278	0.159	0.284
3	0.615	0.186	0.202	0.282	0.172	0.291
4	0.563	0.195	0.181	0.485	0.066	0.298
5	0.577	0.192	0.164	0.300	0.155	0.278
6	0.531	0.181	0.117	0.480	0.064	0.275

Compared with the observed abundance index and based on the mean absolute relative error *MARE*, the standardized indices from the simple spatio-temporal Model 6 is most accurate (the best for 4 out of 5 species) (Table 5.6). Modelling all observed point locations at continuous temporal scales is one of the advantage of Model 6. Also, by treating spatial and temporal variables as random another advantage is its shrinkage estimator, which push extreme values from sub-groups (e.g. at finer scales of sub-area and month) towards the annual mean. However, this shrinkage does not reduce inter-annual variation for the index as Year is modelled as a fixed effect in Model 6.

5.6 Development of Multispecies Catch and Effort Simulators

Simulators are useful for comparing the performances of different methods at estimating known quantities. However, in complex multispecies fisheries it can be difficult to understand the factors that drive the observed dynamics, including catch rates. Different methods for analysing the data may be affected by different features of the real world, so simulators should be flexible to allow investigators to explore and model the effects of a range of real-world features. We have therefore developed several different simulators that address different issues.

Developing a data simulator for testing analytical methods which attempt to recover trends in the abundance of the species used when generating the data is a highly non-trivial task. This is because any simulator developed for this purpose can quite easily show bias if the data structures incorporated into the simulator match the analytical framework of one (or more) of the methods to be tested. For example, if one generated data on a spatio-temporal framework which was the same as the spatio-temporal framework of the analyser, with additional linear terms to account for the influence of different gear settings, then one would be surprised if that analyser (such as a GLMs) could not recover the abundance trends. In order to help overcome this problem we have attempted to develop a generic data simulator which is based on observed data inputs instead of model-based assumptions.

5.6.1 Empirical-Data Approach

Outline of Approach

Catch rates are known to be influenced by several factors other than general abundance and knowledge of those factors which influence CPUE underlies the statistical analysis known as effort standardisation. Most analyses make use of Generalised Linear Models (GLMs) or Generalised Additive Models (GAMs) which are well known statistical methods for estimating the relationship between a dependent variable (CPUE) and a number of independent variables.

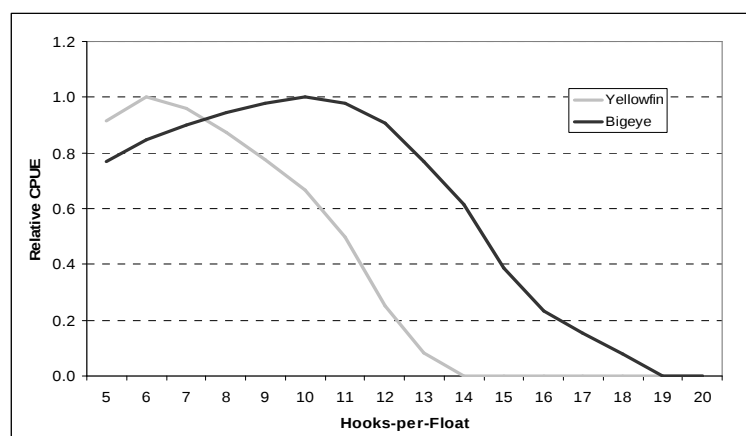
For standardising longline CPUE a number of variables are known to influence CPUE and, where the information is available, the following variables can be included in such analyses:

- Time stratum (e.g. Year-quarter)
- Spatial stratum (e.g. Area effect)
- Hooks-per-float (HPF)
- Bait-type
- Time of day
- Use of fish attracting devices (e.g.)
- Adoption of alternative fishing gears (e.g. hook-type, line material)
- Use of fish finding technologies (e.g. use of ocean temperature plots to identify fronts)

In fitting these variables within a GLM (or GAM) framework, the data is used to estimate the nature of the relationship between each independent variable and the dependent variable after controlling for the effects of the other fitted variables. This is a generally a purely statistical exercise as no process model is invoked to connect the two variables. However, one may use the results of such analyses to infer the nature of the underlying processes that contribute to the observed relationship.

For example, consider the illustrative GLM result shown in Figure 5.13, which displays the relative influence of the number of hooks-per-float on CPUE. From this result it is seen that the CPUE of Yellowfin Tuna is highest when 6 HPF are deployed and that CPUE decreases when a higher HPF is deployed. On the other hand, the CPUE of Bigeye Tuna is highest when 10 HPF are deployed. In order to understand this statistical result, we need to combine this result with the knowledge that (all other features constant) the number of hooks-per-float (HPF) is an indicator of the depths fished by the longline gear, with these depths generally increasing with the number of HPF (Suzuki *et al.* 1977). Hence, the GLM result can be used to infer that the number of Yellowfin Tuna available to the longline gear is highest at the depths fished by those gears set with 6 HPF, whilst the greatest number of Bigeye Tuna are found at the relatively greater depths fished with 10 HPF. The conclusion that gear fishing deeper in the water column is more effective in targeting Bigeye Tuna has been confirmed by a number of studies (Hanamoto 1987; Boggs 1992) and is thought to be due to a preference by Bigeye Tuna for cooler 10 to 15°C water (Holland *et al.* 1990; Brill 1994). In this manner, one can combine the knowledge about the depth distributions of both the hooks deployed by a longline and the target species to interpret, and understand, the purely statistical results of the GLM.

Figure 5.13: Hypothetical result of a GLM illustrating the relationship between the number of hooks-per-float and relative CPUE of both Yellowfin and Bigeye Tuna.



An alternative to the above statistical approach to standardising CPUE is to instead use our knowledge of the distribution of the target species and the fishing gears more directly. Indeed, this was the approach developed by Hinton and Nakano (1996), who presented a method of standardising CPUE which combined information on the spatial and depth distributions of the target species (using information on habitat preference and mapping of this habitat provided by oceanographic models) with information on the depths fished by longline hooks. This method was first applied to Pacific Blue Marlin, before being applied to Bigeye Tuna in the WCPO by Hampton *et al.* (1998). The method was further developed in the early 2000's and was applied to both Bigeye Tuna and Yellowfin Tuna within the context of the stock assessments undertaken for these species within the WCPO (see Langley *et al.* 2005). The approach is generally known as the habitat-based-standardisation (HBS) method.

Basic Equations

A form of the basic equation for applying the habitat model was presented by Hampton *et al.* (1998). However, as the actual derivation of this equation was poorly described in that paper, here we present a more complete description of the rationale behind the development of the required equations.

Consider the volume of water fished by the longline gear during a single set. From the catch equation the number of fish in the catch, C , is related to the total fishing effort, E , and the average fish population density in this volume of water, D as follows:

$$C = qED \quad (5.6.1)$$

where q is a fixed constant of proportionality known as the catchability coefficient and is related to the efficiency of the fishing gear. From this equation:

$$CPUE = \frac{C}{E} = qD = \frac{qN}{V} \quad (5.6.2)$$

where N is the number of fish and V is the volume of water fished. Without loss of generality, we assume this volume is divisible into n_d depth stratum each of depth d and cross-sectional area A and volume V . Let N_k be the number of fish within the depth stratum k so that the average density of fish within this stratum is $D_k = N_k/V$. If E_k is the effort (number of hooks) within stratum k , and q_k is the corresponding catchability, then from Eqn. (5.6.1) the catch, C_k , within stratum k is:

$$C_k = qE_k D_k = q_k E_k \frac{N_k}{V}$$

If q_k is considered to be a constant across all stratum then the total catch over all stratum is:

$$C = \sum_{k=1}^{n_d} C_k = \frac{q}{V} \sum_{k=1}^{n_d} E_k N_k \quad (5.6.3)$$

Finally, if $E_k = h_k E$ where E is the total effort (number of hooks) deployed and h_k is the proportion of these hooks within stratum k , and $N_k = p_k N$ where N is the total number of fish in all depth stratum and p_k is the proportion of these fish in stratum k , then Eqn. (5.6.3) can be expressed as follows:

$$C = \sum_{k=1}^{n_d} C_k = \frac{qEN}{V} \sum_{k=1}^{n_d} h_k p_k$$

When there are multiple species in the catch, the catch C_s of any single species s can be expressed as follows:

$$C_s = \sum_{k=1}^{n_d} C_{sk} = \frac{q_s E N_s}{V} \sum_{k=1}^{n_d} h_k p_{sk}$$

where N_s is the total number of fish of species s in all depth stratum and p_{sk} is the proportion of these fish in stratum k .

Finally, where the density of fish ($D_s = N_s/V$) varies both spatially and temporally, and assuming that the depth profiles of fish can vary hourly (to account for diurnal behaviour) but does not vary spatially (possibly unrealistic but simplifies the data required for the multispecies simulator), the last expression can be generalised so that the catch (C_{si}) of species s by longline set i within spatial area a , month m and deployed at time-of day (hour) t can be expressed as:

$$C_{si} = q_s E_i D_{sam} \sum_{k=1}^{n_d} h_{ik} p_{smtk} \quad (5.6.4)$$

where: E_i is the total number of hooks deployed by longline set i ,

D_{sam} is the density of species s within area a during month m ,

h_{ik} is the proportion of hooks for longline set i within depth stratum k , and

p_{smtk} is the proportion of the fish of species s during month m and hour t within in depth stratum k .

Note, Eqn. (5.6.4) assumes that the depth profile of the fish (and the hooks) does not change during the duration of the longline set. In practice the catch will be dependent upon an integration of these two profiles over the duration of the set, which could be as long as 6 to 12 hours, but again this assumption simplifies the data simulator. Indeed, the simulator assumes that the catch is dependent on the depth profile of the fish at the start time of the longline set.

Data Inputs

In order to use the above equations, information on each of the inputs required in the above equation was ascertained as follows:

- a) Hook depth profiles, h_{ik} . Direct observations of hook depths using time-depth recorders (TDRs) were used to estimate the proportion of time spent within pre-specified depth stratum by all hooks within various hooks-per-float (HPF) configurations. These observations were based on the 2050 individual TDR recordings obtained between August 2004 and May 2007 from vessels deploying longlines in the ETBF during a previous FRDC funded research project “*Determination of effective longline effort in the Eastern Tuna and Billfish Fishery*” (Campbell and Young 2010). Time-at-depth profiles for each HPF configuration observed were binned into 20m depth strata and the profiles for the following eight HPF configurations used in the data simulation were collated:
 - i) HPF=6 (combination of observed HPF=6 and HPF=7)
 - ii) HPF=8 (combination of observed HPF=8 and HPF=9)
 - iii) HPF=10 (combination of observed HPF=10 and HPF=11)
 - iv) HPF=12 (combination of observed HPF=12 and HPF=13)
 - v) HPF=15 (combination of observed HPF=14 and HPF=15)
 - vi) HPF=20
 - vii) HPF=25
 - viii) HPF=30

The observed profiles for each of these HPF configurations are shown in Figure 5.14. A LOWESS (LOcally WEighted Scatterplot Smoothing) function was fitted to each profile to help smooth the observed profiles (again shown in Figure 5.14 for two different smoothing parameters). Finally, the profiles selected for use in the multispecies data

simulator were the observed profiles for HPF=6, 8 and 10 and the LOWESS curves for HPF=12, 15, 20, 25 and 30 and are shown in Figure 5.15.

- b) Fish depth profiles, p_{smtk} . Direct observations of hourly depth profiles for four species (YFT, BET, SBT and SWO) were obtained from archival tags deployed on fish caught and tagged within the ETBF. These observations were based on tag deployments made during several projects (Gunn *et al.* 2005; Evans 2010; Evans *et al.* 2011, Basson *et al.* 2012). A hybrid profile, based on re-configuration of the profile for Yellowfin Tuna towards deeper depths, was also constructed in order to provide a fifth profile for use in the data simulator. A comparison of indicator depth profiles for each of the five species used in the multispecies data simulator is shown in Figure 5.16 while an example of the mean observed hourly depth profile (binned by 20m strata) for Yellowfin Tuna during both January and August is shown in Figure 5.17.

Figure 5.14: Observed hook-depth profiles for eight different hooks-per-float longline configurations. Fitted LOWESS curves are also shown for two different smoothing parameters.

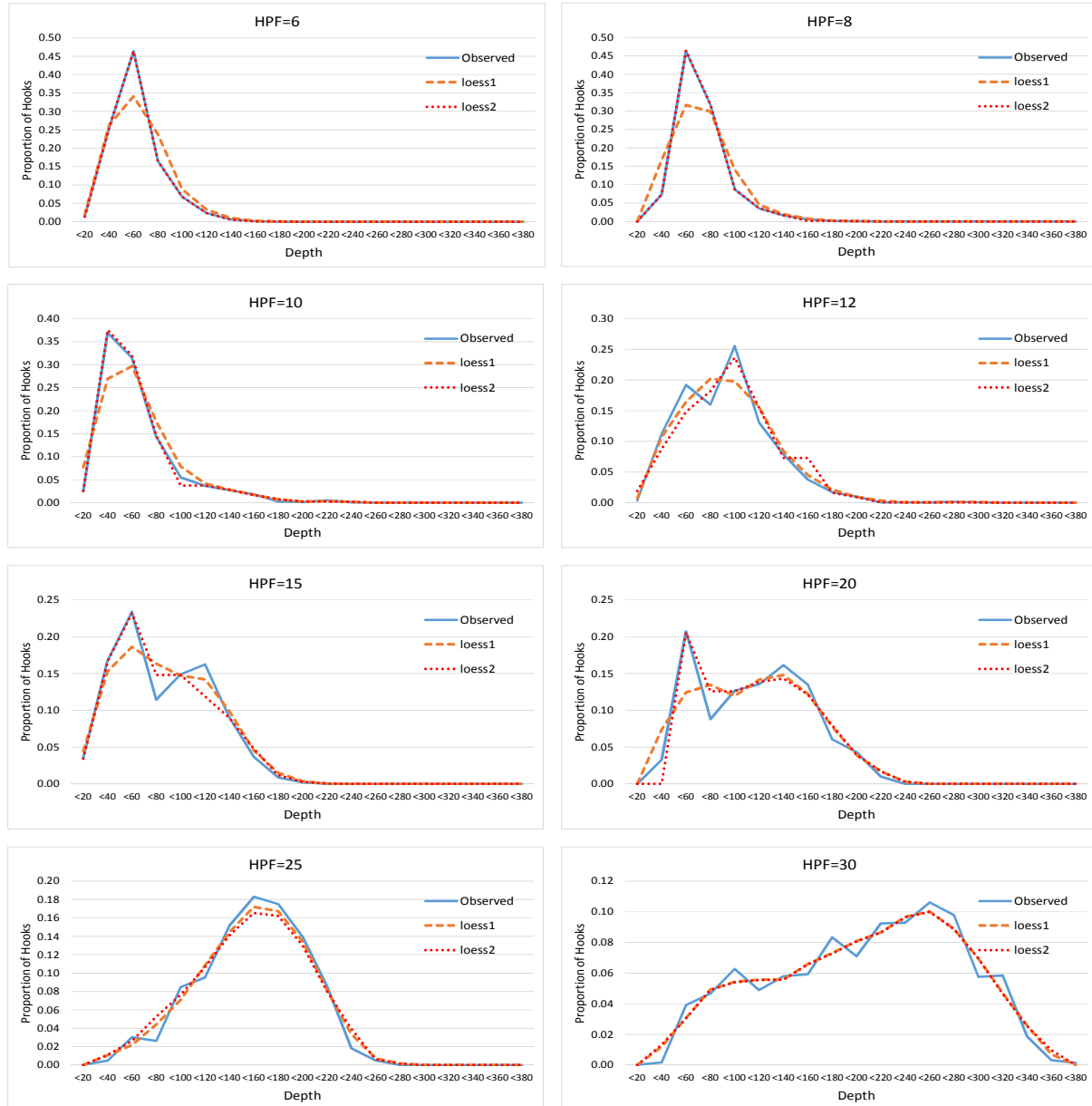


Figure 5.15: Hook-depth profiles used in the multispecies data simulator.

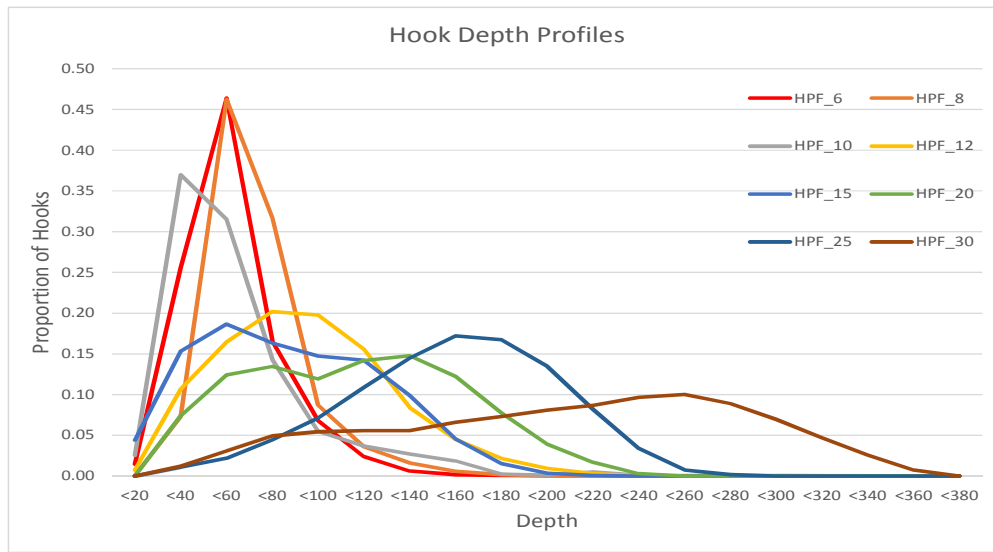


Figure 5.16: Mean observed depth profiles by month for YFT, BET, SBT and SWO based on data collected from archival tags. A fifth hybrid profile used in the multispecies data simulator is also shown.

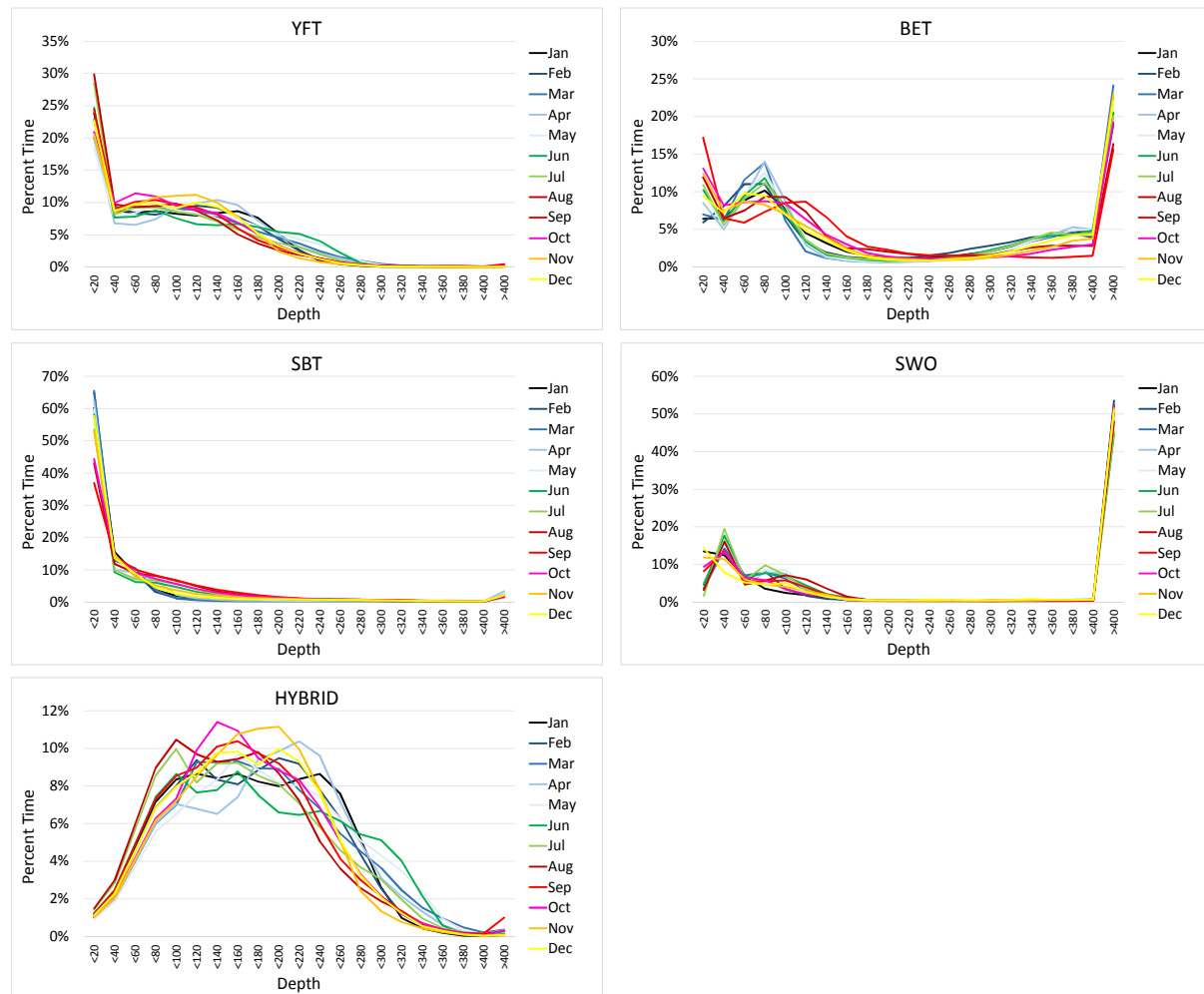
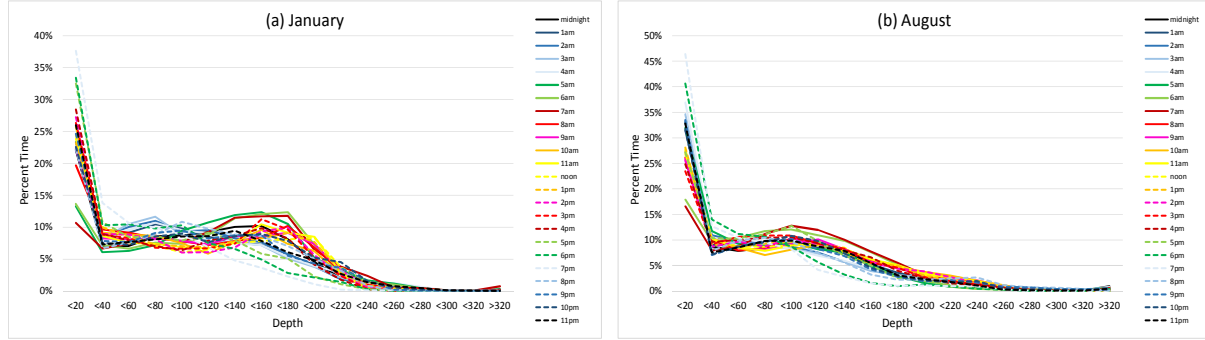


Figure 5.17: Hourly mean observed depth profiles for Yellowfin Tuna during (a) January and (b) August.



- c) Fish density distributions, D_{sam} : As with the depth profiles adopted above, each of these are based on *in-situ* observations, the spatial-temporal density surfaces for each species were modelled on the monthly distributions of nominal CPUE observed in the ETBF. The spatial scale used was a 1x1-degree square and the fishery (for the purposes of the simulator) was limited to the 139 squares that had been fished for a least 9 of the 15 years between 2000 and 2014. The monthly distributions of nominal CPUE were then smoothed using the Excel based spatial smoother RegularizeData3D (setting the smoother stiffness parameter to 3) found at:

<https://mathformeremortals.wordpress.com/2013/09/02/regularizedata3d-the-excel-spread-sheet-function-to-regularize-3d-data-to-a-smooth-surface/>.

An example of the density surface for Swordfish during June 2014 across the 139 1x1-degree squares incorporated into the multispecies data simulator is shown in Figure 5.18. Finally, the parameter D_{sam} was redefined as:

$$D_{sam} = N_{sy} d_{saym}$$

where N_{sy} is the total number of fish of species s across the fishery in year y , and d_{saym} is the proportion of fish within area a within year y and month m . The distributions d_{saym} were obtained by re-scaling the smoothed nominal CPUE such that within each month $\sum_{a=1}^n d_{saym} = 1$. The following five species were chosen for input to the simulator:

- i) Yellowfin Tuna (YFT) - matched with the YFT depth profile,
 - ii) Bigeye Tuna (BET) - matched with the BET depth profile,
 - iii) Albacore Tuna (ALB) - matched with the HYBRID depth profile,
 - iv) Broadbill Swordfish (SWO) - matched with the SWO depth profile, and
 - v) Dolphin fish (DOL) - matched with the SBT depth profile.
- d) Effort distribution, E_i . Finally, the spatio-temporal distribution of fishing effort across the fishery was taken to be the same as that observed in the ETBF during the years 2000 to 2014. While the logbooks for the ETBF record a number of factors relating to the manner in which the fishing gear is deployed (e.g. start-time, HPF, bait-type, use of light-sticks), for the purposes of the simulator the 'gear factors' were limited to just start-time (hour of the day) and the number of HPF (grouped to the list of eight HPF levels listed previously). Selected sets were restricted to the 139 one-degree squares chosen previously and to those sets where the number of hooks was greater than 200 and the HPF between 4 and 40. In total 113,711 sets were selected. Across the 139 one-degree squares constituting the fishery, the mean number of years fished per square was 14.0 and the number of longline sets per square averaged 818. The number of one-degree squares fished per month averaged 72 and ranged between 37 and 110. This indicates that the

Figure 5.18: An example of the density surface for Swordfish during June 2014 across the 139 1x1-degree squares incorporated into the multispecies data simulator.

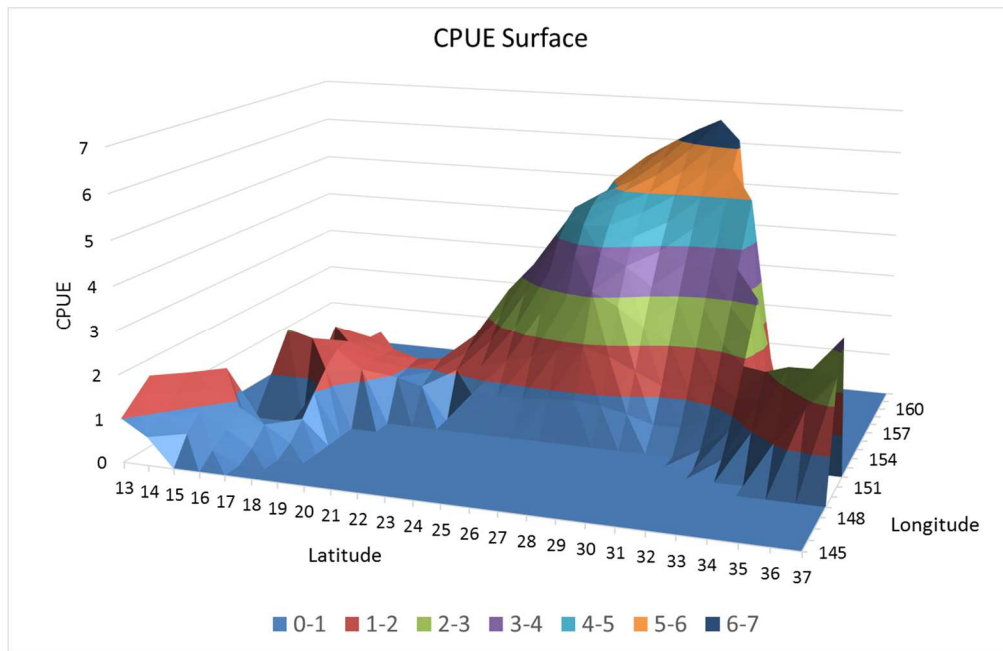
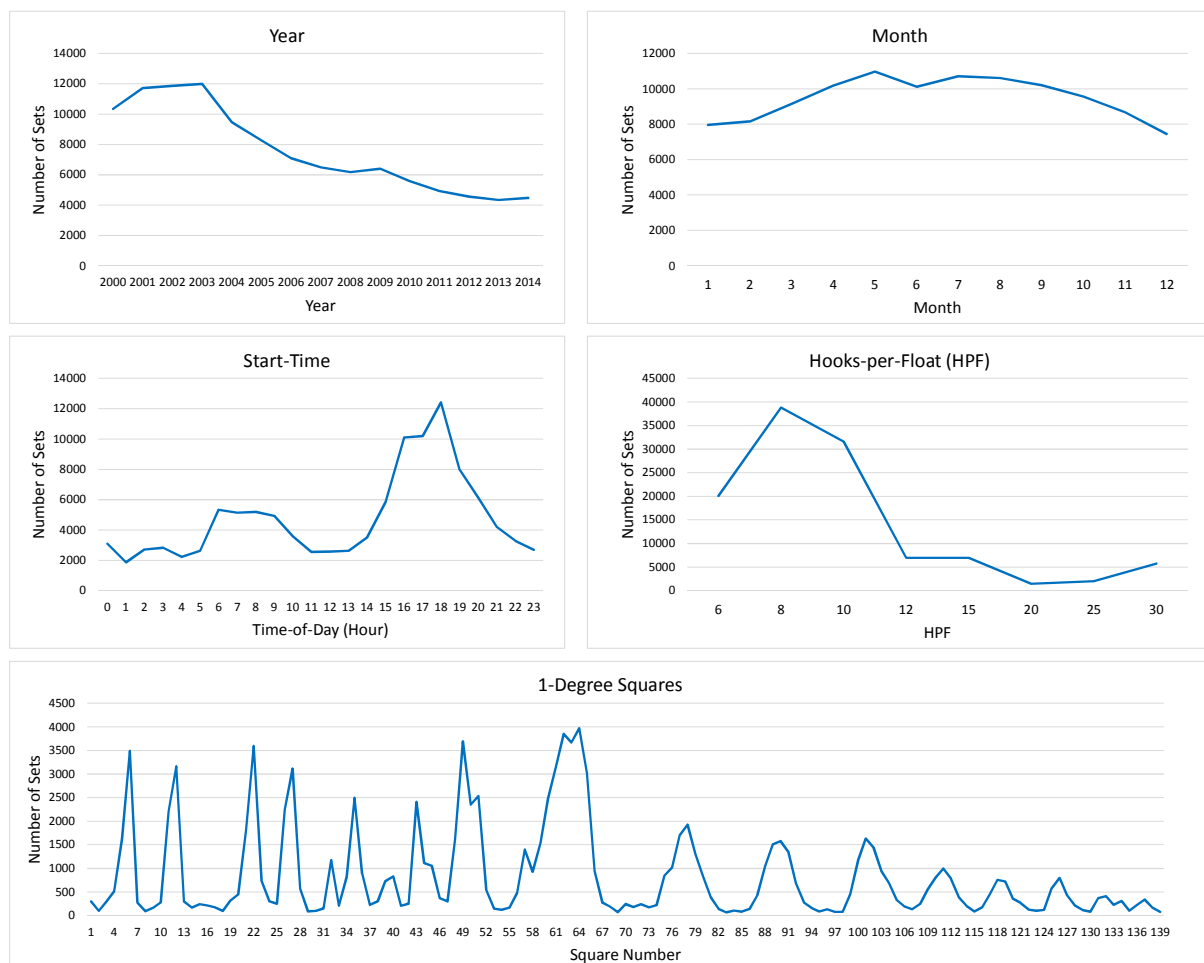


Figure 5.19: Distribution of sets used in the simulation across (a) Year, b) Month, (c) Start-time, (d) Hooks-per-float, and (e) 1x1-degree squares.



distribution of effort across the fishery was highly heterogeneous with potential for the spatial coverage to vary substantially between months. The distributions of the 113,711 longline sets used in the simulation across (a) Year, (b) Month, (c) Start-time, (d) Hooks-per-float, and (e) 1x1-degree squares are shown in Figure 5.19.

Simulated Catches

Inserting the reparametrized variable $D_{sam} = N_s d_{sam}$ into Eqn. (5.6.4) gives the final form of the catch equation:

$$C_{si} = q_s E_i N_{sy} d_{saym} \sum_{k=1}^{N_d} h_{ik} p_{smk} \quad (5.6.5)$$

With the four distributions h_{jk} , p_{kj} , d_{saym} , E_i characterised as described above, the final parameter values needed for calculation of the catch for each longline set is that of the catchability parameters q_s and the assumed annual abundance of each species N_{sy} . Towards this end, the values of q_s were chosen such that when $N_{sy}=1$ for all years, the mean simulated catch over all sets was similar to the mean observed catch. The values of N_{sy} for each species were then set to equal the annual values of the standardised CPUE for that species calculated from the logbook data (as described in Section 5.3) and re-scaled so that $\sum_{y=1}^{15} N_{sy}/15 = 1$. (Note, the approach described here means that the number of fish available to the fishery remains constant throughout a year, but the spatial distribution varies by month). Finally, in order to simulate the stochastic characteristic of any catch, the final simulated catch was selected from a negative binomial distribution, with the mean catch given by Eqn. (5.6.5) and the probability of success for each species selected so that the proportion of zeros in the distribution of simulated catches for each species was similar to that in the distribution of observed catches. Note, two sets of catches were simulated, the first set based on the deterministic mean catch given by Eqn. (5.6.5) (and known as the D-catches) and the second set based on those randomly selected from a negative binomial distribution described previously (and known as the R-catches).

A comparison of the distribution of simulated catches for the 113,711 sets with the distribution of corresponding observed catches is shown in Figure 5.20, while a comparison of the assumed annual abundance and the nominal CPUE based on the simulated catch and effort data for each species is shown in Figure 5.21. The percentage of sets deployed each year stratified by (a) start-time, (b) hooks-per-float, (c) longitude, (d) latitude and (e) month is shown in Figure 5.22.

Despite basing many of the parameters in the simulation model on a range of *in situ* observations made in the fishery, nevertheless the simulated catches generated are premised on a number of assumptions and likely simplifications about the distribution of fish within the fishery, the depths of the fishing gear, and the capture process. For example, it has been assumed that the depth profiles of the fish do not vary spatially which is unlikely, as these profiles will most likely vary according to changes in oceanographic conditions (e.g. the depth of the thermocline). The depth profiles fished by the various HPB settings of the longline is also assumed to be constant, though again these will be influenced by other factors such as how the line is deployed (e.g. line-setting speed of the vessel) and oceanographic conditions (e.g. current speed which will induce shoaling of the gear). The increased stochasticity introduced in the resulting catches by these and other factors not accounted for in the simulation model is assumed to be taken account of by randomly sampling the catches from the negative binomial distribution. To this end, the comparison of the simulated and observed catches shown in Figure 5.20 can provide some guidance as to whether the simulator can be deemed reasonable. Whether or not the simulated data passes the Turing Test (i.e. is the simulated data indistinguishable from the real data, Turing 1950) remains unknown, but the fact that the

Figure 5.20: Comparison of the distribution of simulated catches, both with (Random) and without (Determin.) the random component added, with the distribution of observed catches.

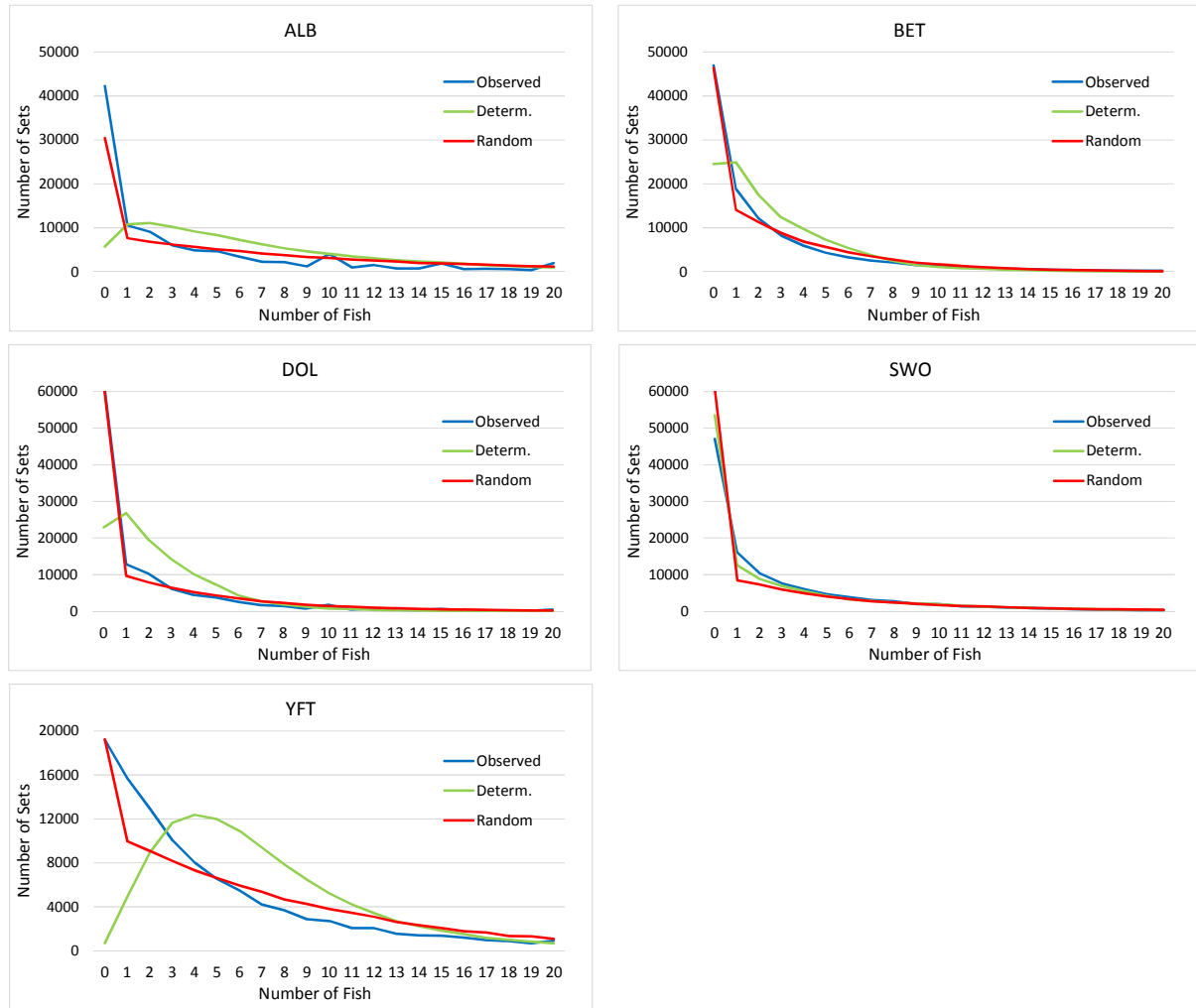


Figure 5.21: Comparison of the assumed annual abundance and the nominal CPUE based on the simulated catch and effort data (where both have been scaled so that the mean is one over the time-series).

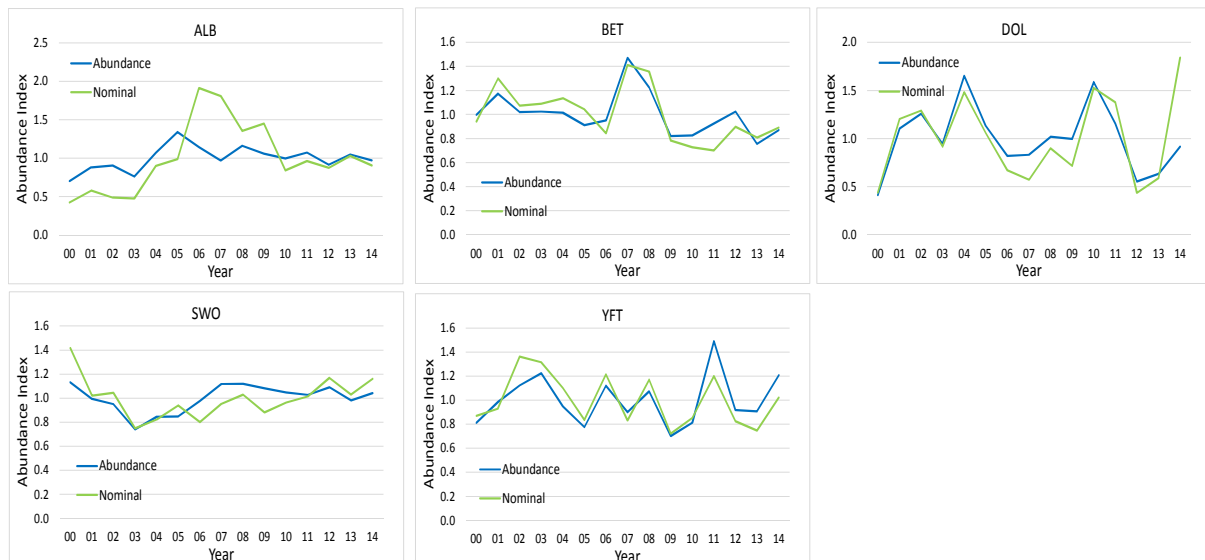


Figure 5.22: Percentage of sets deployed each year stratified by: (a) Start-time, (b) Hooks-per-float, (c) Longitude, (d) Latitude and (e) Month.



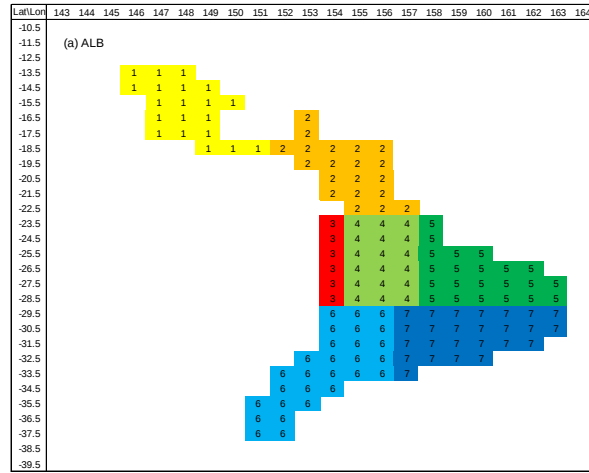
distributions of the two sets of catches are similar provides some reassurance.

Another feature of the simulated catch is that the N_{sy} trends chosen for each species represent only one hypothetical trend, with no underlying simulated population dynamics processes that they can be attributed to. While, these trends were chosen to simulate a distribution of catches and abundance trends that mimicked those observed in the fishery, nevertheless there is no reason to believe that these trends represent the true abundance trends of these species in reality. As such, in the following we will refer to the modelled abundance trends as the assumed relative abundances in the ETBF. While we have only modelled one assumed abundance trend for each of the five species used in the simulation, of course other trends could also be assumed, and this would allow the models used to standardise CPUE to be evaluated across a greater range of hypothetical conditions (assumed plausible).

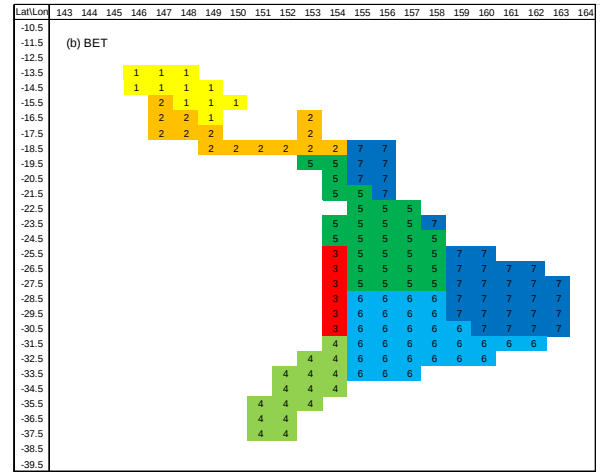
Finally, to model the spatial distribution of catch rates in the standardisation analyses, seven *Area*-effects were identified for each species. These *Area*-effects were determined according to the method outlined in Section 5.3 and are shown in Figure 5.23. The spatial distribution of the areas differ for each species and reflects the differences in the spatial distribution of the associated species. The distribution of the data records and fishing effort across these seven areas is shown in Figure 5.24, together with the distribution of the D-catches and associated catch rates for one realization of the simulated data sets.

Figure 5.23: Distribution of the 139 one-degree squares included in the simulated data sets across the seven *Area* effects used in the associated GLM analyses.

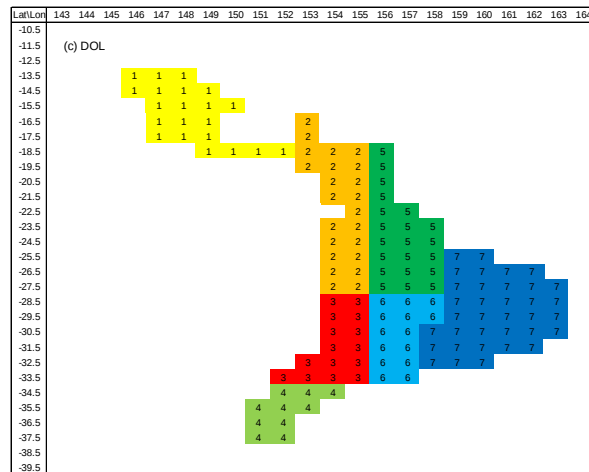
(a) ALB



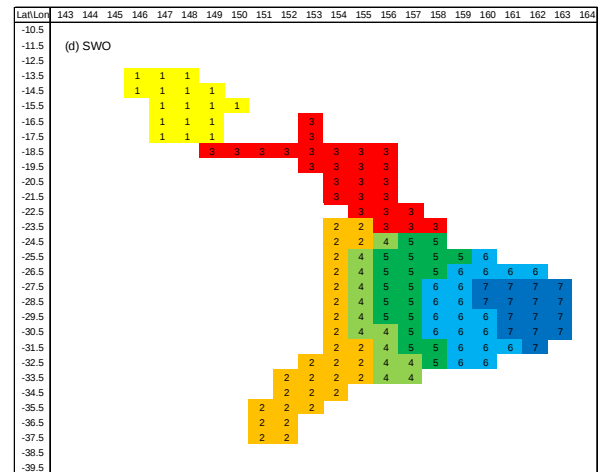
(b) BET



(c) DOL



(d) SWO



(e) YFT

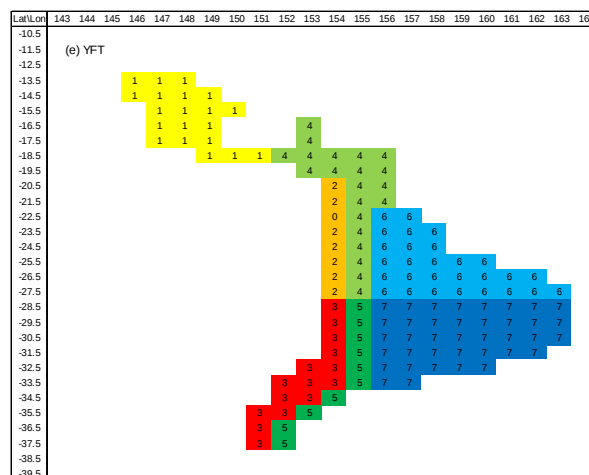


Figure 5.24: The distribution of the data records and fishing effort across these seven areas together with the distribution of the D-catches and associated catch rates for one realization of the simulated data sets.



Illustrative Analyses

The input parameter sets used in the above simulator are quite variable (with the density surfaces D_{saym} varying by 1-degree square, year and month and the fish depth profiles p_{smk} varying by month, hour and depth) due to the nature of the observed data upon which the parameter sets are based. This variability introduces a large degree of heterogeneity into the simulated data, both on a spatial and temporal basis, and this variability is also likely to introduce a challenge to any analyser chosen to recover the assumed annual abundance trends. In lieu of this challenge, a series of alternative data-sets were created where a number of aspects in the input parameter sets were altered in order to reduce the degree of spatial and temporal variability in the simulated data.

First, the density surfaces d_{saym} were altered such that the proportion of fish in each 1-degree square was set to the same value for all squares within the larger areas (usually seven) adopted as the spatial effects in the GLMs used to analyse the data. This value was the mean of the proportions across the 1-degree squares within each area. Furthermore, this value was also held constant on a quarterly basis, instead of the monthly basis in the original data. This change means that the proportion of fish within each spatial area varies only quarterly (i.e. by year and quarter).

Second, the fish depth profiles p_{smk} were altered such that the proportion of fish within each depth strata was held constant during each 4-hour period of the day. The value of this constant was set equal to the mean of the proportion across the four individual hours in the original data. Furthermore, this value was again held constant on a quarterly basis. This change means that the proportion of fish within each depth strata only varies each 4-hours and by quarter.

Four simulated data sets were generated based on whether or not the two alternatives above were used or not. These four data sets are:

1. Q-Q data-set Density surface varies by quarter and area.
Depth profile varies by quarter and 4-hourly.
2. Q-M data-set Density surface varies by quarter and area.
Depth profile varies by month and hour.
3. M-Q data-set Density surface varies by month and one degree square.
Depth profile varies by quarter and 4-hourly.
4. M-M data-set Density surface varies by month and one degree square.
Depth profile varies by month and hour.

Each of these data sets was fitted to the GLM described in Section 5.3 with the following four model variations fitted to the size of the positive catch:

Model GAM-2

$$CPUE = I + Year * Qtr + Qtr * Area + \sum \beta_j X_j / \text{dist} = \text{gamma link} = \log$$

Model GAM-3

$$CPUE = I + Year * Qtr * Area + \sum \beta_j X_j / \text{dist} = \text{gamma link} = \log$$

Model NEG-2

$$Catch = I + Year * Qtr + Qtr * Area + \sum \beta_j X_j / \text{dist} = \text{negbin link} = \log \text{ offset} = \ln(\text{hooks})$$

Model NEG-3

$$Catch = I + Year * Qtr * Area + \sum \beta_j X_j / \text{dist} = \text{negbin link} = \log \text{ offset} = \ln(\text{hooks})$$

Where I is the *Intercept*, *dist* refers to the assumed distribution of errors and *negbin* refers to the Negative-Binomial distribution.

The performance of each analysing model in recovering the assumed annual abundance index, when fitted to each of the four simulated data-sets, was evaluated by calculating the following log-relative error for each result:

$$abs(LRE_{s,y}) = \log \left(\frac{I_{s,y} / \frac{1}{ny} \sum_{y=1}^{ny} I_{s,y}}{T_{s,y} / \frac{1}{ny} \sum_{y=1}^{ny} T_{s,y}} \right)$$

where $abs(LRE_{s,y})$ is the absolute value of the log-relative error for species s in year y , $I_{s,y}$ and $T_{s,y}$ are the estimated and assumed abundance, respectively, for that species and year, $ny=15$ is the total number of years, such that $I_{s,y} / \frac{1}{ny} \sum_{y=1}^{ny} I_{s,y}$ and $T_{s,t} / \frac{1}{ny} \sum_{y=1}^{ny} T_{s,y}$ are the estimated and assumed indices of abundance after rescaling each to have a mean of one (i.e. given their treatment as relative indices of abundance). The mean annual error (*MAE*) for each species was also calculated as follows:

$$MAE_s = \frac{1}{ny} \sum_{t=1}^{ny} abs(LRE_{s,t})$$

The *MAE* for each of GLM models fitted to the above deterministic catch data set for each of the five species is provided in Table 5.7. The *MAE* for the nominal CPUE is also shown for

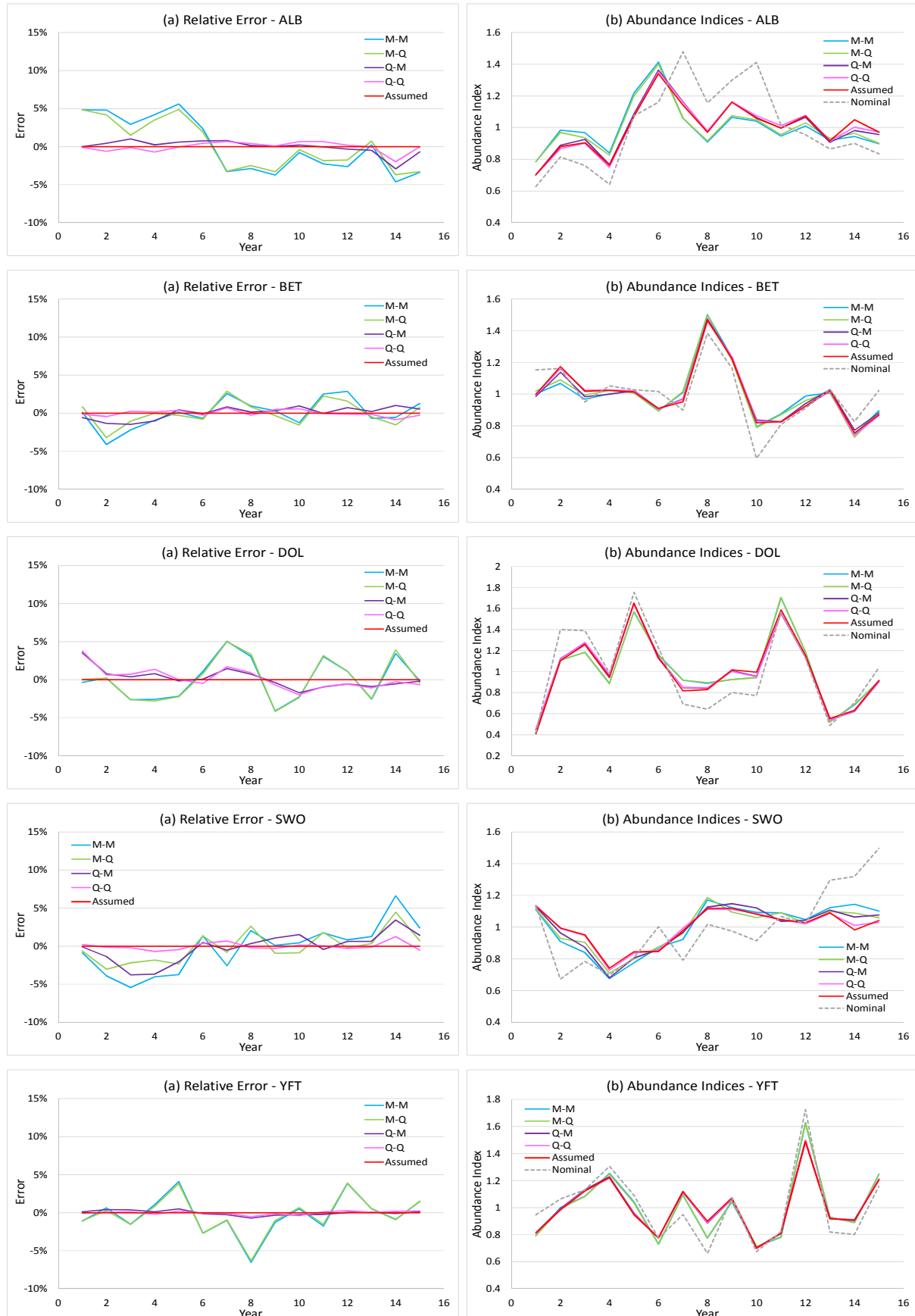
Table 5.7: Mean Annual Error (MAE) for each GLM fitted to the four versions of the simulated data sets. For each data-set and species, the GLM with the smallest MAE is highlighted in light grey.

Model	GLM	ALB	BET	DOL	SWO	YFT
M-M	NOMINAL	5.84%	3.36%	5.35%	6.75%	4.05%
M-M	GAM-2	6.44%	2.14%	2.94%	1.89%	3.20%
M-M	GAM-3	3.82%	1.37%	2.14%	2.28%	1.96%
M-M	NEG-2	6.39%	1.86%	2.21%	3.19%	2.29%
M-M	NEG-3	3.23%	1.41%	2.25%	2.49%	1.93%
M-Q	NOMINAL	5.68%	3.47%	5.32%	6.26%	4.03%
M-Q	GAM-2	5.93%	2.61%	2.85%	1.54%	3.10%
M-Q	GAM-3	3.20%	1.30%	2.23%	1.47%	1.88%
M-Q	NEG-2	5.89%	1.83%	2.29%	2.53%	2.27%
M-Q	NEG-3	2.77%	1.23%	2.31%	1.61%	1.84%
Q-M	NOMINAL	5.11%	3.93%	5.74%	5.60%	3.30%
Q-M	GAM-2	3.98%	1.71%	1.63%	1.20%	1.95%
Q-M	GAM-3	0.80%	0.67%	0.93%	1.56%	0.20%
Q-M	NEG-2	4.37%	1.63%	1.38%	2.43%	1.92%
Q-M	NEG-3	0.58%	0.65%	0.88%	1.43%	0.25%
Q-Q	NOMINAL	5.10%	3.97%	5.78%	5.31%	3.29%
Q-Q	GAM-2	3.57%	2.25%	1.67%	0.77%	1.78%
Q-Q	GAM-3	0.46%	0.54%	1.14%	0.36%	0.18%
Q-Q	NEG-2	3.81%	2.09%	1.41%	2.09%	1.99%
Q-Q	NEG-3	0.45%	0.34%	1.05%	0.38%	0.17%

comparison and is seen to be appreciably larger than the corresponding MAE for most of the GLM results (with the ALB results being the exception). The two GLMs incorporating the single 3-way interaction ($Y*Q*A$) are found to perform better than the two GLMs incorporating the two 2-way interactions ($Y*Q+Q*A$) with the relative performance varying between the different species. There is also some variation between species as to whether the best fit is found with the Gamma or Negative-Binomial distribution. However, for four of the species the results when fitted to the Q-Q data set indicate that the NegBin3 model performs best, with the MAE being less than 1.05% for all species (varying from 1.05% for DOL to 0.17% for YFT). This result is not unexpected, as the structure of the data (homogeneity within GLM areas and quarters) closely matches the structure of the model, and this result provides a check that the simulator and analyzing models are working correctly.

The time-series of the annual log-relative error for each species for the NEG-3 model fitted to each of the four data-sets are plotted in Figure 5.25a. The substantially smaller errors noted for the models fitted to the two data-sets, where the modelled density surface varies by quarter and area (instead of by month and 1-degree square), are clearly seen. This indicates that the spatial structure of the data has a greater impact on the ability of the model to infer the assumed annual abundance of a species than the depth profiles of the fish. For the model with the smallest degree of spatial and depth variation (the Q-Q model) there is still some residual error in the annual time-series, indicating that the GLM model is not capturing the full extent of the variation in the data. For example, the GLM assumes that the relative effect of the different hooks-per-float (HPF) categories on the data are constant across all spatial and temporal strata,

Figure 5.25: Annual time-series of (a) the log-relative error and (b) the relative abundance index for each species for the NEG-3 model fitted to each of the four data-sets.



whereas there may in fact be some variation in these relative effects. This could be investigated by fitting appropriate interaction terms but as the residual errors are small this will not be explored further here.

The resulting time-series of annual abundance indices for each species for the NEG-3 model fitted to each of the four data-sets are plotted in Figure 5.25b. These indices are compared with both the true modelled index and the nominal catch rate for each year. Again, the indices for the Q-M and Q-Q models are seen to be the closest to the assumed abundance index for each species. Furthermore, and perhaps more reassuring, the resulting indices for the GLM fitted to the M-M data set also are seen to be substantially closer to the true index than the nominal index. This is seen especially for the results for ALB and the SWO where there are large temporal biases in nominal indices. These results indicate that the current GLM used in the ETBF to standardise CPUE is performing reasonably well, and as such, the model structure is accounting for much of the structured variation in the CPUE within the data.

Finally, the annual time-series of the log-relative error and the relative abundance index for each species for each of the four GLMs fitted to the M-M data-set in Figures 5.26a and 5.26b respectively. For each species there are a number of issues to note:

- ALB: While the GLMs have removed the large temporal variation in the sign of the annual error noted in the nominal index (negative at the start of the time-series, positive in the middle years, then negative again at the end of the time-series), there remains a temporal discontinuity in the errors between the GLM-based indices and the true index (positive for first five years, then negative for the last nine years). While the size of the errors remains similar within each of these periods, taken over the entire time-series there is a negative trend in the relative errors. This indicates that the GLM-based indices have a higher-than-assumed relative abundance during the early period and a lower-than-assumed relative abundance during the latter period.
- BET: Each of the GLM-based indices performs reasonably well and over-comes some of the large biases noted in the nominal index (e.g. years 10 and 15). The GAM-3 and NEG-3 indices are seen to be similar and lack the temporal trend noted in the ALB indices. However, a small positive trend remains in these relative errors (0.2% per year for the NEG-3 index), indicating a slightly lower-than- assumed relative abundance during the early period and a slightly higher-than- assumed relative abundance during the latter period.
- DOL: Again, each of the GLM-based indices performs reasonably well and removes much of the temporal bias noted in the nominal index. For most years the relative error is less than 3% though as for BET there remains a small positive trend (0.13% per year) over the time-series.
- SWO: Apart from the first year, there is a large temporal trend (1.6% per year) in the bias between the nominal and true indices, implying that the nominal index infers a higher-than- assumed relative abundance during the early period and lower-than- assumed relative abundance during the latter period. While this temporal bias remains in each of the GLM-based indices, it is substantially reduced, being 0.5% for the NEG-3 index.
- YFT: As for the results for BET, the GLM-based indices perform reasonably well and over-comes some of the large biases noted in the nominal index (e.g. years 1 and 8). The small negative trend in the relative errors (-0.5% per year) noted in the nominal index is also substantially reduced (being 0.07% for the NEG-3 index). The index nevertheless, still displays some short-term biases, particularly for the years before and after year 8.

While the standardised CPUE indices displayed in these illustrative analyses are seen to

Figure 5.26: Annual time-series of (a) the log-relative error and (b) the relative abundance index for each species for each of the four GLMs fitted to the M-M data-set.

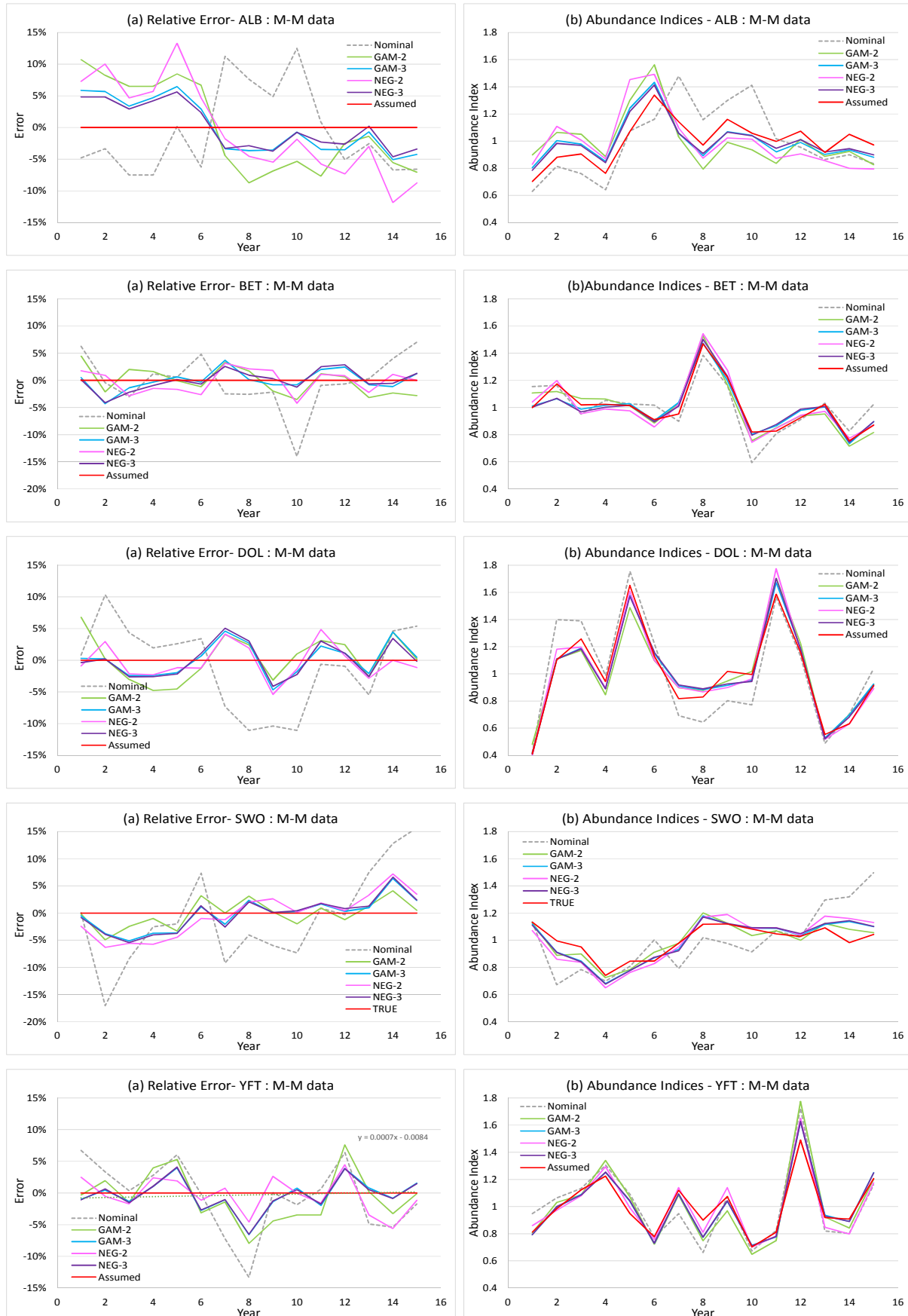


Figure 5.27: (a) Distribution of the annual abundance index for each year based on 100 random analyses with the mean indicated by the solid black line (and assumed index shown by the red line). The mean of the nominal CPUE for each simulated data set is shown by the dashed grey line. (b) Distribution of the error for each year with the mean indicated by the solid black line.



perform reasonably well (i.e. remove some of the more substantive biases seen in the nominal indices), as noted above, there are still a number of annual and temporal biases remaining in these indices. The reasons for these biases and attempts to identify improved methods to standardise the CPUE will be the focus of the results outlined in the next chapter.

Further Testing of GLM Models

Further testing of the GLM methods was undertaken by replicating 100 simulated data-sets from the M-M data-set, where the catches were randomly sampled from the negative binomial distribution. Again, the performance of each estimator was calculated using the log-relative error for each year and replicate. The distributions of the annual abundance indices calculated for each year and data replicate using the NEG-3 method are shown in Figure 5.27a, together with the assumed annual abundance and the mean nominal CPUE and the GLM-based index across each of the data replicates. The distributions of the log-relative error (LRE) between the estimated and true annual index for each year and data replicate are shown in Figure 5.26b, together with the mean across all data replicates.

The main trends in the results are similar to those previously discussed for the deterministic catches, and again vary to some extent across the five species shown. For ALB, the mean of the estimated annual indices (c.f. Figure 5.27a) is higher in the earlier years (before 2004) than the true index, but closely follows the trend of the true index after this time. This result indicates that the estimator, to a large extent, is able to account for the large shift in fishing practices in 2006 when there was a shift to deeper sets (using more HPF, c.f. Figure 5.22b) and concomitant higher catch rates of Albacore Tuna (as indicated by the high nominal CPUE indices after this time). The reasons for the poorer performance of the estimator before 2004 (when the LRE is as high as 10%) remains uncertain and warrants further investigation.

The results for BET are encouraging. Across all replicates, the mean annual $\text{abs}(\text{LRE})$ for each estimator is generally small (4%, c.f. Figure 5.27a), while for all estimators the mean across all years is less than 2.2% (Figure 5.27b). The mean annual index over all replicates also closely matches the true annual index and generally performs well in those years when there are large discrepancies between the nominal CPUE and the true index (e.g. 2009 and 2014, c.f. Figure 5.27a). On the other hand, the results for DOL are generally poorer, with the error between the estimated and true index for individual years being as high as 20% in 2000 and averages around 5% across all years and replicates. Nevertheless, the mean annual index over all replicates is similar to the true annual index and again accounts for discrepancies between the nominal CPUE and the true index.

The results achieved for SWO are the best of the five species, where the error between the estimated and true index for all but one years is less than 3% and averaging around 1.5% across all years and replicates (compared to the average of around 7% for the nominal CPUE). The estimated annual index closely matches the true index, and it is quite encouraging to see that the estimated index does not follow the increasing trend seen in the nominal CPUE, especially during the last few years.

Finally, the results for YFT are similar to that for BET. Again, the estimated annual index closely follows the true index and generally performs well for those years where the nominal CPUE shows a greater discrepancy, though the relative performance is slightly poorer given that the mean error across all years and replicates is closer to 3.2% than the 2.0% achieved for BET.

5.6.2 Second Simulator: Use of Random Distributions

We developed a second simulator that modelled catch and effort data with a generic, flexible and individual vessel based approach. The simulator was designed to generate catch rate data that reflected species abundance, targeting practises, and (unlike the ETBF simulator) individual vessel efficiency to capture the fact that different vessels have their own characteristic catchability and fishing behaviour. There is also spatial and seasonal variation in catch rates, in a species specific manner.

The methods often used to investigate fishing strategies, such as cluster analysis and principal components analysis, assume that fishing strategies can be inferred from the proportions of each species in the catch. This simulator is designed to explore the reliability of this assumption when independent factors, such as location and fishing strategy, also affect species-specific catch rates.

This simulator may also be used to evaluate management strategies, if adapted to include feedback from fishing mortality to abundance.

General approach

The simulator has multiple components: the spatial domain, the fish population dynamics, and the fishing process. The spatial domain is modelled as a grid $loc_{x,y}$ of n_x by n_y cells, with n_x and n_y initially set to 10. The model covers a period of nt years, with nt initially set to 20 and each year made up of 12 months.

Fish population

The fish population is modelled as ns species, with ns initially set to 10. The biomass $B_{s,t}$ of species s at time t is modelled very simply as an exponential trend through time:

$$B_{s,t} = B0_s \cdot e^{r_s t}$$

Initial biomass $B0_s$ is created by sampling from a lognormal distribution with median M_s of either 50 or 200, such that:

$$B0_s = M_s e^{dev}$$

where dev is a normally distributed random number with mean 0 and standard deviation 0.5. Population growth rate r_s is sampled from a uniform distribution between -0.1 and 0.1. Note, as with the ETBF simulated data there is no need to model how the catches (or any other processes) impact on the fished population of any species, as all we need to know is the assumed size of each population at any time from which the simulated catches are sampled. The known size of each population over time is also used to calculate the assumed abundance index for each species.

Each species is distributed across space with a constant proportion $p_{s,x,y}$ of the population in each cell. The spatial distribution $p_{s,x,y}$ has three options. Firstly, the *random* approach samples density by cell from a uniform distribution between 0 and 1. Secondly, the *ew-trend* approach samples the densities of the westernmost and easternmost longitudes from a uniform distribution between 0 and 1. Density at intermediate longitudes follows a uniform linear trend. Densities are the same at each latitude. Thirdly, the *ew-random* approach generates densities from the *ew-trend* and *random* approaches, then takes the inverse logit of $\text{logit}(ew-trend) + \text{logit}(random)/5$. For each of the three approaches, density is normalized so that $p_{s,x,y}$ sums to 1 for each s , giving density in each location $loc_{x,y}$ as:

$$density_{species,loc,t} = B_{species,t} \cdot reldensity_{species,loc}$$

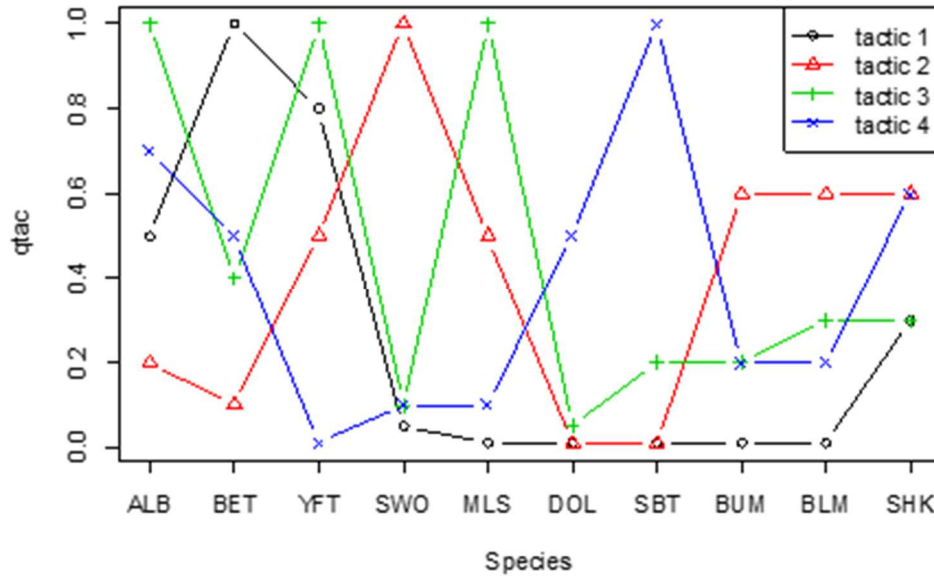
The species are labelled for convenience as ALB, BET, YFT, SWO, MLS, SBT, DOL, BLM, BUM, and SHK.

The catchability of each species $q_{mon_{sp,mon}}$ varies seasonally by month within each year. The monthly catchability variation of each species is based on the relative monthly catch rates for the equivalent species in the ETBF (c.f. Figure 6.1).

Fishing Effort

The unit of effort is the set. Fishing effort is distributed among $ntac=4$ sets of fishing tactics. Tactics may differ in a number of ways, such as the spatial distribution of effort and the relative catchability of each species. Initially the only feature implemented is relative catchability by species, $qtac_{tactic,species}$. Tactics are allocated predetermined relative catchabilities (see Figure 5.28).

Figure 5.28: Relative catchability by species and fishing tactic ($qtac_{species,tactic}$), as assumed in the simulation model.



Fishing effort is associated with $n_{vess}=15$ vessels. Each vessel is assigned a catchability deviate qv_{vessel} that applies to all species. Vessel catchability deviates are sampled from a lognormal distribution with mean 1 and standard deviation 0.3.

Each vessel is assigned a preferred tactic. Each trip by that vessel is allocated a main tactic, and each set within the trip is assigned a tactic. In the initial implementation of the simulator there is no tactic variation or switching, with each trip and each set using the same tactic. Preferred tactics are allocated randomly to vessels by equal probability sampling with replacement.

Each vessel carries out $n_{trip_{vess,t}}$ trips in a year, with each trip comprising $n_{set_{trip}}$ sets. Sets within a trip occur on consecutive days. Initially each vessel has 8 to 12 trips per year, with the number for each vessel-year sampled from a uniform distribution, and each trip comprising 5 to 13 sets, with the number for each trip also sampled from a uniform distribution. Trips in each vessel-year occur at random non-overlapping times through the year.

Each trip begins in a randomly selected cell, and subsequent sets on a trip may move to an adjacent cell. Movement probabilities on the x axis are modelled as $pmove_x = (0.05, 0.9, 0.05)$ of respectively moving 1 cell left, not moving, or moving one cell to the right. The same probabilities apply on the y axis, sampled independently. For cells against a boundary, probabilities of moving outside the area are set to 0 such that $pmove_x = (0, 0.95, 0.05)$.

Catch

The catchability for each species associated with each set is a function of the fishing tactic, the vessel, and the season:

$$q_{s,vessel,month,tactic} = qtac_{species,tactic} \cdot qvess_{vessel} \cdot qmon_{species,month}$$

Expected catch by species per set is equal to:

$$expcatch_{species,loc,yr} = density_{species,loc,yr} \cdot q_{species,vessel,month,tactic}$$

Observed catches by species per set are sampled from a Tweedie distribution with parameters μ =expected catch, $p=1.3$, and $\phi=10$. Deviates have variance equal to $\phi \mu^p$, so that the data are very overdispersed compared to the Poisson distribution. Deviates are generated using the function *rTweedie* from the R package *mgcv* (Wood 2006). As the set is taken as the unit of effort, the CPUE (catch per set) associated with each set is equivalent to the catch. Alternatively, one can set the number of hooks deployed per set to be the same for all sets (e.g. 1000) and then calculate an associated CPUE based on the catch-per-hook as done for the simulated ETBF data.

Datasets

One-hundred data sets were generated each sampling from the random distributions described above. Each data-set contained between 28,000 and 30,000 records and for the remainder of this report these will be known as the DIST (for distributions) data-sets to distinguish them from the simulated ETBF data-sets described in the previous Section.

5.7 Models Selected for Comparative Analysis

5.7.1 Listing of Models

Based on:

- 1). the review of CPUE standardising methods outlined in Section 5.2,
- 2) further investigation of the cluster approach in Section 5.4,
- 3) development of the new spatio-temporal Bayesian models in Section 5.5, and
- 4) consideration of some of the pertinent features of the data relating to the ETBF,

the following standardising models were selected for comparative performance using the two simulated data-sets developed in the previous Section. The models selected were broken down into the following three groups with the analysis of all models within each group undertaken by separate analysers (R.C., S.H. and S.Z. respectively).

Group-A Models

Model: A-1 (Base Model)

This is the same as the NEG-2 model currently used for standardising CPUE in the ETBF and described in Sections 5.3 and 5.6. Specifically, the GLM fitted to the probability of obtaining a positive catch is:

$$\text{Prob}(\text{catch} > 0) = I + \text{Year} * \text{Qtr} + \text{Qtr} * \text{Area} + \text{HPF} + \text{STIME} + \text{logit}(\ddot{H})$$

/ dist= binomial link=logit

while the GLM fitted to the size of the positive catch is:

$$\text{Catch} = I + \text{Year} * \text{Qtr} + \text{Qtr} * \text{Area} + \text{HPF} + \text{STIME}$$

/ dist= negbin link=log offset=ln(hooks)

where for both models I is the *Intercept*, HPF and $STIME$ are the number of hooks-per-float and the start-time for each fishing operation (set) respectively and $\text{logit}(\ddot{H})$ is the function as described in Section 5.3. Following the protocol used in the CPUE standardisations currently undertaken in the ETBF (as outlined in Section 5.3) all effects were fitted as categorical variables (corresponding to the related categories defined in Table 5.1) and the model was fitted to the data using the SAS GENMOD procedure (SAS 2008).

Model: A-2 (Total Cluster by Set)

A cluster analysis was undertaken across all sets based on the species composition of the catch using the *Ward* clustering method (SAS 1999) outlined in Section 5.4. The untransformed catch composition based on all species included in the simulated data sets was used for the cluster analysis and the number of clusters selected was set to five (to correspond to the five principal target species in the ETBF). A cluster effect was then added to the linear models described in the Base Model (A-1).

Model: A-3 (Total Cluster by Trip)

This is similar to model A-2, except the cluster analysis was undertaken based on the species composition of the catch for each fishing trip (instead of set). Set level data is likely to contain high variability in species composition due to the randomness of chance encounters between fishing gear and schools of fish. This variability is likely to lead to some misallocation of sets using different fishing strategies. Aggregating the data tends to reduce the variability, and therefore reduce misallocation of sets.

Model: A-4 (Month Cluster by Set)

This is similar to model A-2 except the cluster analysis was undertaken in two stages. First, separate cluster analyses were undertaken for the data for each month of the year with each set assigned to one of four clusters. This gave a total of 48 clusters across all months. In a second cluster analysis, these 48 clusters were then clustered into five final clusters based on the species composition of each cluster. This approach (as illustrated in Section 5.4) was used to investigate the possibility that species associations may change (and as a result the manner that the sets will clustered based on their catch compositions) during the different months of the year. These changes, more fully explored in Section 6.3, may be due to seasonal changes in availability and distribution of the different species due to prevailing oceanographic conditions.

Model: A-5 (Bi-Month Analysis)

This model is an extension of the approach used in model A-4 (and uses the same as model A-2) except separate (cluster and GLMs) analyses are undertaken on the data for each the six 2-month periods in the data. The resulting annual indices obtained for each bi-monthly period are then averaged to give a final annual index across all the data. Again, this approach was used to investigate the possibility that the species association may change during the seasons of the year due to changes in both the availability of the different species and the prevailing oceanographic conditions (c.f. Section 6.3).

As the GLMs are fitted to the bi-monthly data, there is no need to include the *Quarter*-effect (and in particular the two associated interactions) in the linear models fitted to the data. Indeed, this reduces the linear model fitted to each stage just the sum of the following main effects:

$$y = I + Year + Area + HPF + STIME$$

While this may appear as a 'simplification' this approach has the further advantage that the gear effects (*HPF* and *STIME*), which are constrained in Models A-1 to A-4 to have the same relative effects across all spatio-temporal strata, are now free to have different relative effects within each of the bi-monthly periods. This is similar to including a *QTR*GEAR* interaction effect in the previous models. It is important to note, however, in taking the mean of the six indices corresponding to the six bi-monthly periods it is important to keep the scale of the standardised CPUE the same for each period. This can be achieved by ensuring that the same set of standardizing reference levels for the fitted *HPF* and *STIME* effects are used for each period, i.e. the annual index constructed for each bi-monthly period will need to be standardised against the same specific reference level for each of the main effects (e.g. using SAS this is the last categorical level defined for each effect).

Group-B Models

Model: B-1 (Base Model)

This is the same as model A-1 but without fitting the delta (Binomial) component, i.e. all sets (including those with a zero catch) are fitted to the following single GLM.

$$\begin{aligned} Catch &= I + Year*Qtr + Qtr*Area + HPF + STIME \\ / dist &= \text{negbin link=log offset=ln(hooks)} \end{aligned}$$

The model was fitted using the function *glm()* in the R statistical package (R Core Team 2014).

Model: B-2 (PCA-ns Model)

This model is based on the Principal Components Analysis (PCA) model initially developed by Winker *et al.* (2013, 2014) to identify groups in the data. In this method, the proportional species compositions are first transformed by taking the fourth root, in order to reduce the dominance of individual species. Principal components are estimated using the function *prcomp()* in the R statistical package which transforms the data by centering and scaling, so as to reduce the dominance of species with higher average catches. Centering is performed by subtracting the column mean from each column, and scaling is performed by dividing the centered columns by their standard deviations. The fitted model (a GAM) has the form:

$$\begin{aligned} Catch &= I + Year*Qtr + Qtr*Area + HPF + STIME + ns(PC1,3) + ns(PC2,3) + ns(PC3,3) \\ / dist &= \text{negbin link=log offset=ln(hooks)} \end{aligned}$$

where $PC1-PC3$ are the first three principal components of the variation in the catch composition and $ns(,3)$ is a cubic spline with 3 degrees of freedom.

Model: B-3 (PCA-linear Model)

This is a variation on the Model B-2 where the principal component variables are fitted with linear effects rather than cubic splines.

Model: B-4 (PCA-binomial- Model)

This is variation on the Model B-3 where the values of the principal component variables are transformed from continuous space to zeroes and ones. Values less than zero are set to 0, and values greater than zero are set to 1.

Model: B-5 (*kmeans*-untransformed Cluster by Set Model)

This model uses the base model B-1 and adds a cluster effect, where clusters are identified using the *kmeans* method applied to untransformed species proportions. The *kmeans* method minimises the sum of squares from points to the cluster centers, using the algorithm of Hartigan and Wong (1979). It was implemented using function *kmeans()* in the R stats package (R Core Team 2014). The number of clusters was determined after applying the function *nScree* from the R package nFactors (Raiche 2010), and taking the minimum of the numbers selected according to the Kaiser rule and the optimal coordinates test. At least three clusters were always used. The same number of clusters was used for all cluster-based methods.

Model: B-6 (*clara* Cluster by Set Model)

This is the same as model B-5 but the *clara* method is used to undertake the clustering, and data for each species are transformed by scaling and centring before clustering. The *clara* method is an efficient clustering approach for working with large data-sets (Kaufman and Rousseeuw 2009). It was implemented with the function *clara* in package cluster (Maechler *et al.* 2014), with samples equal to 100 and sample size equal to 250.

Model: B-7 (*kmeans*-transformed Cluster by Set Model)

This is the same as model B-5 but the data for each species are transformed by scaling and centring before clustering.

Model: B-8 (Ward Cluster by Set Model)

This is the same as model B-5 but the *Ward* method is used to undertake the clustering. The hierarchical clustering method Ward *hclust* was 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 and Legendre 2014). However, in practice the method gave similar patterns of clusters to the other methods, more reliably than ward.D2 in the cases we examined. Note, this model was generally not used because it takes too long to run in R.

For the next set of models (B-9 to B-12) we aggregated the data by vessel-month, and refer to this level of aggregation as a trip, assuming that individual vessels tended to follow a consistent fishing strategy during a month. One trade-off with this approach is that vessels may change their fishing strategy within a month, which would result in misallocation of sets.

Model: B-9 (*kmeans* Cluster by Trip Model)

This is the same as model B-5 but the data aggregated at the 'trip'- level was used to undertake the clustering.

Model: B-10 (*clara* Cluster by Trip Model)

This is the same as model B-6 but the data aggregated at the 'trip'-level was used to undertake the clustering.

Model: B-11 (*Ward* Cluster by Trip Model)

This is the same as model B-8 but the data aggregated at the 'trip'- level was used to undertake the clustering.

Model: B-12 (*Ward* Cluster by Trip + Vessel Model)

This is the same as model B-11 but the vessel identifier was added as another categorical effect to the linear predictor. Note, an individual vessel effect was not used to generate the simulated ETBF data and as such a Vessel variable could not be included in the models fitted to these data.

Group-C Models

Model: C-1 (Base Model)

This is the same as model B-1 but with the $\ln(\text{hooks})$ term fitted as a linear effect instead of an offset. The fitted GLM has the form:

$$\text{Catch} = I + \text{Year} * \text{Qtr} + \text{Qtr} * \text{Area} + \text{HPF} + \text{STIME} + \ln(\text{hooks})$$

/ dist= negbin link=log

The model was fitted using the function *glm()* in the R statistical package (R Core Team 2014).

Note, use of the offset term is equivalent to setting the coefficient to 1 for $\ln(\text{hooks})$ term and is equivalent to fitting to CPUE on the left hand side of the equation (which is not possible with a negative binomial distribution which requires a discrete response). The use of the offset preserves the assumed 1-to-1 relationship between catch and effort (i.e. doubling hooks doubles the catch per set) and uncertainty in the relationship between catch and effort goes into the residual error, rather than the effort error. If the true relationship is not 1 to 1 then one can explore the nature of the true relationship by adding $f(\text{effort})$ as an effect in the linear predictor. While the relation between catch and effort (and therefore the definition of CPUE) using this approach become more nuanced, this does not impact on the calculation of the abundance index, as the $f(\text{effort})$ will only be a scaling factor that will be cancelled out when the relative index is determined. However, there may be problems with this approach. For example, there may not be enough data to estimate the effort relationship. Also, when the data-sets are unbalanced, e.g. when cluster X consistently uses more hooks than cluster Y, the effort effect would be confounded with the cluster effect. The effort effect may also be confounded with the year effect, if hooks per set has increased through time (as has been observed in the ETBF).

Model: C-2 (INLA Model)

This model is the same as Model C-1 except that it is implemented in INLA (see Section 5.5).

Model: C-3 (INLA continuous spatial Model)

This is a Bayesian spatial model implemented in INLA. The fitted model has the form:

$$Catch = I + Year*Qtr + HPF + STIME + \ln(hooks) + f(s)$$

where $f(s)$ is spatially structured effects based on continuous latitude-longitude point locations where longlines were deployed. Note that the discrete area (region) is not used in this Gaussian Markov Random Field Model that uses the stationary Matern function for the spatial effect.

Model: C-4 (INLA discrete and continuous spatial Model)

This model is the same as Model C-4 except that it also includes an interaction term $Qtr*Area$ as in the Base Model.

$$Catch = I + Year*Qtr + Qtr*Area + HPF + STIME + \ln(hooks) + f(s)$$

Model: C-5 (INLA continuous spatio-temporal Model)

The difference between this model and Model C-2 is that it not only includes a random spatial effect but also a random temporal effect. The random spatial effect is modelled by the Matern function while the temporal effect is modelled as a first order autoregressive process AR(1). The fitted model has the form:

$$Catch = I + Year*Qtr + HPF + STIME + \ln(hooks) + f(s,t)$$

Model: C-6 (INLA discrete and continuous spatio-temporal Model)

This model is the same as Model C-5 except that it also includes an interaction term $Qtr*Area$ as in the Base Model. The fitted model has the form:

$$Catch = I + Year*Qtr + Qtr*Area + HPF + STIME + \ln(hooks) + f(s,t)$$

Model: C-7 (INLA simple and spatio-temporal Model)

This model is the same as Model C-6 except that it does not include the discrete Qtr or $Area$ effects. The fitted model has the form and is the simplest of the models in Group-C:

$$Catch = I + Year + HPF + STIME + \ln(hooks) + f(s,t)$$

Note: The performance of one further model was also investigated by this project. This model, known as Spatial Dynamic Factor Analysis (SDFA, Thorson *et al.* 2016), was described in Section 5.2.3. However, as it was only recently published (available online since April 2016), and as the initial version of the model required considerable lead- and run-time to get the model working on the large ETBF data set, it was not possible to include it in the above list of models to be analysed. Nevertheless, some analyses using the SDFA model have since been completed and the results are reported in Appendix D.

5.7.2 Model Analyses and Comparisons

The following sets of analyses using the models listed above and the two simulated data-sets were undertaken to compare the relative performance of each model.

- 1D-Trial Group-A and Group-B models were fitted to the ETBF simulated data-set using the deterministic D-catches.
This trial involved fitting each model once, as there is only one set of D-catches. Furthermore, in order to investigate the influence of the inclusion of the two gear

effects (HPF and STIM) on the performance of each model, two sets of analyses were completed. First, for each of the five Group A models and seven Group B models the analysis was undertaken with the gear effects included in all models. Second, a similar set of analyses were then undertaken with the two gear effects (HPF and STIME) removed from all models. These models are denoted G-Na where G and N are the model Group and model Number respectively.

- 1R-Trial Group A and Group B models were fitted to the ETBF simulated data-set using the randomly generated R-catches.
This trial involved fitting each model to 100 replicates of the data sets, each generating a different set of R-catches. As with the first set of analyses, two sets of analyses were completed with and without the gear effects included in the models.

- 2R-Trial Group A and Group B models were fitted to the DIST simulated data-set using the randomly generated R-catches.
This trial involved fitting each model to 100 replicates of the data sets each generating a different set of R-catches. As with the first set of analyses, two sets of analyses were completed with and without the gear effects included in the models.

- 3D-Trial Group A and Group C models were fitted to a subset of the ETBF simulated data-set using the deterministic D-catches.
As noted in Section 5.5, even though INLA is deemed faster than the more traditional MCMC approach, running a multispecies model that involves random effect terms with only a few covariates is still time-consuming. Indeed, fitting the Group-C INLA based models to the full simulated ETBF data-set proved impractical even with the use of CSIRO's High-Performance computers. To reduce the computing time and overcome this difficulty a subset of the ETBF simulated data was selected. This was achieved by limiting the number of data records to being three or less in all year, month and one-degree cells. This resulted in a data-set with 31,959 records and for the remainder of this report is known as the ETBF-N3 data-set.

As with the first set of analyses, two sets of analyses were completed with and without the gear effects included in the models.

5.7.3 Model Area Effects

For all models fitted to the ETBF simulated data-sets that included a categorical *Area* effect, the areas correspond to the seven GLM-areas shown in Figure 5.23. As noted previously, these *Area* effects were determined externally from the model by visual inspection of the distributions of nominal CPUE in the data. Furthermore, the same *Area* effects could be used in each analysis of the ETBF simulated data-sets, as the spatial-distribution of the fishing effort and the fished resource remained the same across all data-sets. Such an approach, however, was not possible with the DIST simulated data-sets, as the distributions of effort and the resource varied between simulations, and it would have been too time consuming to visually inspect the distributions across all 100 data-sets. Consequently, the following two different approaches were used to model spatial effects in the models fitted to these data-sets.

Group-A models

For each data-set, ten *Area* effects were identified using the following algorithm. First, for each year the nominal CPUE (total catch divided by total effort) was calculated within each of the 100 cells in the simulated data-set. Second, for each cell the mean annual nominal CPUE across all years was then calculated. Third, the distribution of the mean annual nominal CPUE across all 100 cells was ranked from highest (rank=1) to lowest (rank=100). Each cell was then allocated to one of ten *Area* effects using the following formula:

$$Area = 1 + floor \left[\frac{rank_i - 1}{10} \right]$$

where $rank_i$ is the ranking of the nominal CPUE of cell i , as determined in the third step above. This approach makes each *Area* a composite of ten individual cells (so each *Area* is the same size), though this approach does not guarantee that each *Area* is composed of ten cells that are contiguous. While this may be seen as unusual, and perhaps an undesirable feature of this approach, at the same time one can argue that all this is doing is aggregating together spatial cells which have similar characteristics, and which therefore may behave in a similar and coherent fashion over time – which one can also argue is exactly what a well-chosen *Area* effect is supposed to do.

Group-B models

A simpler approach was adopted for this group of models, where the *Area* effects were simply taken to be equivalent to the 100 individual spatial cells used in the simulation. However, due to the large number of *Area* effects using this approach no interactions with this effect were included in the linear predictor. The base model therefore took the following simple form:

$$Catch = I + Year + Qtr + Cell$$

5.7.4 Model Comparisons

After constructing the estimated abundance index, $I_{s,t}$, for species s in year t based on the model parameters, the performance of each model in recovering the assumed annual abundance index, $T_{s,t}$, when fitted to each simulated data-set was evaluated by first calculating the following four annual errors:

Error-1: Absolute Log-Relative Error

$$ALRE_{s,t} = abs \left[\log \left(\frac{\hat{I}_{s,t}}{\hat{T}_{s,t}} \right) \right]$$

Error-2: Squared Relative-Error

$$SRE_{s,t} = \left(\frac{\hat{I}_{s,t} - \hat{T}_{s,t}}{\hat{T}_{s,t}} \right)^2$$

Error-3: Absolute Error

$$AE_{s,t} = abs(\hat{I}_{s,t} - \hat{T}_{s,t})$$

Error-4: Absolute Relative-Error

$$ARE_{s,t} = abs \left(\frac{\hat{I}_{s,t} - \hat{T}_{s,t}}{\hat{T}_{s,t}} \right)$$

where $\hat{I}_{s,t} = I_{s,t} / (\frac{\sum_{y=1}^{ny} I_{s,t}}{ny})$ and $\hat{T}_{s,t} = T_{s,t} / (\frac{\sum_{y=1}^{ny} T_{s,t}}{ny})$ are the estimated and assumed abundance, respectively for species s and year t after rescaling to have a mean of one (i.e., given their treatment as relative indices of abundance) and where ny is the number of years in the analysis. The mean annual error (MAE) across all years for each error i and species s was then calculated as follows:

$$MAE_{s,i} = \frac{1}{ny} \sum_{t=1}^{ny} Error(i)_{s,t}$$

For those trials based on fitted to the R-catches and for which 100 replicate analyses were completed, the MAE for each model was taken to be the mean across all replicate analyses, i.e.

$$MAE_{s,i} = \frac{1}{100} \sum_{r=1}^{100} \left[\frac{1}{ny} \sum_{t=1}^{ny} Error(i)_{s,r,t} \right]$$

where r is the replicate number ($r=1, \dots, 100$).

For each species s and error i , the $MAE_{m,s,i}$ for each model m was tabulated then ranked in order with the best performing model (i.e. that with the smallest MAE) given a ranked score (RS) of 1 and the worst performing model (i.e. that with the largest MAE) given a ranked score of N where N is the number of models being compared. A total ranked score for each species (Species Ranked Score, SRS) was then calculated by summing the individual scores across all four errors:

$$SRS_{m,s} = \sum_{i=1}^{ne} RS_{m,s,i}$$

where ne is the number of error types (4). Finally, a total ranked score for each model (Model Ranked Score, MRS) was calculated by summing across all species:

$$MRS_m = \sum_{s=1}^{ns} SRS_{m,s}$$

The relative performance of each model was then determined based on the MRS value for each model (smallest = best model, largest = worst model).

Each of the above errors is reported where deemed appropriate.

6. Results

6.1 Review of Factors Influencing CPUE

From the discussion provided in Section 3.1, the relationship between CPUE and the true abundance of fish (B) within a given spatial region (A) is influenced by both the availability of the fish to the fishing gear (a) and the efficiency, or catchability, of the fishing gear (q):

$$CPUE = \frac{aqB}{A}$$

In this section we review some of the factors that influence either availability or catchability.

6.1.1 Availability

The availability of fish to the fishing gear will be influenced by a number of factors – principally the abundance of fish in the area being fished. This in turn will be influenced by a range of factors (apart from overall stock abundance), such as seasonal migration patterns of the fish and prevailing oceanographic conditions (e.g. currents and the characteristics of the surrounding water masses) that influence the spatial distribution of the habitats preferred by the targeted fish (and therefore the spatial distribution of the fish themselves). These influences are obviously beyond the control of the fisher. Given that fish are available in the area being fished, availability will also be influenced by the relationship between the depth distribution of the species being targeted and the depth distribution of the fishing gear.

For example, demersal fishing gears target fish species which live on or near the seafloor, and while the distribution of the depth above the seafloor for some of the target species caught by such gears may vary from time to time (possibly based on prevailing oceanographic conditions), the distribution of the depths fished by the fishing gears remains relatively constant. Consequently, the fish available to the fishing gear of a demersal trawl will be generally limited to those fish which are located between the seafloor and the head-height of the trawl net. On the other hand, for fishing gears which target pelagic species (such as pelagic trawl or pelagic longline), both the depth distribution of the target species and the depth distribution of the fishing gear can vary. Consequently, the nature of the catch will be dependent on the relationship at the time of the fishing operation between these two depth distributions. Indeed, an understanding of such relations based on observations made in the ETBF was used in generating the simulated ETBF data (c.f. Section 5.6.1).

Again, the depth distribution of the fish caught will be influenced by a number of factors, including the prevailing oceanographic conditions, which may influence the vertical distribution of the habitats preferred by the target species (e.g. depth of the thermocline), and possible behavioural characteristics of these species, such as diurnal vertical migrations (e.g. Bigeye Tuna prefer deeper depths during the day and shallower depths during the night). These factors again remain outside the influence of the fisher but various variables related to prevailing environmental and oceanographic conditions can be included in any GLM to model the influence of such factors on fish availability (c.f. Table 5.1). On the other hand, the depth distribution of the pelagic fishing gear is under the influence of the fisher, as it will be largely dependent upon the manner in which the fishing gear is deployed (though it will also be under the influence of prevailing currents). Apart from luck, an understanding of the factors which influence the vertical distributions of both the target species and the fishing gears is therefore required to achieve ongoing success in fishing operations.

6.1.2 Catchability

The ability (or effectiveness) of a given fishing gear to catch the available fish being targeted will also be dependent upon a range of factors. For example, the range of species and the size of fish caught are dependent upon the size of the mesh used by a trawl and the size of the hook used by a longline. The effectiveness of a fishing gear is, to a large extent, under the influence of the fisher, but will also be dependent to some extent upon their understanding of the size and behavioural preferences of the species being targeted. For example, the catch of Broadbill Swordfish increases with the use of light-sticks and squid bait on shallow sets deployed in the afternoon while the catch of Albacore Tuna shows a preference for pilchard baits on deep sets deployed in the morning.

Like most businesses, individual fishers will strive to improve the effectiveness of the fishing operations over time to increase the catch per unit of deployed effort (commonly known as effort creep). Towards this end, and based on their knowledge of the behavioural characteristics of the species being targeted, fishers will 'experiment' with the manner in which the fishing gear is deployed in order to maximize the overall effectiveness of the gear. To this end, the manner in which fishing gears such as a pelagic longline are deployed can vary considerably between vessels and will be dependent on the objectives of the fisher. In order to understand this variability, and attempt to standardise the resulting CPUE for these differences, it is important that the details pertaining to the manner in which the fishing gear is deployed is fully recorded in the vessel logbook. Unfortunately, for many fisheries much of this information is not captured in the logbooks, though for the Eastern Tuna and Billfish Fishery (ETBF) a range of such information has been recorded since 1997. This information includes: (i) start time of set, (ii) the length of the longline deployed, (iii) number of hooks deployed, (iv) the number of hooks-between-floats, (v) number of light-sticks deployed, and (vi) bait type (both species and life-status). Each of these factors (and several other derived factors, e.g. number of hooks per kilometer) is currently included in the analyses used to standardise CPUE in this fishery.

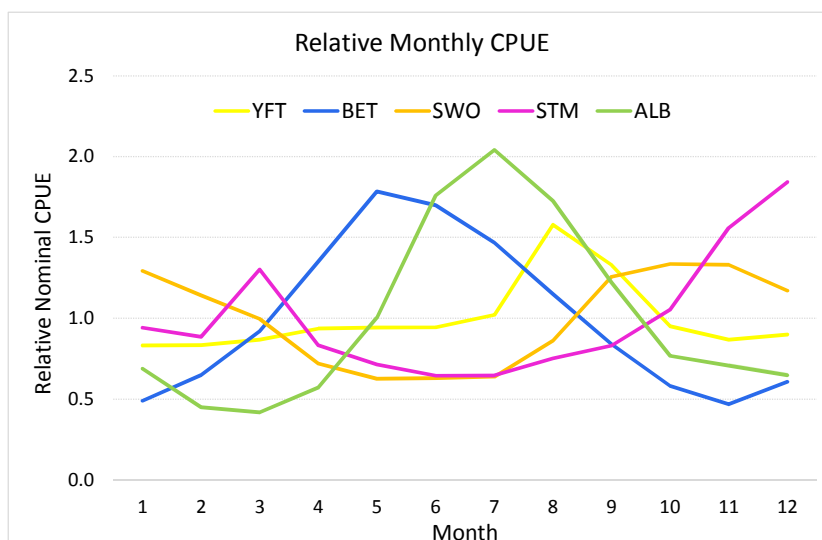
6.2 Targeting and Gear-Setting Practices in the ETBF

A short description of the ETBF together with level of longline effort deployed each year has been provided in Section 3.1 Due to changes in the level of operational data recorded in logbooks used in the fishery, all catch data for the ETBF was summarised by year and species to ascertain continuity of recording of catch by species on logbooks – see Appendix A. This information was discussed at the meeting of Tropical Tuna Resource Assessment Group held 1-2 October 2015 which concluded that the recording of several species in the logbook is unreliable (mainly for species which are predominately discarded). It was also considered best to combine the catch of shark species into several major groups.

6.2.1 Gear Configurations

Fishers can target different species in a number of different ways. If there are seasonal differences in the availability of different species, then fishers can target these different species by fishing at different times of the year. While each of the main species targeted in the ETBF are generally available all year around, there is some evidence that the level of availability for some of these species may change on a seasonal basis. For example, the monthly nominal CPUE for each of the primary target species in the ETBF (aggregated over the years 2006 to

Figure 6.1: Index of availability of the five primary target species in the ETBF as measured by nominal CPUE.



2013 and scaled such that the mean over all months is one for each species) is shown in Figure 6.1. From this Figure it appears that the availability of Bigeye Tuna is higher in autumn and lower over summer, the availability of Yellowfin Tuna is generally highest during spring, while the availability of Swordfish is highest during spring and summer.

Difference in availability can also be noted on a spatial basis, as evidenced by the spatial distribution of aggregate nominal CPUE across the ETBF. For example, the availability of Yellowfin Tuna is generally higher along the continental shelf off eastern Australia (c.f. Figure B.2a, Appendix B), possibly due to these fish being entrained in the Eastern Australian Current.

On the other hand, when all species are generally available during the same season and region, fishers can attempt to target the different species by alternating the manner in which they deploy their fishing gears. In order to understand and identify such practices, it is instructive to examine the gear settings deployed in the ETBF. For this purpose, we make use of the deployed gear information recorded in logbooks and by observers on vessels. While some gear setting information is available on logbooks completed by fishers, we make use of the observer data here as these data incorporate a greater range of information.

Information on longline gear settings deployed by vessels in the ETBF is recorded in the logbooks completed by skippers for each fishing operation. While very limited information was recorded on logbooks before 1997, the introduction of the AL04 logbook in 1997 resulted in a greater number of gear-setting information being collected. Since 1997 the number of vessels operating in the ETBF each year has varied between 173 in 1999 and 39 in 2015 (c.f. Figure 3.2a) while the number of longline sets deployed each year has varied between 13,227 in 2003 and 4,593 in 2013. AFMA observers have also been recording data on longline fishing operations in the ETBF since 2001, and data for 4990 observed sets and 1659 fishing trips were extracted from in the ETBF Observer database managed by CSIRO in Hobart.

Recorded fields for each fishing operation (FOP) include:

1. Vessel
2. Date of fishing operation
3. Location of FOP

4. Start-Time of FOP
5. Number of hooks deployed
6. Number of buoys deployed (observer only)
7. Number of hooks observed hauled (observer only)
8. Catch by species (retained, discarded)
9. Number of hooks-per-float
10. Number of light-sticks used
11. Use of line-shooter (observer only)
12. Line-setting speed (observer only)
13. Timer setting for attaching branch-lines (observer only)
14. Distance between branch-lines (observer only)
15. Bait species
16. Bait life-status
17. Primary target species (observer only)
18. Secondary target species (observer only)
19. Line configuration (observer only)

A number of additional fields are also recorded on a per-trip basis. These include:

20. Length of mainline
21. Length of buoy-line (observer only, up to 2 lengths)
22. Length of branch-line (observer only, up to 2 lengths)
23. Hook-type (observer only, up to 3 types)
24. Hook-size (observer only)
25. Hook-weights (observer only, up to 3 weights)
26. Presence of line-shooter on vessel (observer only)

The percent of longline sets deployed in the ETBF each year since 1997 deploying different categories of gear-types based on information recorded in ETBF logbooks is shown in Figure 6.2. For each gear-type varying degrees of change in usage are seen across the years. In particular, there was a significant change in the number of hooks-per-float (HPF) used in 2006 when vessels began deploying more than 25 HPF. Commensurate with this change, one can also see shifts in the bait usage (more pilchard), light-stick usage (fewer) and set start-time (earlier in the day). The introduction of this new 'deep-longline' technique, and other associated gear changes, was due to a significant change in fishing strategy to incorporate the direct targeting of Albacore Tuna in this fishery. Additional information on the distribution of the length of the mainline deployed across sets each year is shown in Figure 6.3, together with the mean number of hooks deployed per set. One can use this information to derive additional features relating to the deployed fishing gear, such as the distance between floats and the number of hooks deployed per kilometer of longline. Again varying degrees of change in usage are seen across the years, together with a substantial change in 2006.

Figure 6.2: Percent of longline sets deployed in the ETBF each year using different categories of gear-settings. Information based on ETBF logbooks. (NR denotes Not Recorded, Mack=mackerel, Pilch=pilchard, D=dead bait, A=live bait, M=mixed dead and live bait).

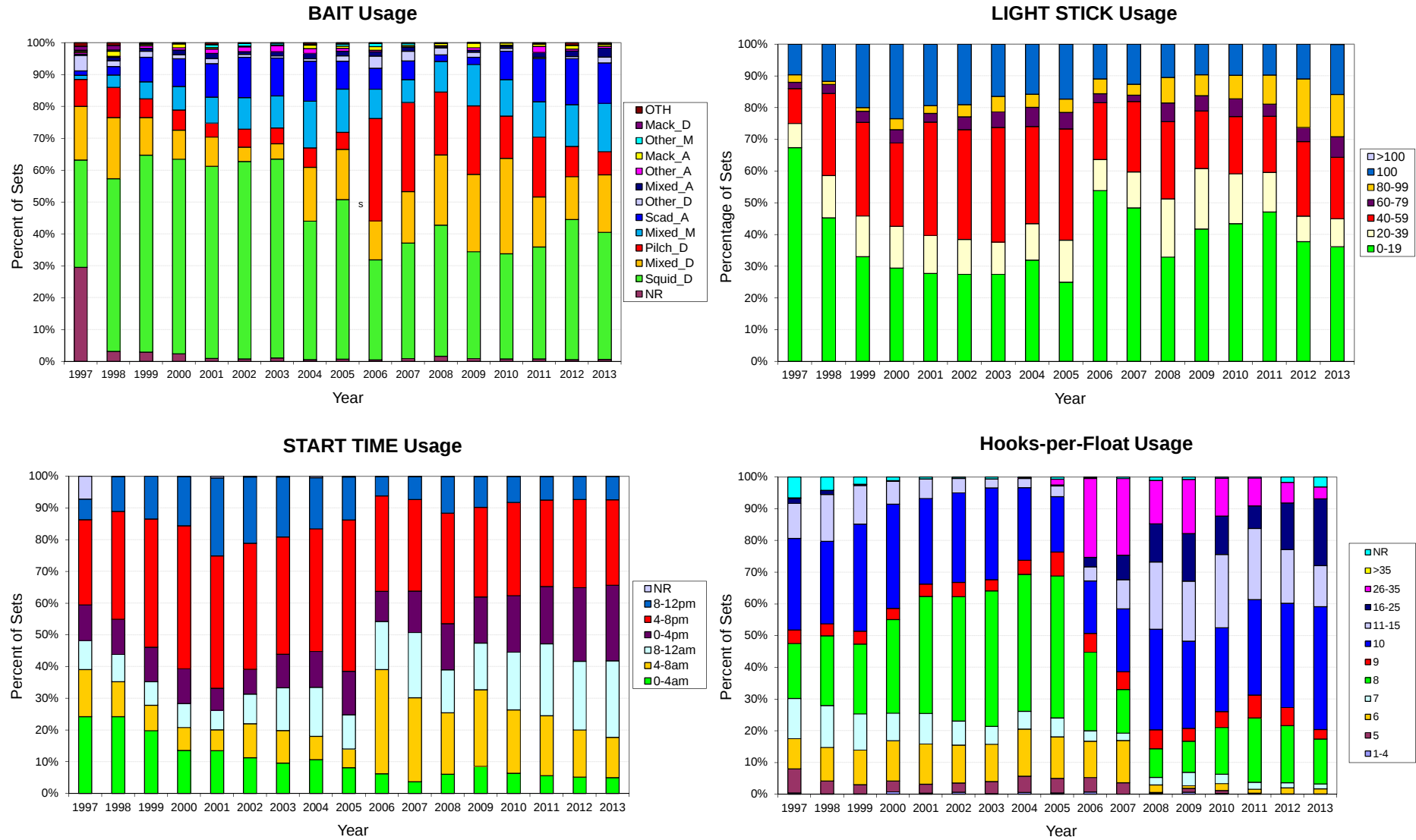


Figure 6.3: Percent of longline sets deployed in the ETBF each year using different categories of gear-settings. Information based on ETBF logbooks. (UNK denotes Unknown).

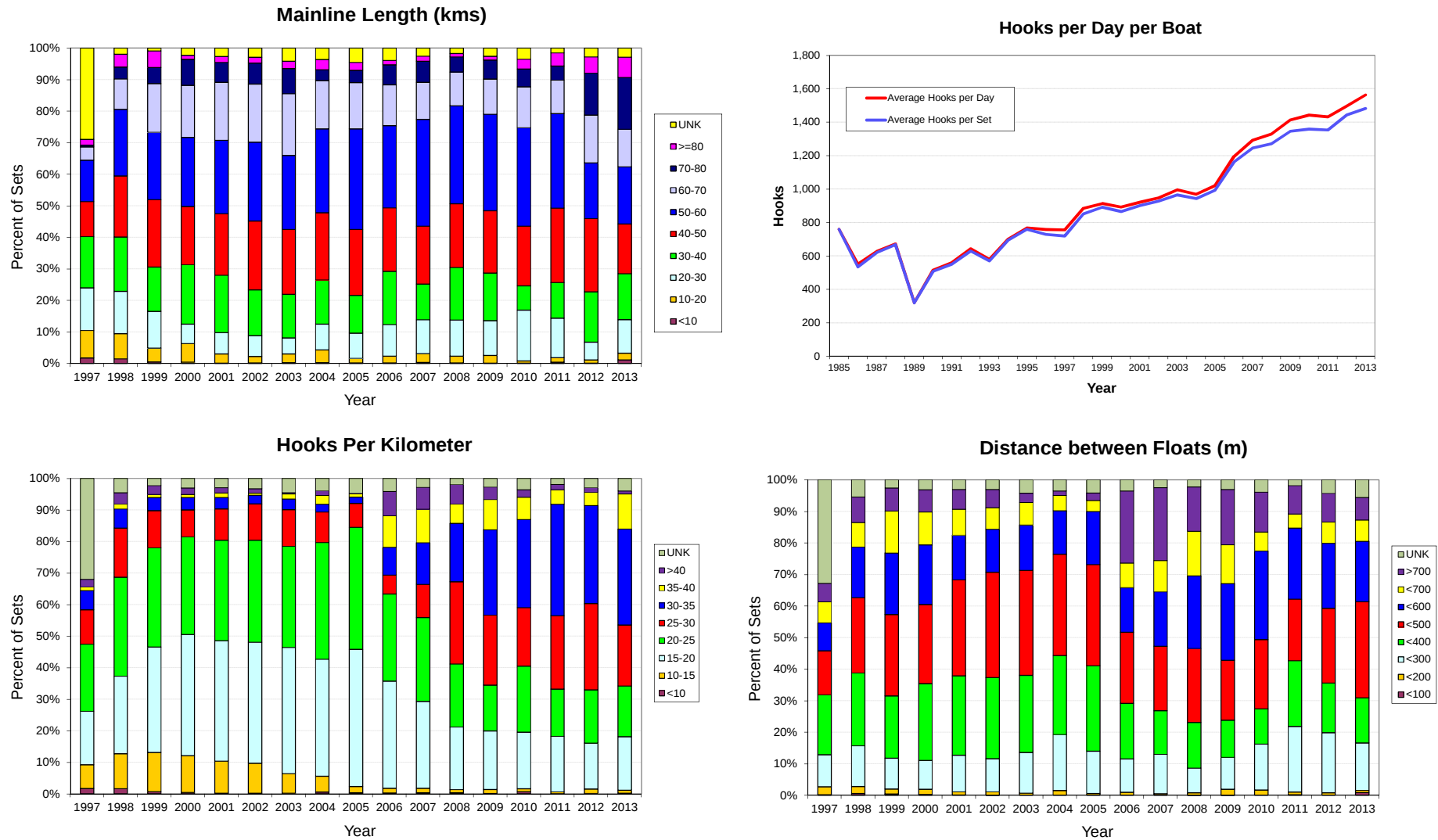


Figure 6.4: Average setting (and standard deviation) of individual gear-types during each year based on observer data collected from longline vessels operating in the ETBF (solid blue line). The number of observer voyages from which data was available is also shown (green columns).



Figure 6.5: Time-series of branch-line lengths deployed on five different vessels operating in the ETBF. Each dot represents an individual fishing trips and the length recorded by an observer for that trip.

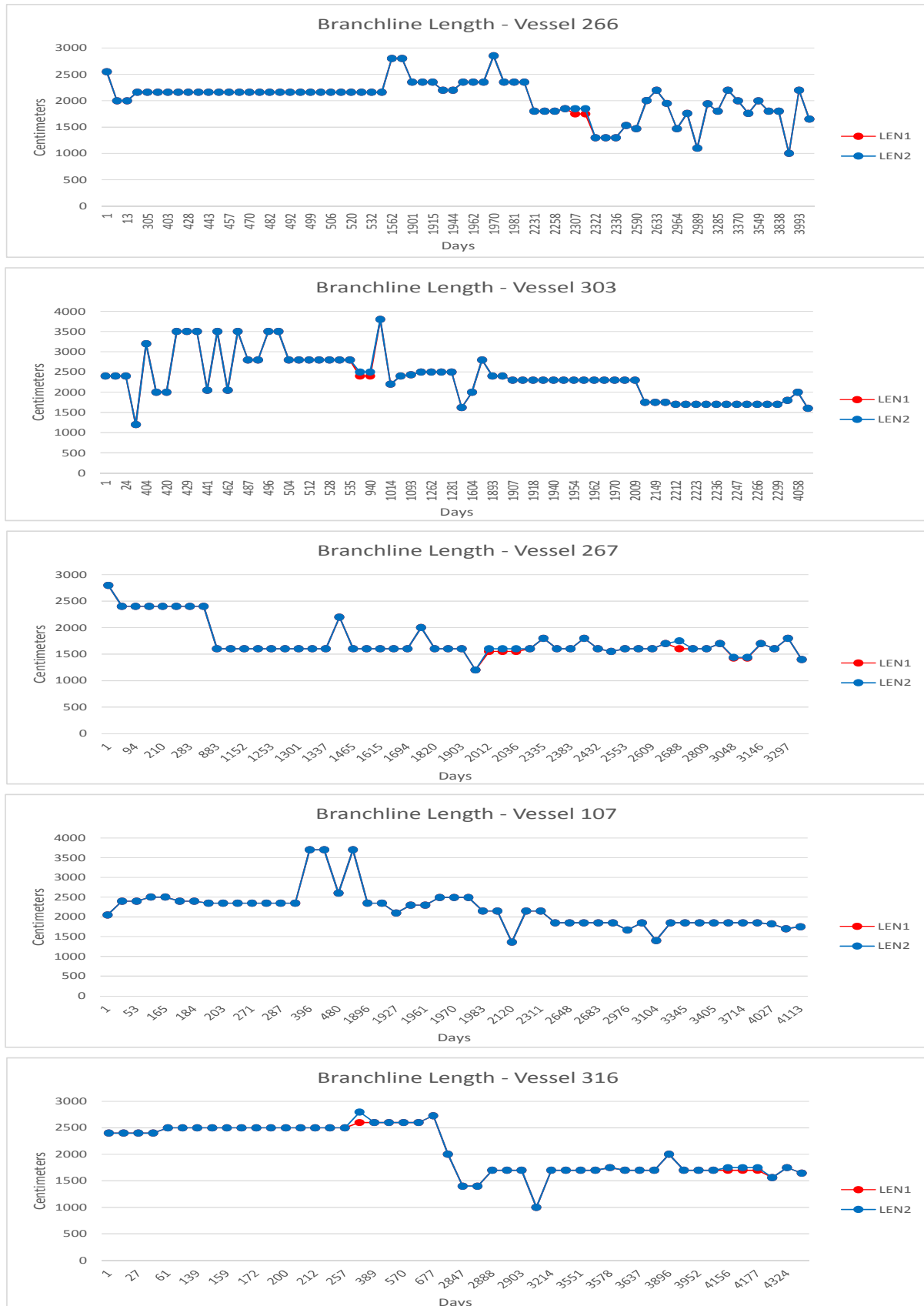
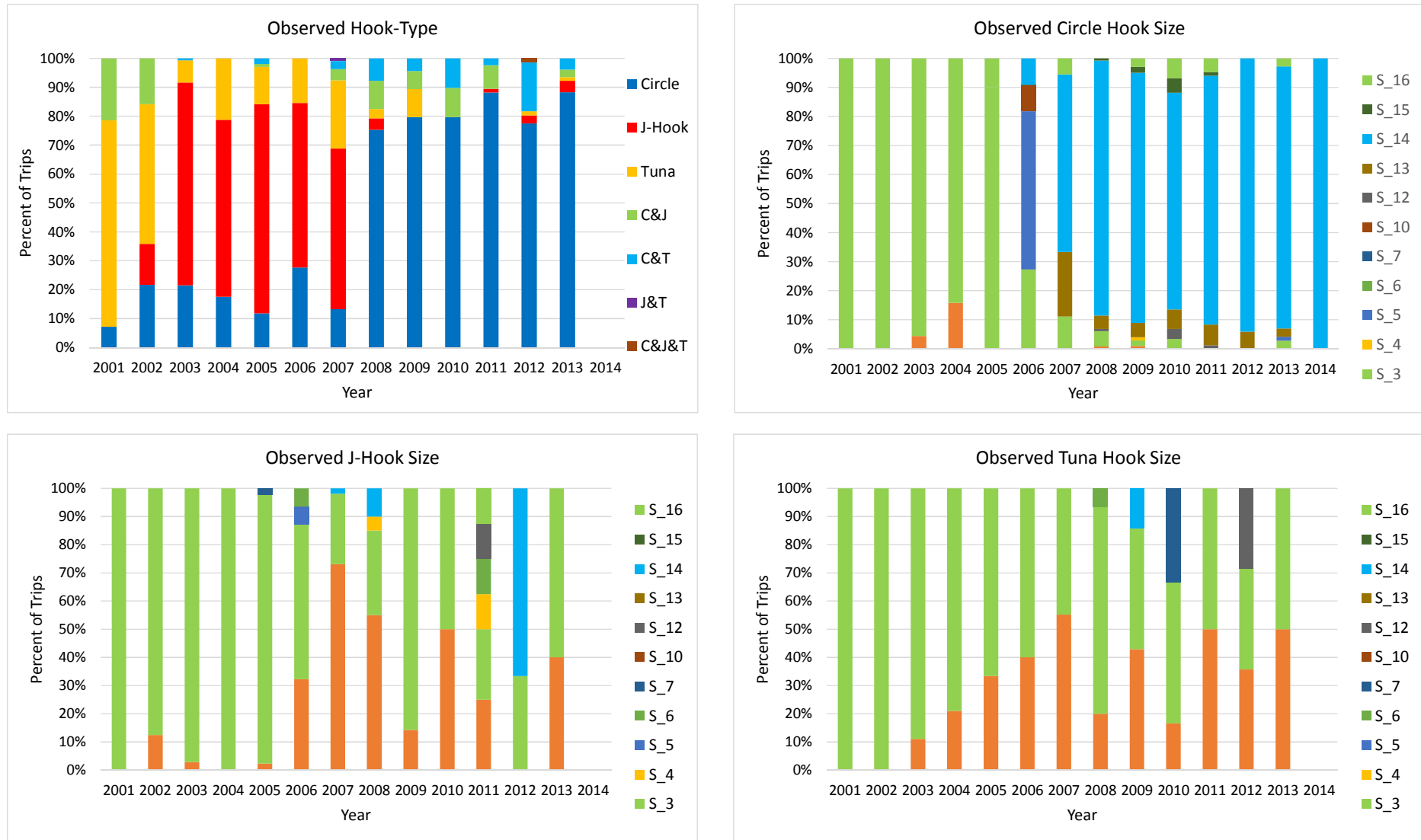


Figure 6.6: Annual distribution of hook-types deployed on observed longline vessel operating in the ETBF.



The average (and standard deviation) characteristic for several gear-settings only recorded by observers during each year is shown in Figure 6.4. The number of observer voyages from which data was available is also shown. There appears to have been an increase in the mean length of the buoy-line deployed from around 10m to 15m during the 2000s, while the mean length of the branch-lines has decreased. The weights placed on the branch-lines (to assist in sinking the hooks to the fishing depth) have also increased over time. It is also possible to investigate the use of different gear settings at the individual vessel level, and an example is shown in Figure 6.5 where the time-series of branch-line lengths deployed on five different vessels are shown. Finally, the annual distribution of hook-types deployed on observed longline vessels operating in the ETBF is shown in Figure 6.6 indicating a substantial shift to circle hooks after 2007.

The above results indicate that there can be considerable variability in the manner that longline gears are deployed in the ETBF, both between years and within years on individual vessels.

6.2.2. Gear Usage during Trips

In the absence of knowledge about the fish species being targeted in each fishing operation, one can attempt to use information recorded in the logbook on the manner in which the fishing gear is deployed, and the species composition of the catch, to infer what the targeting intention of the skipper may have been. Set level data contains variability in species composition due to the randomness of chance encounters between fishing gear and schools of fish. This variability can lead to possible misallocation of sets using different fishing strategies. Aggregating the data tends to reduce the variability, and therefore reduce misallocation of sets. The next level of aggregation above the individual set is the vessel-trip. We therefore investigated the variability in fishing strategies (i.e. gear setting practices) per fishing trip.

For ordinal variables (i.e. those which can be naturally ordered) such as '*Number of hooks deployed per Set*', which can range from 100 to 3000, it is unlikely that exactly the same number of hooks will be deployed each set during a trip. What we seek to identify are sets that deploy a similar number of hooks (e.g. 1000 and 1050 may be considered similar). For this purpose the mean, M , and range, R , of hooks deployed across all sets during a trip were calculated and all sets within a trip were categorized as being similar if the range was within 10% of the mean (i.e. $R \leq 0.1M$). This approach was used for the following ordinal data-types:

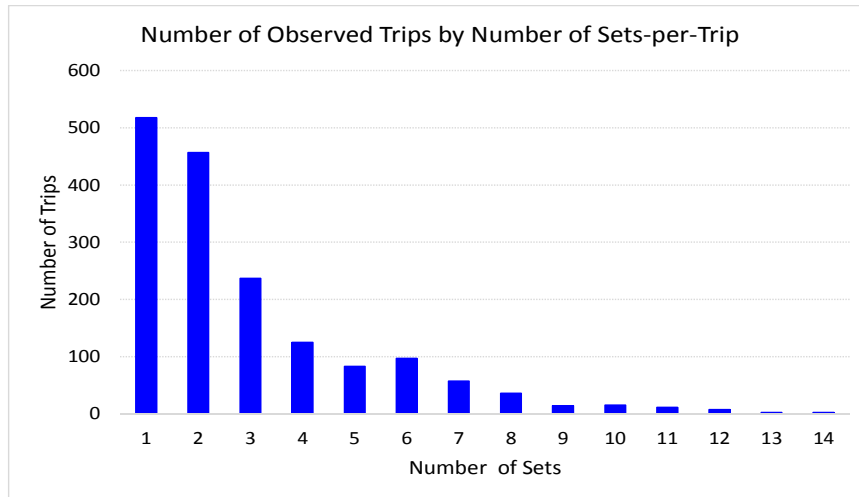
- 1) Hooks-deployed
- 2) Number of Hooks-per-float
- 3) Start-time of set
- 4) Percentage of hooks with light-sticks
- 5) Speed of line-shooter
- 6) Timer speed (timing between attachment of branch-lines)
- 7) Distance between branch-lines

For categorical variables such as 'Bait-species', 'Bait-life-status' and 'Primary target species', sets were identified as being similar when the same category was used for a given gear type on all sets during a trip. A similar approach can also be used for each of the ordinal variables listed above by first binning these variables into categories. An example is categorizing Start-Time into six 4-hourly bins (e.g. 0-4am, 4-8am,..., 8pm-12pm).

The number of observed trips extracted from the database was 1661 and the number of sets deployed per observed trip is shown in the Figure 6.7 below. The results of the above analyses, shown in Figures 6.8 and 6.9, indicate variable degrees of consistency in the deployment of each gear type during a trip. For example, similar hooks-per-float setting are used for 60-80%

of all sets during a trip, while generally similar start-times are used for less than 20% of sets. Note that the analysis was only undertaken for those trips where the data on the given gear type was available for all sets.

Figure 6.7: Histogram of the number of sets per observed trip.



6.2.3. Gear Usage by Primary Target Species

The observer data is particularly useful for investigating targeting practices as the Primary and Secondary target species are recorded for each set. In this Section we investigate the types of gear settings associated with each recorded primary target species.

The previous Section indicated that for many trips the primary target species is not always the same for all sets deployed during that trip (c.f. Figure 6.9). There may be a range of reasons for this outcome. This indicates that any grouping of fishing strategies into target clusters may need to be undertaken at the level of the individual fishing operations. As such, in order to investigate whether there are particular gear settings associated with particular target species, the distribution of settings for each gear-type at the set level against the identified primary target species was calculated. The results are shown in Figures 6.10 and 6.11.

The results indicate that there is a broad range of settings for each gear-type associated with each recorded primary target species. Furthermore, several characteristics of how the gears are deployed (e.g. buoy-line length and line-setting speed) do not display any appreciable variation between the different target species. However, on closer inspection one can discern particular groups of gear settings which have a higher usage when targeting particular species. For example, Figures 6.10 indicate that more hooks and hooks-per-float are deployed when targeting Albacore Tuna in comparison to the other species. Albacore targeting also generally uses fewer light-sticks, a shorter distance between branch-lines and longer buoy-lines, early morning sets, circle hooks and dead pilchard bait. On the other hand, when targeting Swordfish there is a preference to deploy the gear during late afternoon, use a high percentage of light-sticks, a greater distance between branch-lines and dead squid bait. When targeting Striped Marlin there is a preference for shallow sets (i.e. few hooks-per-float), minimum use of light-sticks, live mackerel bait and the use of a line-shooter.

Figure 6.8: Percentage of trips where the setting for each gear type was similar for all FOPS during a trip. The solid line displays the total number of trips for which data on the given gear type was available for all sets while the dashed line displays the number of trips for which the gear-settings were similar for all sets within a trip. Trips are stratified by the number of sets deployed per trip.

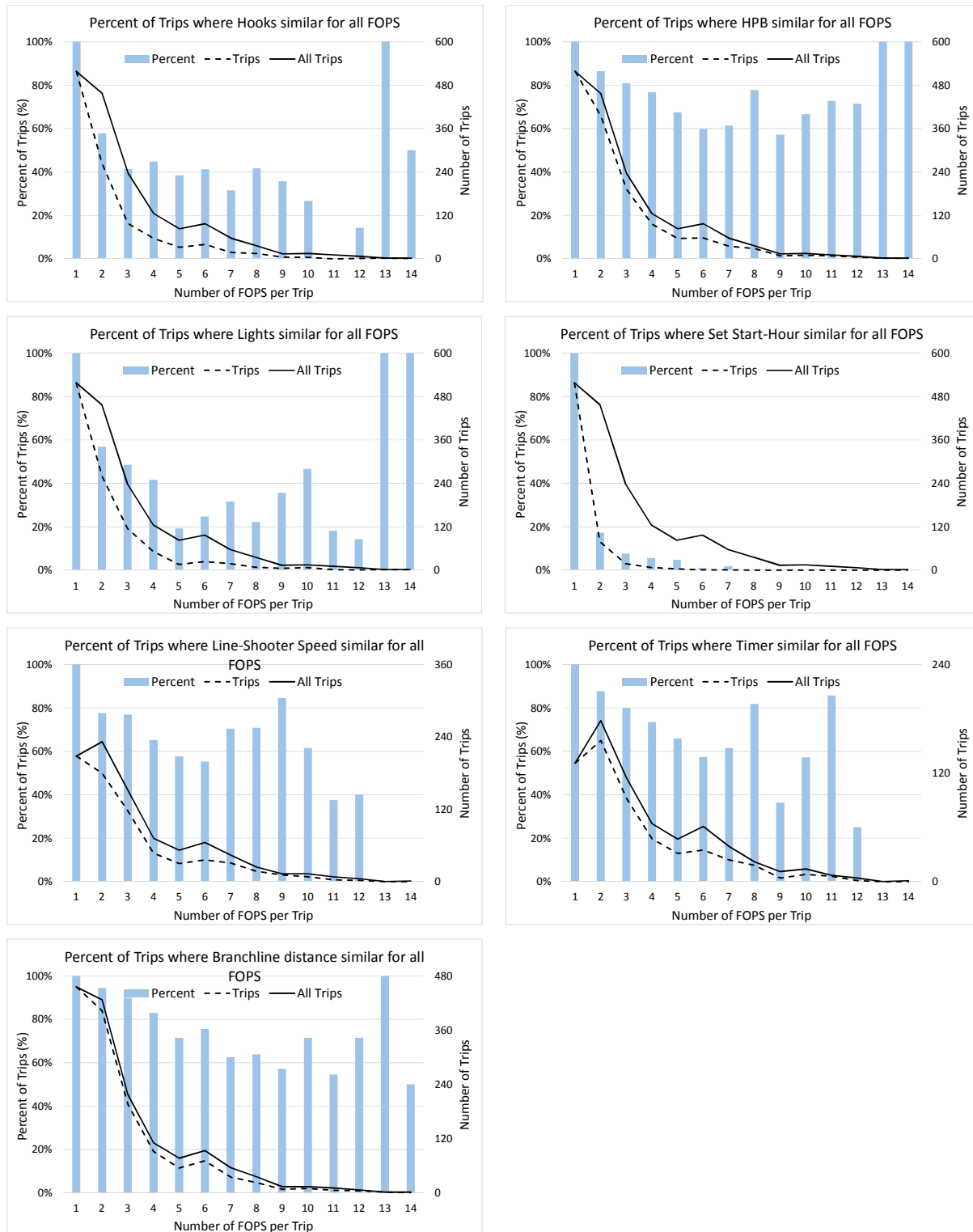
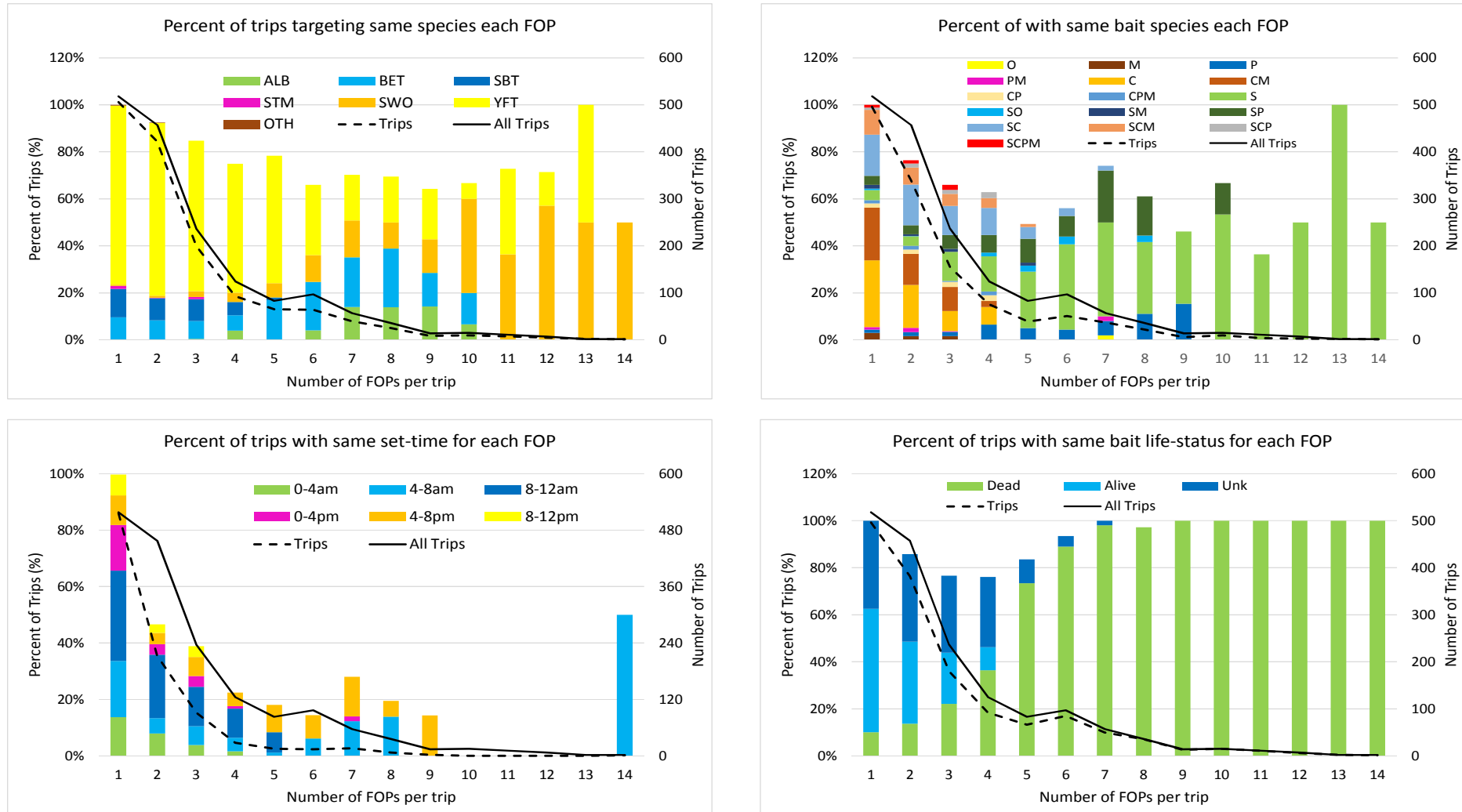


Figure 6.9: Percentage of trips where the setting for each gear type was similar (i.e. the same category level was used) for all FOPs during a trip. The solid line displays the total number of trips for which data on the given gear type was available for all sets while the dashed line displays the number of trips for which the gear-settings were similar for all sets within a trip. Trips are stratified by the number of sets deployed per trip. For those trips deploying the same gear type for all sets the percentage of trips using each gear-type category is also displayed.



The monthly time-series of the percent of observed fishing operations by primary target species is displayed in Figure 6.12a. The targeting of Yellowfin Tuna clearly dominates the time-series, though there are periods when the other tuna species (Bigeye and Albacore) and Broadbill Swordfish are the dominant target species. Note, however, that the observer data is not considered to be a random survey across the entire fleet as Sections of the fleet are targeted for increased observer coverage at certain times, e.g. vessels catching Southern Bluefin Tuna during winter. Aggregated over all years, the monthly distribution of targeting practices shown in Figure 6.12b indicates substantial change in targeting practices across each month. The targeting of Southern Bluefin Tuna over winter, together with the increased targeting of Bigeye Tuna in late autumn, of Yellowfin Tuna in spring and Swordfish over summer, are clearly seen and mimic the changes in relative availability of these species throughout the year (c.f. Figure 6.1).

The catch composition of all observed sets by month is shown in Figure 6.13a and again displays the substantial changes throughout the year, no doubt based on the changing relative availability of individual species and related changes in targeting practices. Note, that targeting practices, whilst influenced by relative availability of given species, will also be influenced by other preferences such as characteristics of the vessel (e.g. how far off-shore it can venture), region location of the fishing operations and market conditions (e.g. prices). Of the 29 species shown in the catch composition (comprising the most dominant species caught), the majority of the catch (>90%) each month is comprised of the following sub-set of twelve species (YFT=Yellowfin Tuna, BET=Bigeye Tuna, ALB=Albacore Tuna, SWO=Broadbill Swordfish, STM=Striped Marlin, SBT=Southern Bluefin Tuna, LEC=Escolar/Oilfish, ALX=Lancetfish, DOL=Dolphin fish, POA=Promfrets/Ray's Bream, and BSH=Blue Shark).

The catch composition of all observed sets stratified by primary target species is shown in Figure 6.13b. It is reassuring to note that for each of the main target species, the proportional catch of that species is highest when that species is recorded as the primary targeted species. This demonstrates that the vessel (skipper) has some ability to target and catch a desired species. Nevertheless, it is also clear that the other species are also usually caught, indicating that it is not possible to just target and catch a single species. Indeed, the target species sometimes is not the dominant catch. For example, on average more Yellowfin Tuna are caught than Striped Marlin when targeting the latter.

Finally, the cumulative percent of sets catching less than or equal to the indicated number of fish stratified by primary target species is shown in Figure 6.14. For example, when targeting Albacore Tuna around 30% of sets catch 25 Albacore or less (with 70% of sets catching more than 25 Albacore), but when targeting Bigeye Tuna around 90% of sets catch 25 Albacore or less (with 10% of sets catching more than 25 Albacore), and when targeting Striped Marlin 100% of sets catch less than 25 Albacore. For the species shown, the highest number of fish are caught when that species is the target species, again indicating that fishers have the ability to target particular species in the fishery.

Figure 6.10: Distribution of various gear-setting stratified by primary target species recorded by observer– ordinal variables.

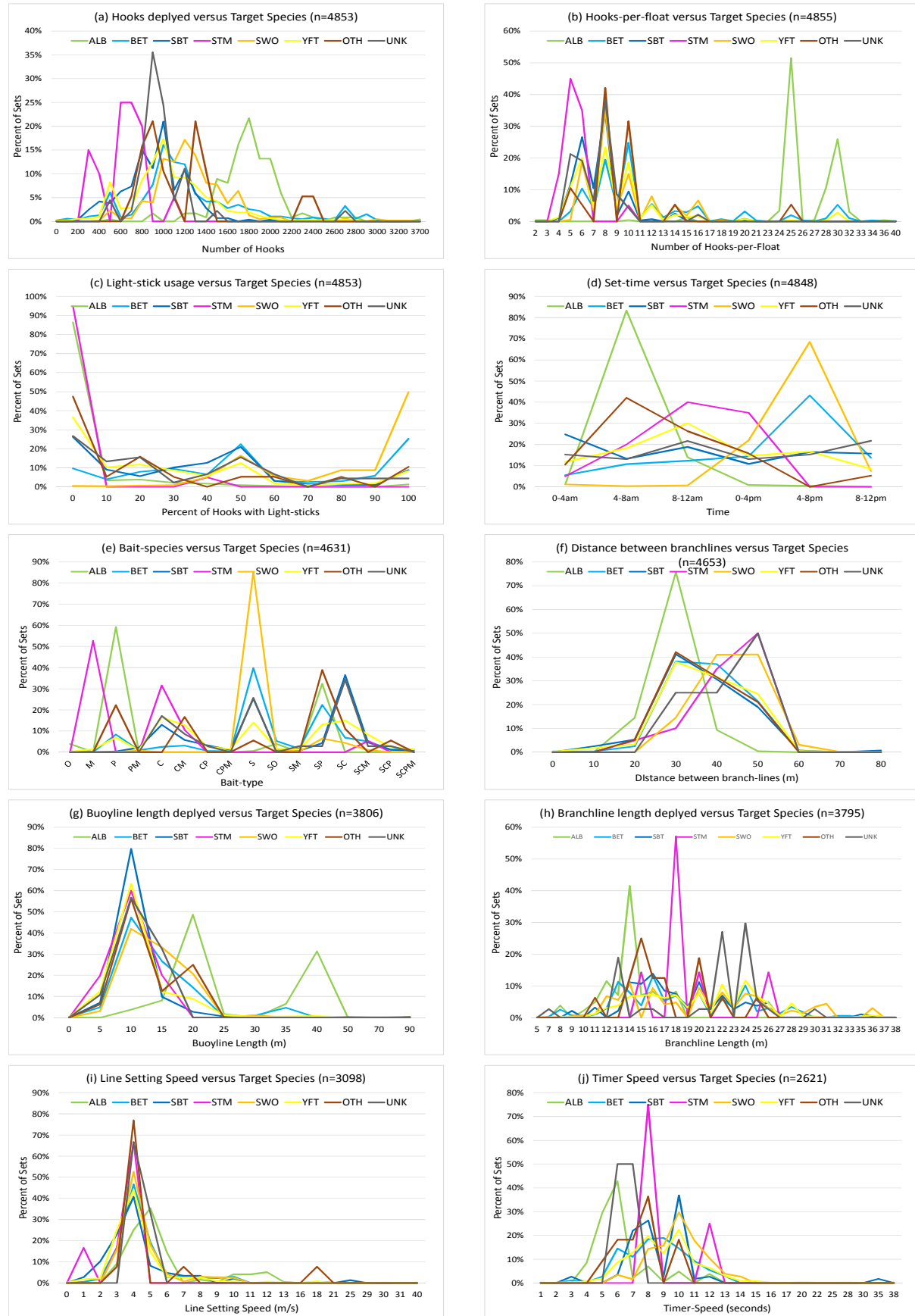


Figure 6.11: Distribution of various gear-setting stratified by primary target species recorded by observer – categorical variables.

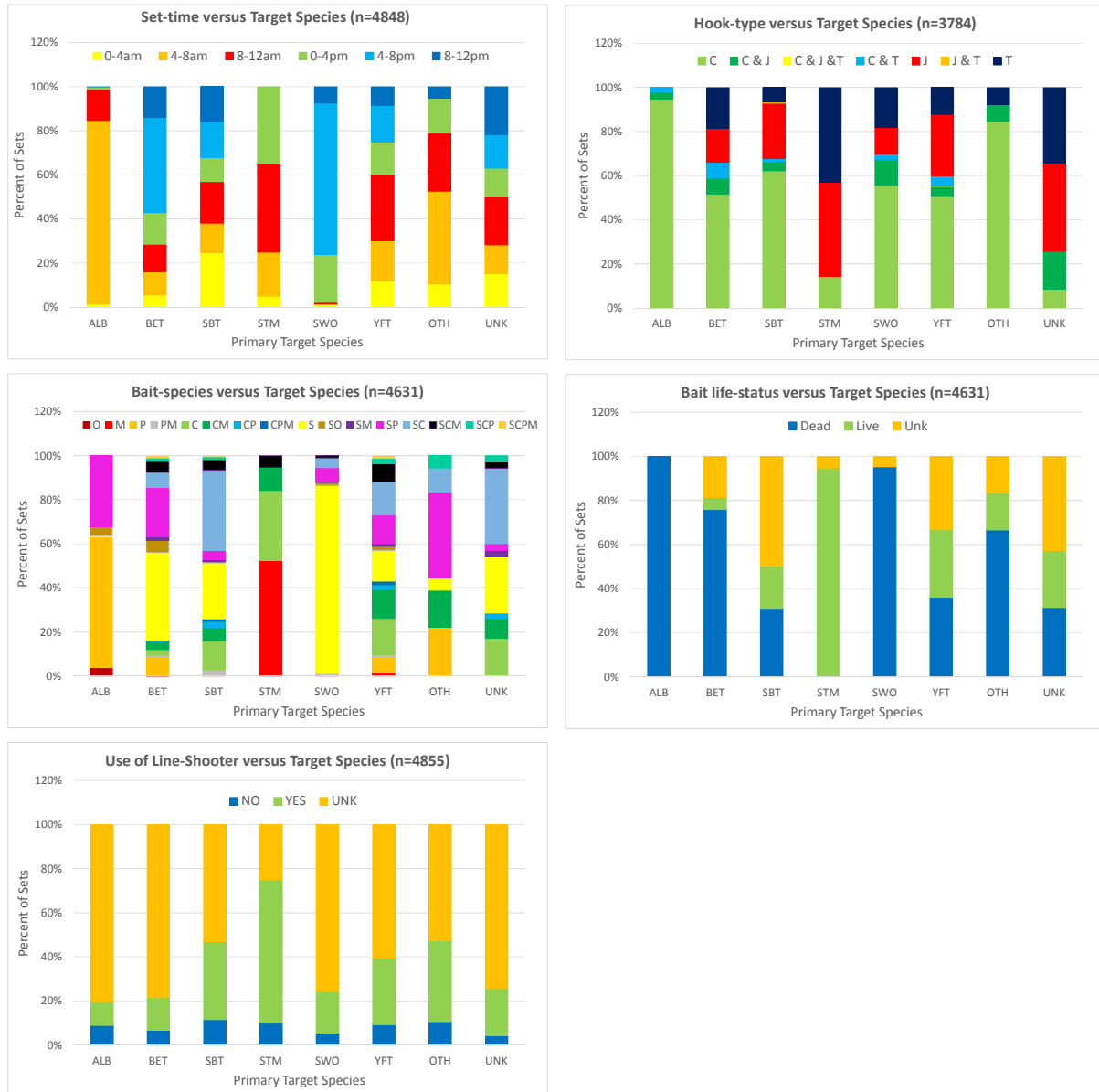


Figure 6.12: Percent of observed fishing operations by primary target species (a) by month and (b) aggregated by month across years.

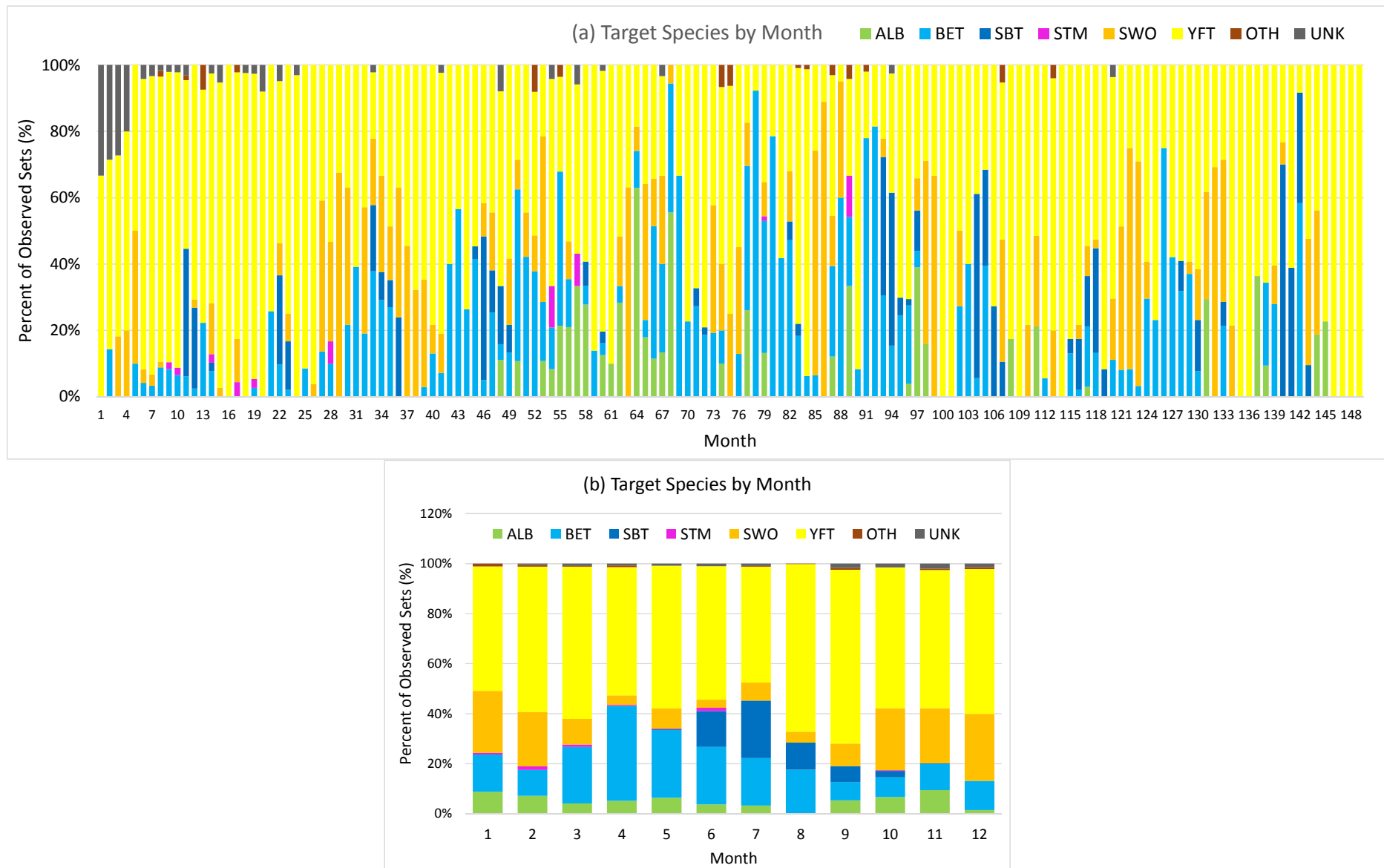


Figure 6.13: (Catch composition of all observed sets by (a) month and (b) primary target species. (YFT=Yellowfin Tuna, BET=Bigeye Tuna, ALB=Albacore Tuna, SWO=Broadbill Swordfish, STM=Striped Marlin, SBT=Southern Bluefin Tuna, LEC=Escolar/Oilfish, ALX=Long-nosed Lancetfish, DOL=Dolphinfish, POA=Ray's Bream/Promfrets, BSH=Blue Shark, SKJ=Skipjack Tuna, WAH=Wahoo, MAK=Mako Shark, GES=Snake Mackerel, ALO= short-nosed lancetfish, MOP=Sunfish, SBS=Short-bill Spearfish, BAM=Black Marlin, BUM=Blue Marlin, OPA=Opah, RAY=Stingrays, TIG=Tiger Shark, BRO=Bronze Whaler, DSK=Dusky Shark, OCS=Oceanic Whitetip Shark, CSH=Crocodile Shark, SKS=Silky Shark and SPN=Hammerhead Sharks).

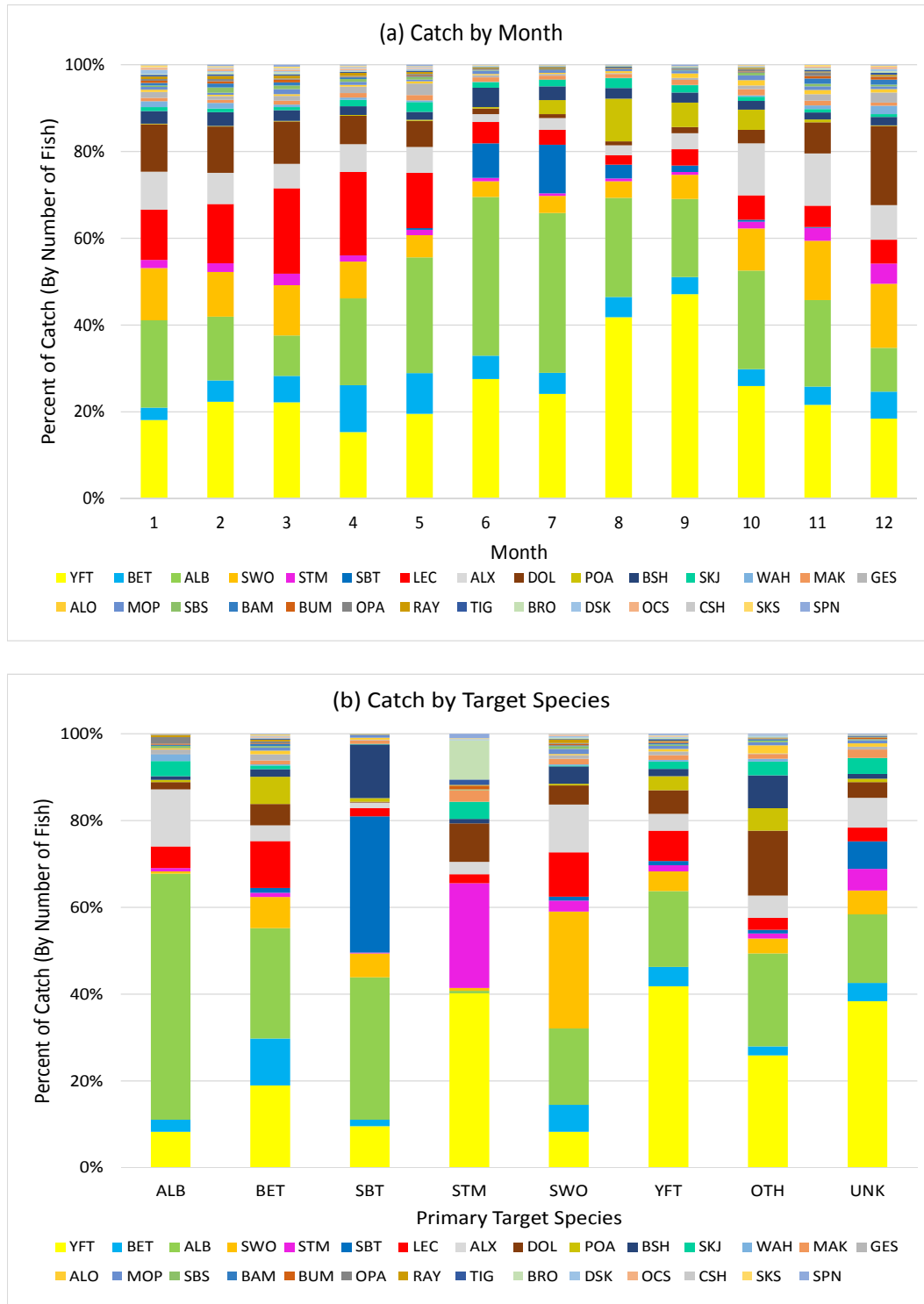
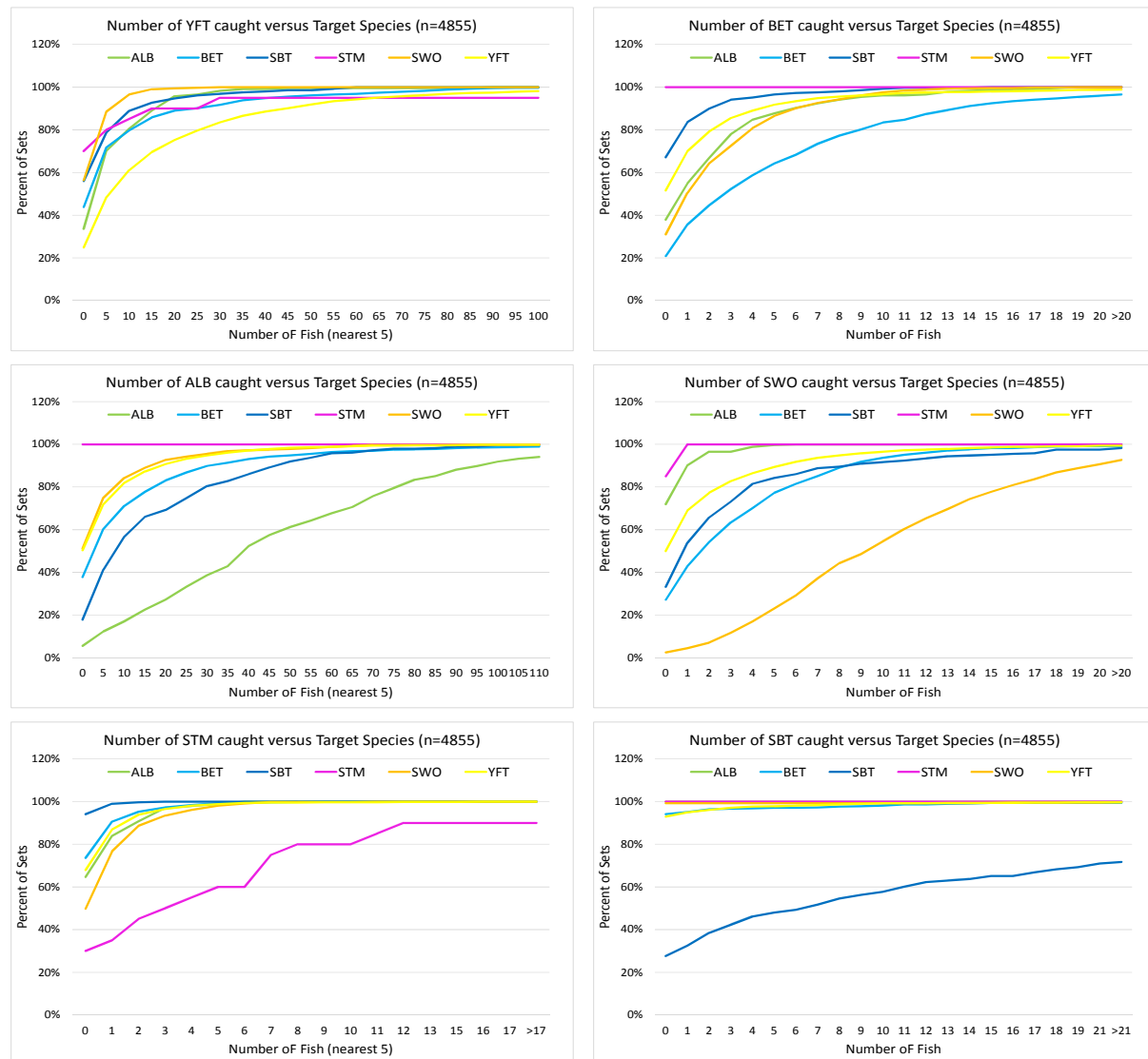


Figure 6.14: Cumulative percent of sets catching less than or equal to the indicated number of fish stratified by primary target species.



6.3 Species Coexistence

Fish are not randomly distributed across the fishing grounds but tend to have distinct species-specific habitat preferences, which cause catch rates to vary according to the habitat that is targeted. These habitats are also likely to change their spatial size and location in response to prevailing currents and seasonal changes in oceanographic conditions. As a consequence, there is likely to be a strong seasonal component in both the distributions of fish species and related fishing and targeting strategies across any fishery.

6.3.1 Correlations between spatial distributions

In order to investigate these changing spatial distributions for the ETBF, spatial plots (on a 1x1-degree basis) of aggregate nominal CPUE over the years 2000 to 2013 for each quarter were produced for 17 species caught in the ETBF and these are shown in Appendix B. The persistence of these spatial distributions for each species was then investigated by calculating the Pearson correlation, R , between quarters and the results are displayed in Figure 6.15.

Figure 6.15: Pearson correlation between quarters of the spatial distributions of CPUE for each species.

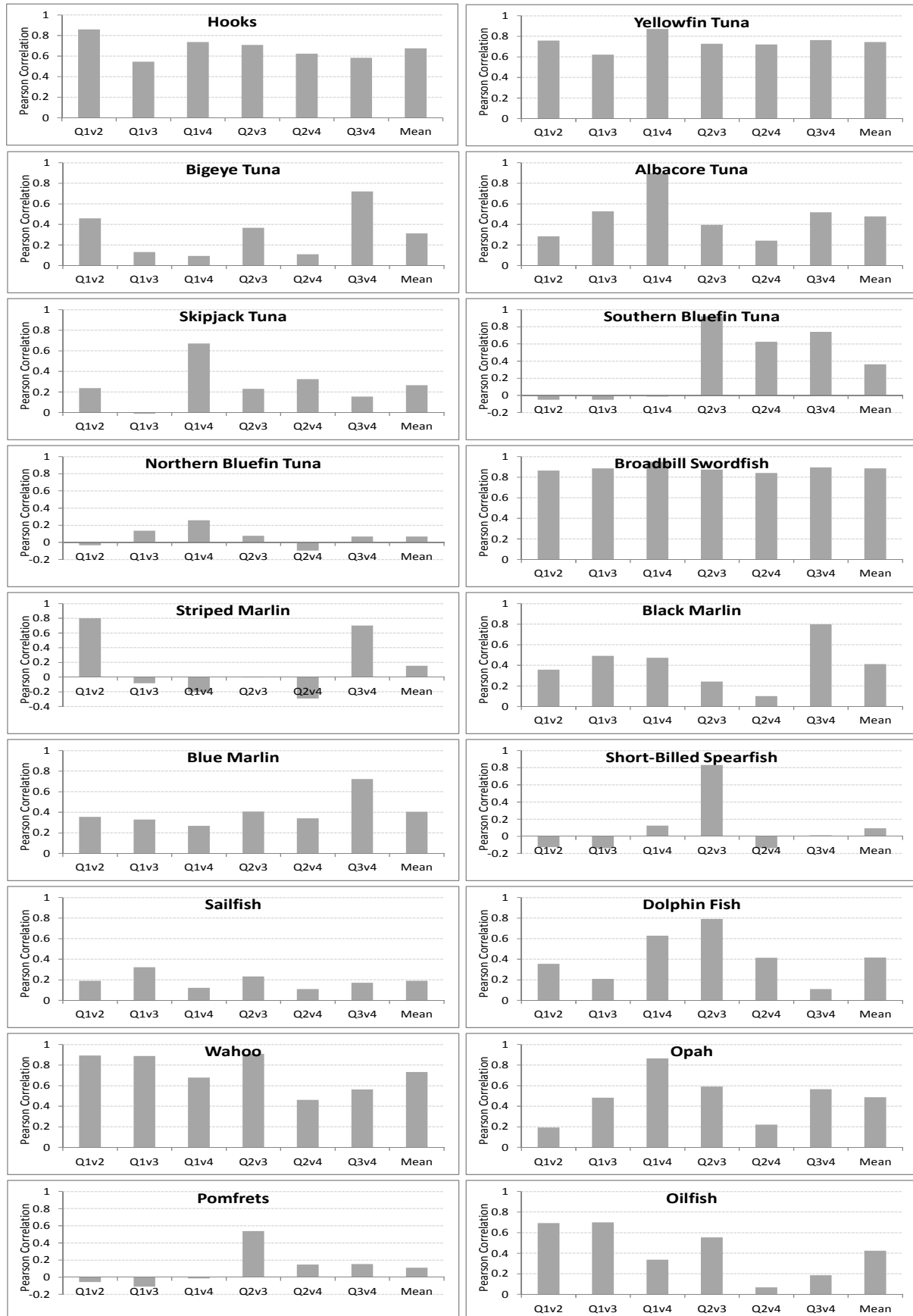


Figure 6.16: Pearson correlation, by quarter, between the spatial distributions of CPUE for a given specie and all other species.

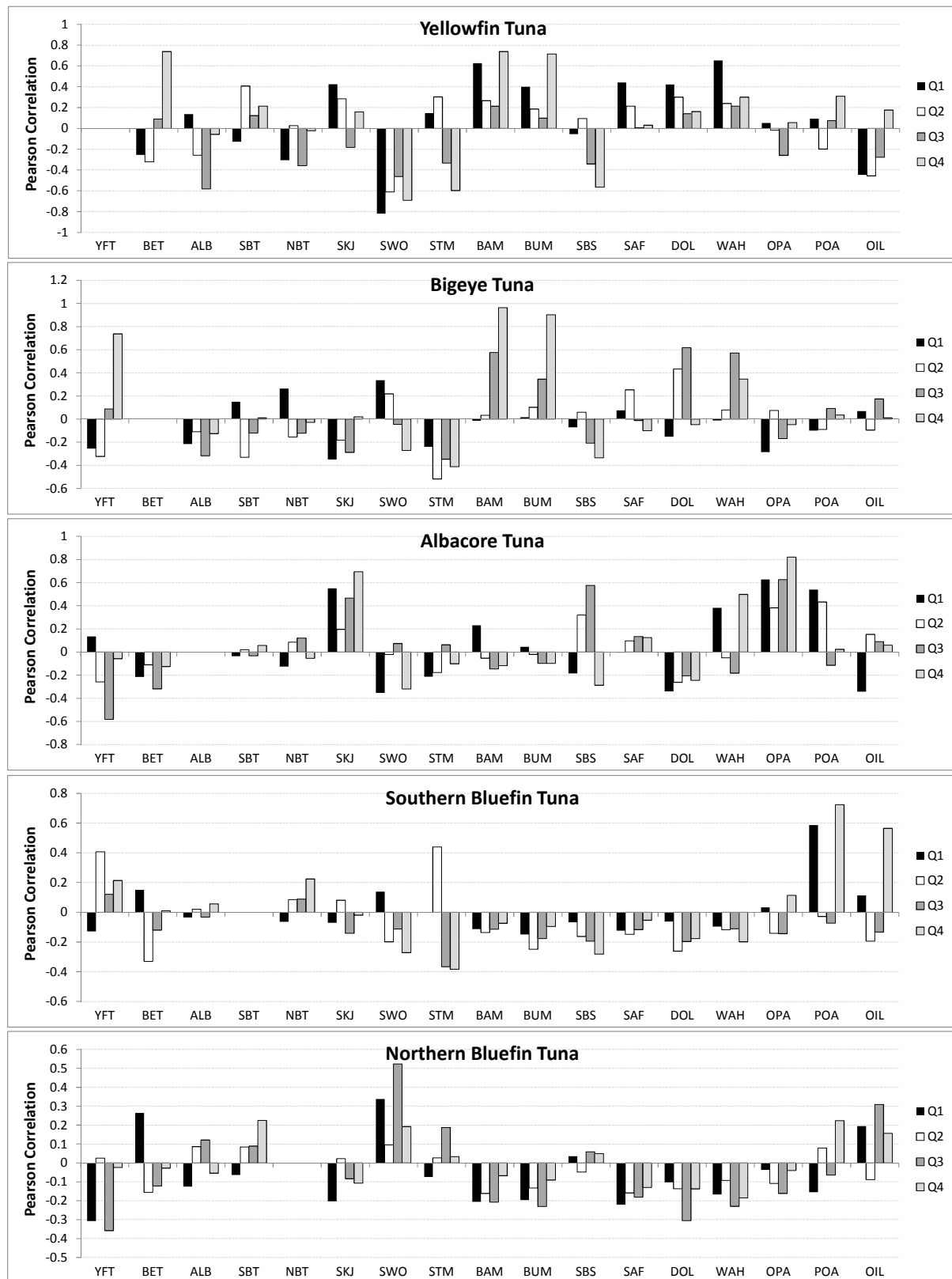


Figure 6.16: (cont'd) Pearson correlation, by quarter, between the spatial distributions of CPUE for a given species and all other species.

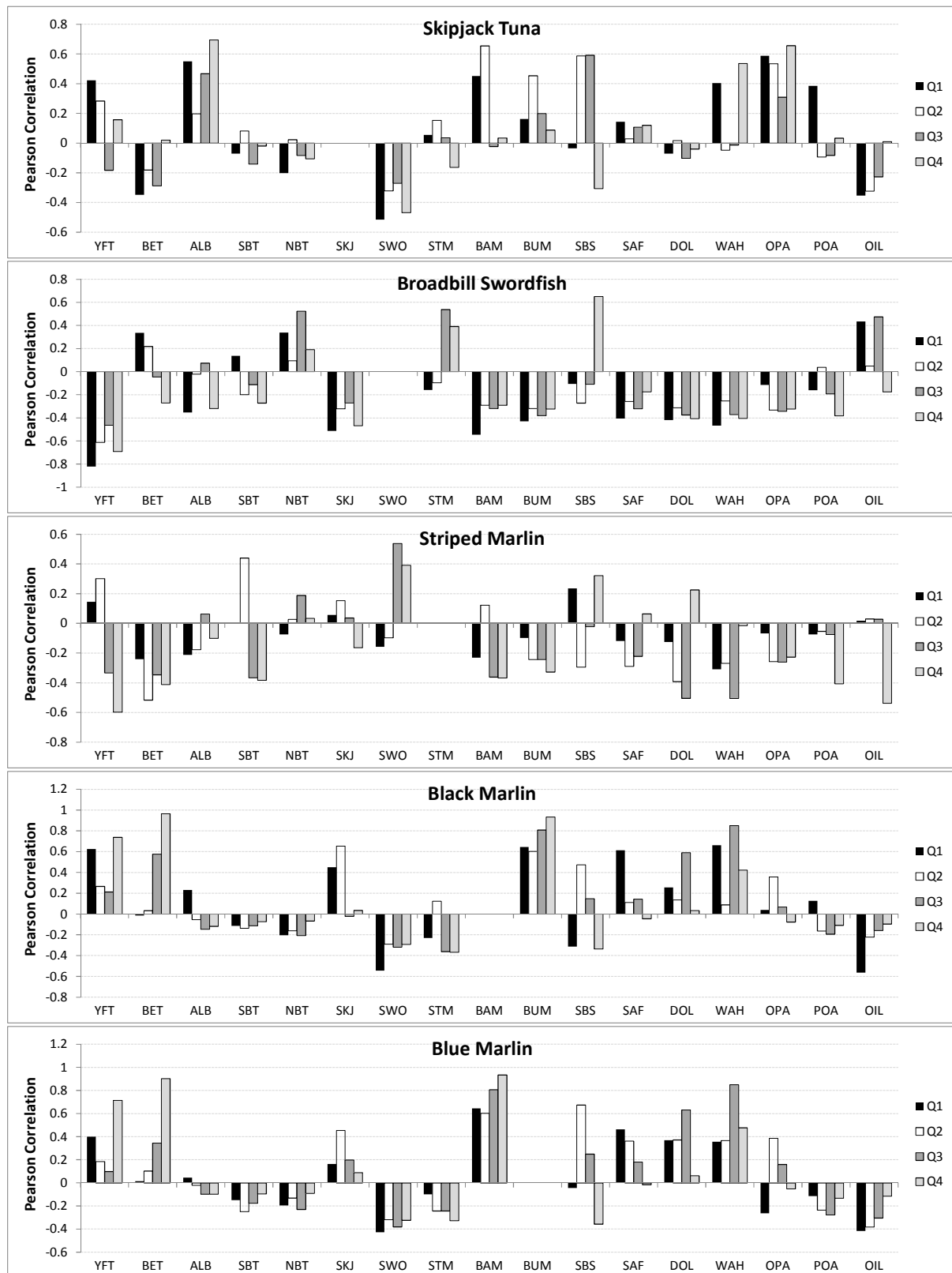


Figure 6.16: (cont'd) Pearson correlation, by quarter, between the spatial distributions of CPUE for a given species and all other species.

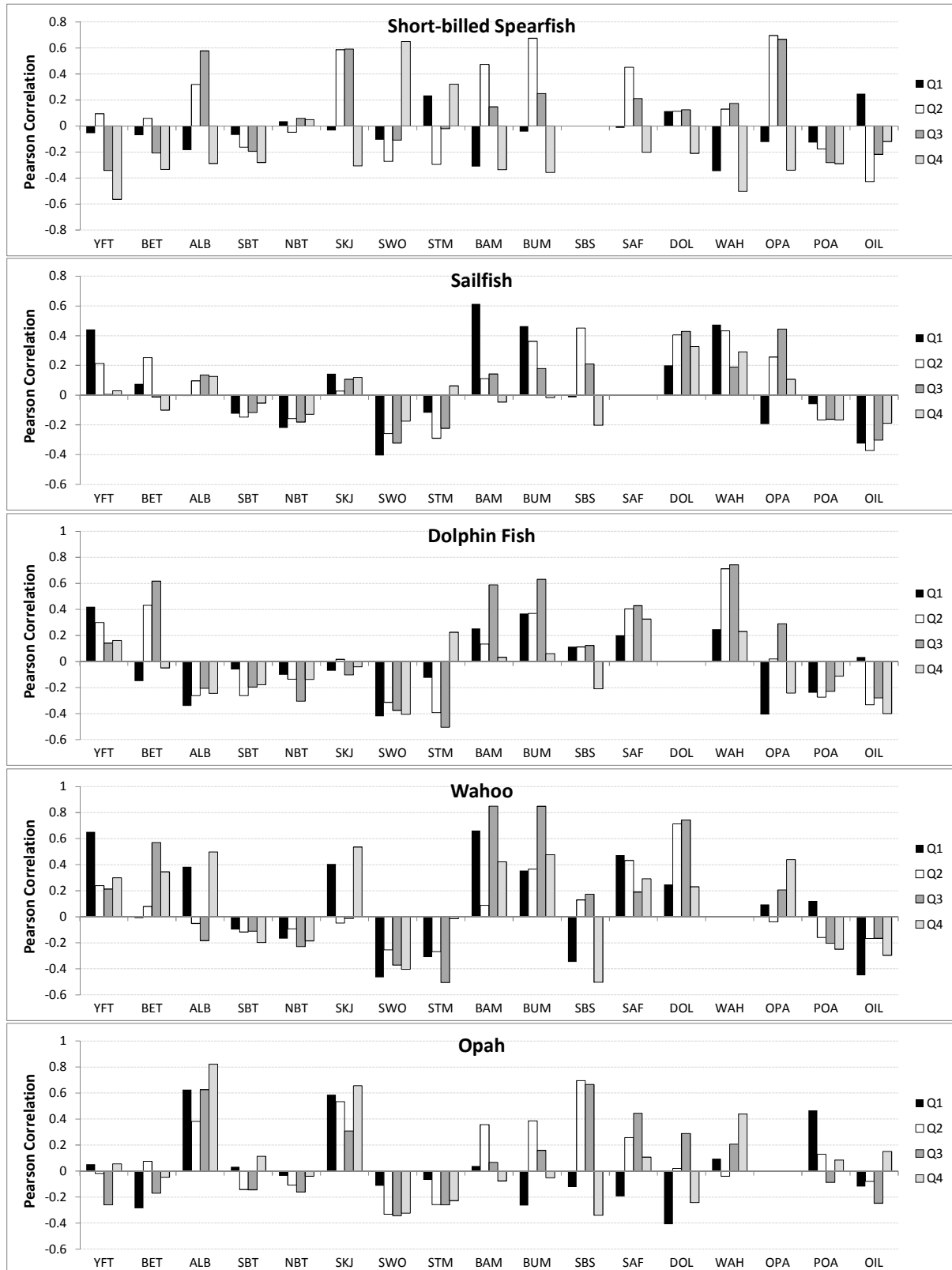
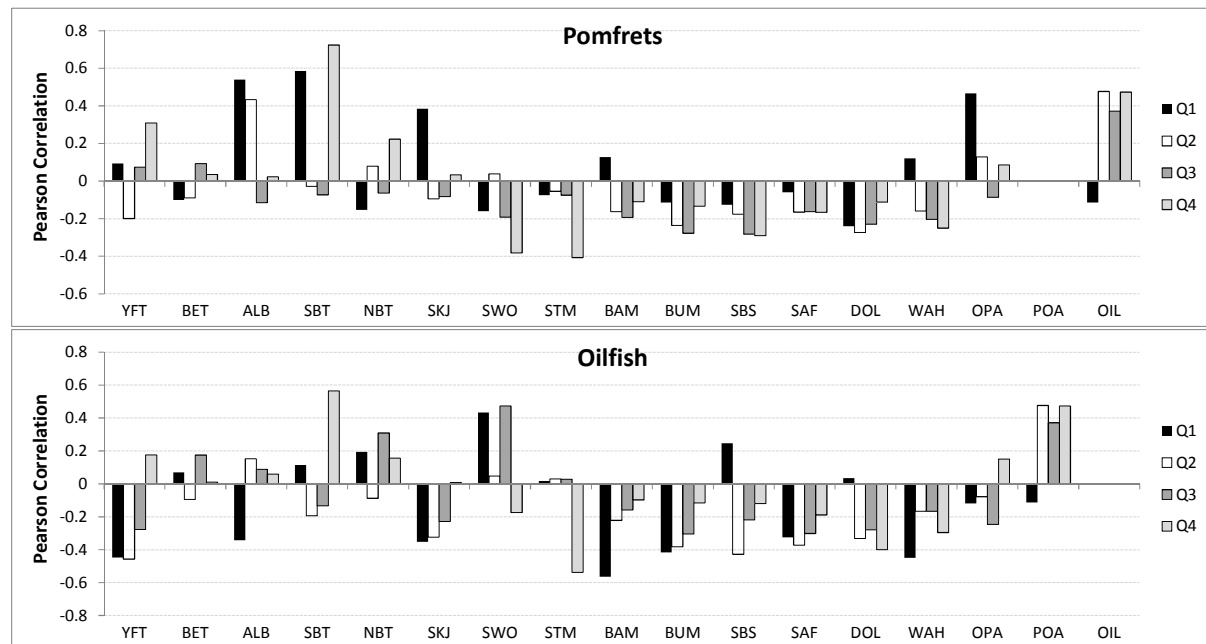


Fig 6.16: (cont'd) Pearson correlation, by quarter, between the spatial distributions of CPUE for a given species and all other species.



The spatial distribution of CPUE (assumed to be a proxy for biomass) was found to be most consistent for Broadbill Swordfish, where for the six quarterly comparisons R was between 0.84 and 0.95 and reasonable high for Yellowfin Tuna (R between 0.62 and 0.95). On the other hand, the results for Bigeye and Albacore Tunas, together with Striped Marlin, indicate variable (and often low) degrees of persistence between the different quarters.

In order to investigate where there are suites of species which have correlated spatial distributions within and across quarters of the year, the Pearson correlation between the spatial distributions of CPUE for a given species, and all other species, was determined between each year. These results (shown in Figure 6.16) can be used to identify companion species, i.e. species which occur in similar habitats and as such often are caught together. For example, results indicate that the distribution of Yellowfin Tuna CPUE is positively correlated with black and Blue Marlin, Dolphin Fish and Wahoo for all quarters (indicating the possible persistence of this species group), and quite strongly negatively correlated with Broadbill Swordfish. The degree of co-habitation identified will help inform the results of other multispecies analyses carried out on the data (e.g. cluster analyses). Furthermore, if the levels of co-habitation are found not to persist across the different seasons, it is likely that such analyses will need to be carried out on a seasonal basis.

Bigeye Tuna and Broadbill Swordfish are reported to be closely linked ecologically, and these two species have been shown to commonly co-occur in longline catches in the central and southern regions of the east coast fishery (Ward 1996). Similarly, Yellowfin Tuna and Black Marlin commonly co-occur in longline catches in the central and northern regions, and it has also been suggested that within the northern region there is a general inverse relationship between Yellowfin Tuna and Striped Marlin. These patterns in the longline catch rates are hypothesized to result from partitioning of the oceanic habitat, as is well known in other species groups. This partitioning has implications both for the efficient targeting of fishing operations and the interpretation of catch and effort data. For example, changes in fishing operations to target different species produces changes in the effectiveness of fishing effort units.

6.3.2 Cluster Analysis

To further explore the coexistence of species available to the ETBF longline fishery, we again use the cluster analysis previously described in Section 5.4 to categorise individual fishing sets using based on similarities in the catch composition. We use the logbook data within the ETBF which provides catch information for a total of 29 individual species (c.f. Appendix A). Furthermore, to investigate possible seasonal differences in cluster types, two sets of analyses were undertaken. The first analysis was limited to the months of January and February (called Summer), while the second analysis was limited to the two June and July (called Winter). Each analysis was also limited to the main central region of the ETBF (i.e. latitudes between 20-35°S). A summary of these constraints is provided below, together with the number of longline sets included in each analysis:

Analysis 1: Summer

Spatial area: $20 \leq \text{latitude} \leq 35$
Months: January and February only
Number of records = 12,345

Analysis 2: Winter

Spatial area: $20 \leq \text{latitude} \leq 35$
Months: June and July only
Number of records = 15,947

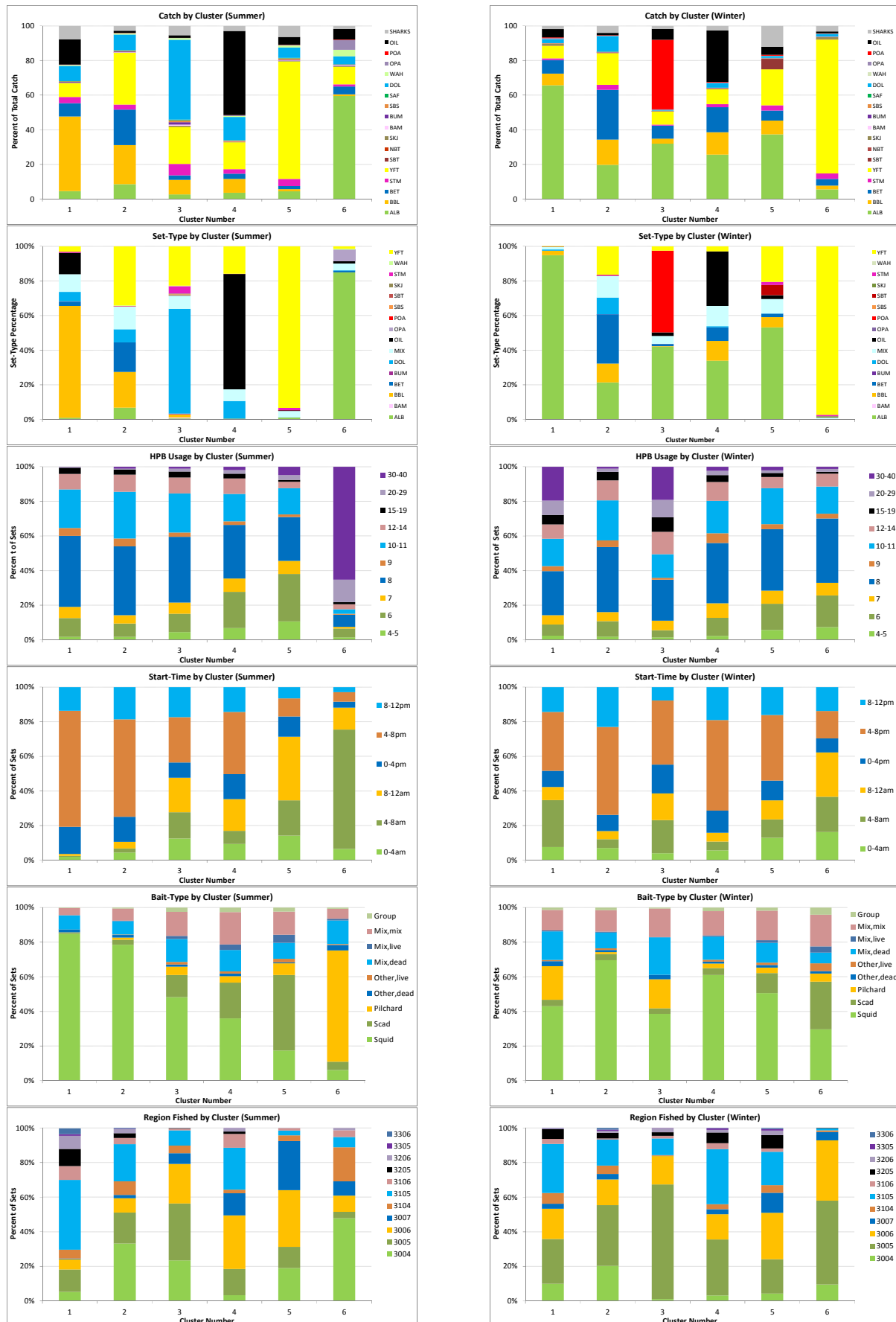
Finally, based on the number of key target species in the ETBF, and results of a number of preliminary trials, the number of clusters to be produced from each analysis was set at six. After production of the clusters, the mean catch compositions of the sets within each cluster were calculated and compared, while other operational characteristics were also compared between the clusters. Each longline set was also given a 'set-type' label corresponding to the dominant species caught (by number). For sets where more than one species was dominant, the set was classified as Mixed (MIX). The percentage of set-types within each cluster were calculated and compared.

Results

A comparison of the catch composition, set-types and operational characteristics comprising each of the six clusters corresponding to each analysis is shown in Figure 6.17. There are clear differences in the catch compositions between the clusters identified, and for most clusters a single species comprises a dominant proportion of the catch ($> 40\%$ for 8 of the 12 clusters). For example, Albacore and Yellowfin Tuna comprise 66% and 77% of the catch for clusters 1 and 6 respectively in the winter analysis, and 60% and 68% of the catch for clusters 6 and 5 respectively in the summer analysis. However, apart from the two respective 'Albacore Tuna' and 'Yellowfin Tuna' clusters which was seen to occur in each season, the composition of the other clusters are quite different. For example, the clusters for which Swordfish, Dolphinfin and Oilfish are dominant during the summer months, but are not seen in the winter months. Instead there is a cluster dominated by pomfrets and several generally mixed-species clusters (Albacore and Yellowfin, Oilfish and Albacore, and Bigeye, Albacore and Yellowfin). This dissimilarity again illustrates the seasonal differences in the co-occurrence of species.

Some clusters display distinct differences in the gear configurations. For example, for summer cluster 1 (the 'Swordfish' cluster) the sets are shallow (< 11 hooks-per-float), deployed mainly in the afternoon (4-8pm), and predominantly use squid bait. On the other hand, summer cluster 6 (the 'Albacore' cluster) is associated with a high use of deep setting (30-40 hooks-per basket), deployment in the morning (4-8am) and the use of pilchard baits, though a different and broader range of gear configurations are associated with the 'Albacore' winter cluster. There are also differences in the mix of gears between the summer and winter clusters comprised primarily of Yellowfin Tuna. For most clusters, nevertheless, there is a relatively wide use of gears configurations, indicating that 'targeting' may depend on more factors (e.g. oceanographic

Figure 6.17: Comparison of results of cluster analysis for summer and winter data

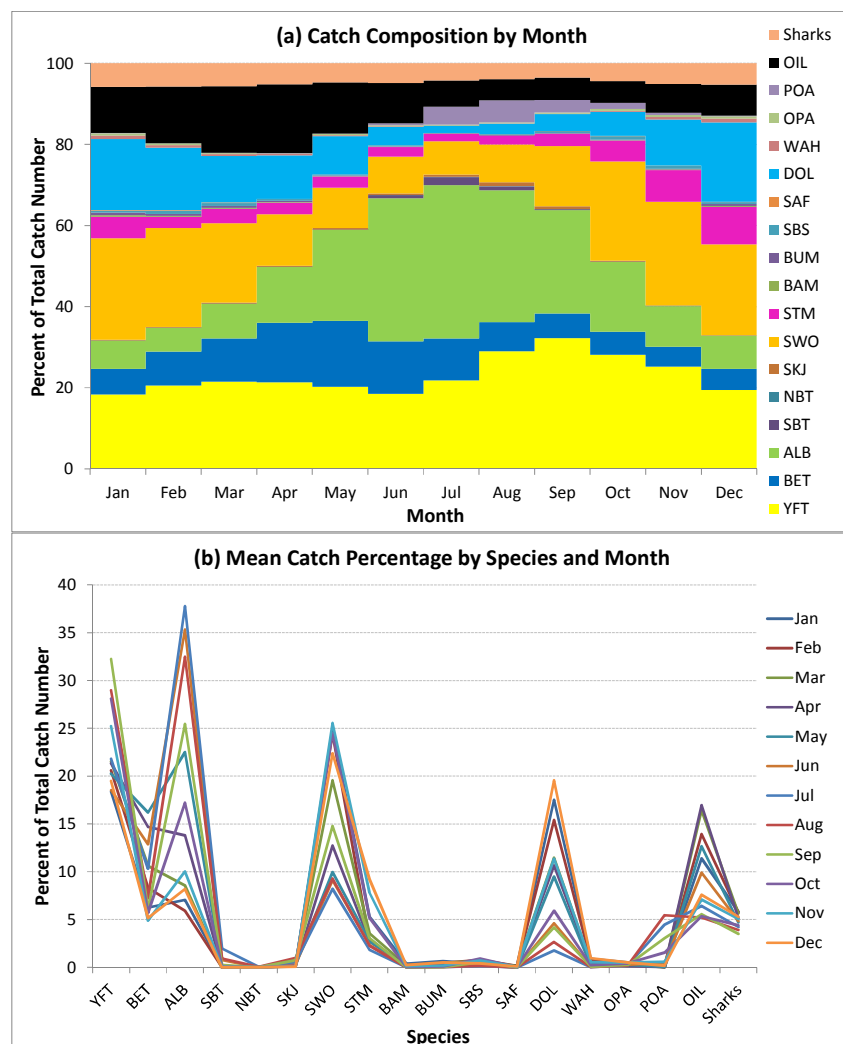


conditions such as water temperatures, eddy features) than just the configuration of the fishing gear. The influence of such factors in helping to delineate 'cluster' types will need to be investigated separately.

6.3.3 Further Considerations

While the logbook data provided by the ETBF logbook provides catch information for a total of 29 species (including 12 shark species), it is likely that some of the less frequently caught by-product and predominantly discarded species (c.f. Appendix A) have not been accurately recorded throughout the times-series of available data. Hence, there may be some merit in limiting the analysis of the data to the catch of the principal catch species only. Whilst limiting the analysis to these species may help to simplify the analyses, on the other hand the analysis may be compromised by including species where the catch statistics may not be as reliable. Of the 29 species included in the data, only ten species (YFT, ALB, SWO, OIL, BET, DOL, STM, BSH, POA, MAK) individually comprise more than one percent of the mean total catch across the 87,106 longline sets included in the above analyses. These ten species together comprise 96.5 percent of the mean total catch. Furthermore, apart from these ten species only one other species (SBT) comprises more than one percent of the mean total catch in any month (c.f. Figures 6.18a&b).

Figure. 6.18: (a) Composition of the mean monthly catch recorded in ETBF logbooks, and (b) Percentage of each species recorded in the ETBF logbook comprising the mean monthly catch.

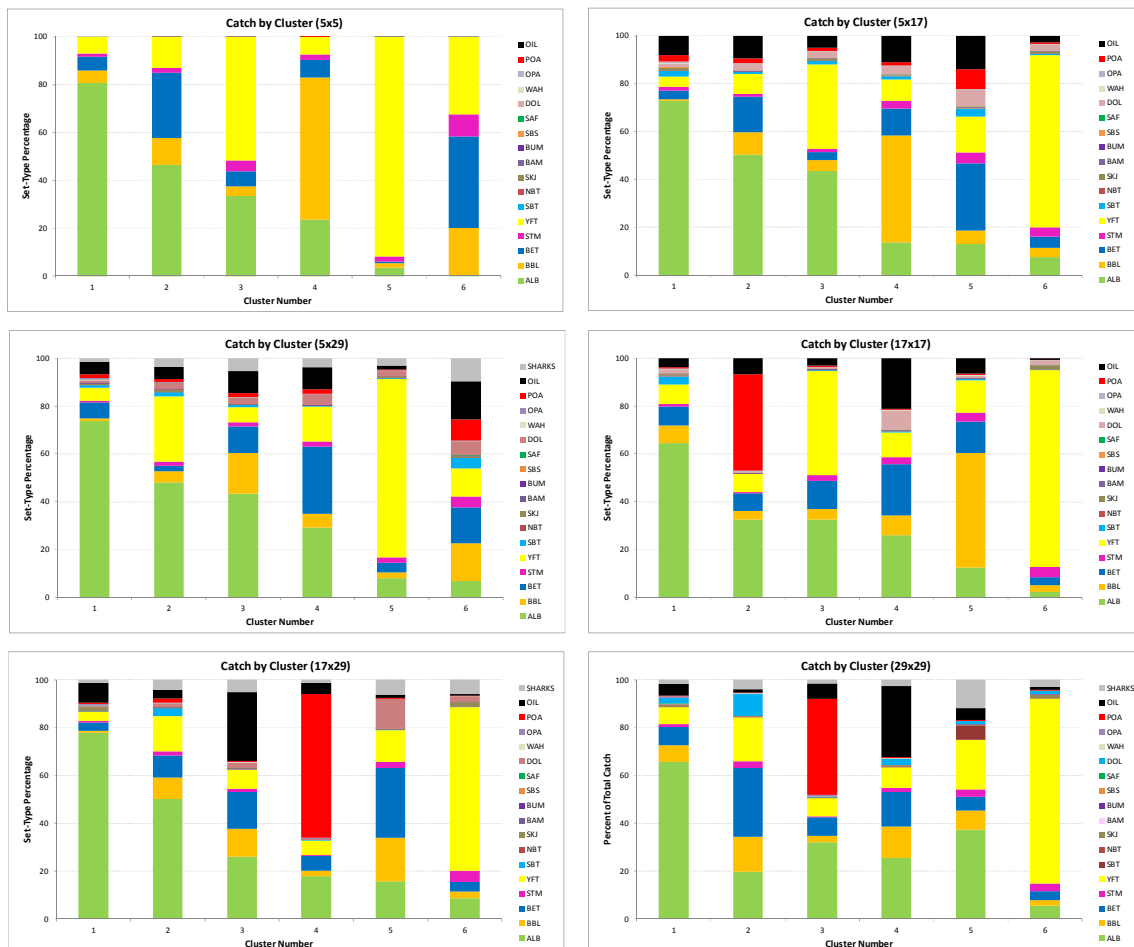


In order to investigate how the configuration of the clusters identified in the data are dependent upon which species are included in the cluster analysis, separate analyses were undertaken using different numbers of species comprising the total catch (B), and different numbers of species upon which the clusters were defined (A). Six analyses were undertaken based on the winter data set defined previously using the following definitions:

- 1) Catch comprised of B=5 species (YFT, BET, ALB, SWO, STM),
Cluster on A=5 species (YFT, BET, ALB, SWO, STM)
- 2) Catch comprised of B=17 species (no sharks),
Cluster on A=5 species (YFT, BET, ALB, SWO, STM)
- 3) Catch comprised of all B=29 species,
Cluster on A=5 species (YFT, BET, ALB, SWO, STM)
- 4) Catch comprised of B=17 species (no sharks)
Cluster on A=17 species (no sharks)
- 5) Catch comprised of all B=29 species,
Cluster on A=17 species (no sharks)
- 6) Catch comprised of all B=29 species,
Cluster on all A=29 species

The results are shown in Figure 6.19. For each of the six analyses one can identify a 'Yellowfin' and 'Albacore' dominant catch cluster and to a lesser extent a 'Bigeye cluster'. Furthermore,

Figure 6.19: Catch comparison of clusters based on analyses using different number of species (B) comprising the total catch and a subset of different species (A) upon which the clusters were defined. Individual results are labelled (Ax1B).



when the 12 sharks species are not included in the analysis (i.e. B=5 or 17), one can also identify a 'Swordfish' dominant cluster. However, the catch composition of the other clusters (not unexpectedly) varies dependent upon the suite of species included in the analyses. This indicates that some thought needs to be given to how best to cluster the catch data for classifying sets, and how to identify suites of species associated with different targeting practices. For example, should the cluster analyses used to identify 'targeting' practices be based only on the main target species in a fishery, or should such analyses also include by-product and by-catch species (even though these species are not targeted). The sensitivity of abundance indices on using the different cluster-types identified by these different types of analyses in procedures to standardise CPUE should also be investigated.

Note, in the simulation analyses undertaken for this project, the results reported in the next section are based on the 5x5 cluster analysis defined above. This was due to the fact that the simulated multispecies data for the ETBF used for these analyses only included five species (called ALB, BET, YFT, SWO and DOL).

6.4 Comparative Performance of Standardisation Models

The comparative performance of each the models listed in Section 5.7.1 using each of the four sets of trials listed in Section 5.7.2 is presented here.

6.4.1 1D-Trial and 1R-Trial: Group A and B models fitted to the ETBF simulated data-sets.

The Mean Annual Error (*MAE*) based on Type-1 and Type-2 errors for the Group A and B models fitted to the ETBF simulated data-set are shown for all 24 models compared in Figures 6.20a&b. Results are shown for each species and for the two sets of analyses fitted to the D-catches and the R-catches (note, the corresponding results for another Trial discussed in Section 6.5 are also displayed). The corresponding results for the Type-3 and Type-4 errors are not displayed, as the *MAE* for each of the four error types are highly correlated as shown in Table 6.1. A summary of range of errors across all 24 models for each species and error type for the

Table 6.1: Pearson correlation coefficient between the *MAE* for each error type for the (a) 1D-Trial and (b) 1R-Trial analyses.

(a) 1D-Trial

Group A	E-1	E-2	E-3	E-4
E-1	100.0%	98.7%	99.3%	99.8%
E-2		100.0%	97.8%	98.6%
E-3			100.0%	99.7%
E-4				100.0%

Group B	E-1	E-2	E-3	E-4
E-1	100.0%	92.1%	98.9%	99.5%
E-2		100.0%	90.5%	94.0%
E-3			100.0%	99.3%
E-4				100.0%

(b) 1R-Trial

Group A	E-1	E-2	E-3	E-4
E-1	100.0%	83.8%	99.4%	99.5%
E-2		100.0%	80.9%	88.0%
E-3			100.0%	98.9%
E-4				100.0%

Group B	E-1	E-2	E-3	E-4
E-1	100.0%	93.7%	98.4%	99.2%
E-2		100.0%	91.4%	94.5%
E-3			100.0%	99.3%
E-4				100.0%

Figure 6.20: (a) Error-1 results for the Group A and B models fitted to the ETBF simulated data-set. Results are shown for each species and for three sets of Trials described in the text.

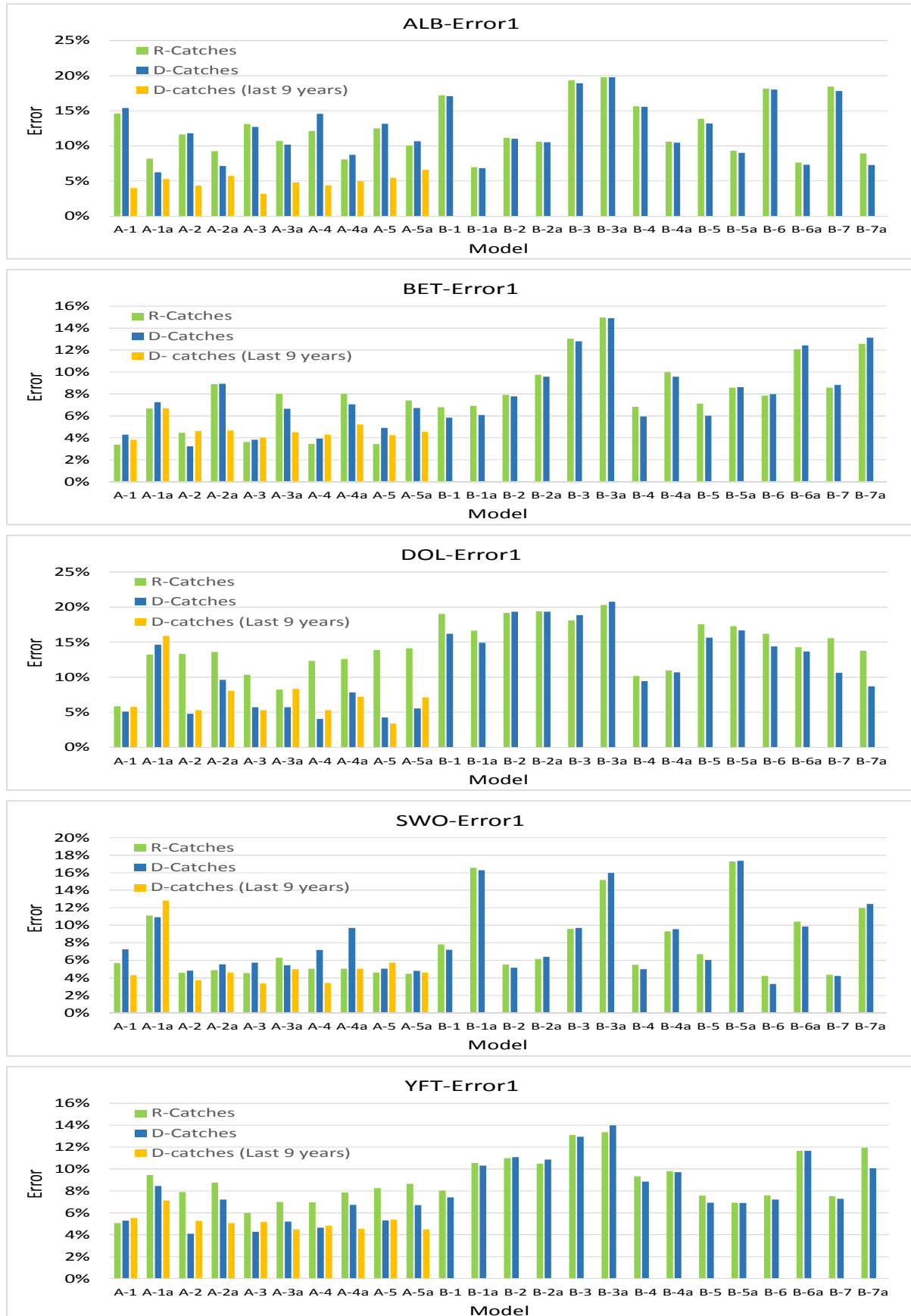


Figure 6.20: (b) Error-2 results for the Group A and B models fitted to the ETBF simulated data-set. Results are shown for each species and for three sets of Trials described in the text.

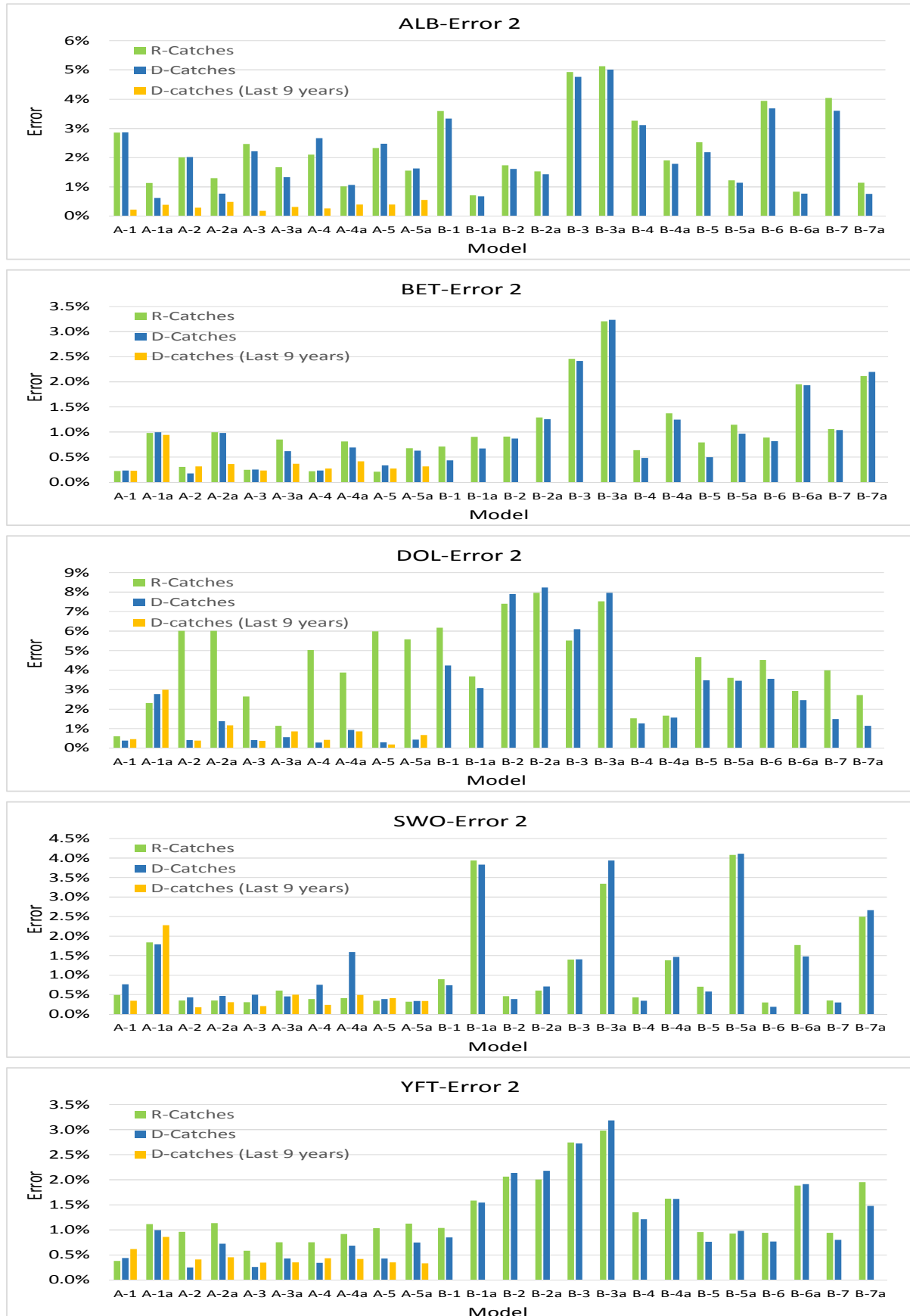


Figure 6.21: Minimum and maximum MAEs for each the four Error-types and species for the 1D-Trials and 1R-Trials.

two sets of 1D-Trial and 1R-Trial analyses is shown in Figure 6.21. For each error-type, the minimum error is seen to be similar for each species for the D-catches but displays some variation for the R-catches (which are larger than the corresponding D-catch errors), with the smallest errors found for the BET analyses and the largest errors for the ALB analyses. This pattern of variation is also seen for the maximum for each error-type and these maximum errors are similar for both the D- and R-catches. Unlike the minimum errors, the smallest errors are found for the YFT analyses and the largest errors for the DOL analyses.

For the analyses fitted to the D-catches the Species Ranked Score (*SRS*) for each species and model is shown in Table 6.2 together with the total Model Ranked Score (*MRS*) across all species, while the *SRS* for each species and the mean across all species is displayed in Figure 6.22. Several results are worth noting.

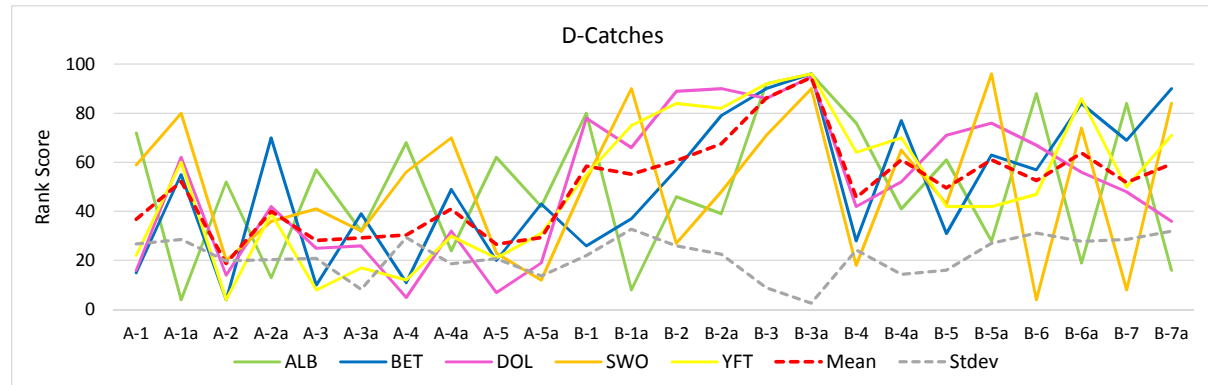
First, the performance of each model varies considerably between the analyses conducted over the five species. For example, Model A-1a is seen to be the best performing model based on the ALB analyses but is ranked 13th, 15th, 20th and 15th in comparative performance based on the BET, DOL, SWO and YFT analyses, respectively. Across all five species it is ranked 13th in performance (as measured by the *MRS*). The reasons for differential performance across the five species remains uncertain but is likely linked to differences in the spatio-temporal distribution of the different fish resources across the ETBF.

Second, the variability in the performance of the 24 models across the five species is relatively consistent (as measured by the standard deviation shown in Figure 6.22). The model with the lowest variation (Model B-3a, the PCA-linear model with no gear effects) is also the worst performing model, so there is no consistently best performing model across all species. The best performing model across all species (Model A-2, the Annual cluster by set model) is only the best model for two of the individual species analyses (BET and YFT), being 3rd and 5th best for the DOL and SWO analysis and only 13th best for the ALB analysis. Of the PCA-based models, the B-4 model (PCA-Binomial) was the best performing of the B-Group models while the two other PCA-based models (B-2 and B-3) were the two worst performing models. Of the

Table 6.2: Species Ranked Scores (SRS) for each species and model and the total Model Ranked Score (MRS) for the 1D-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same SRS.

RANK	ALB		BET		DOL		SWO		YFT		All Species	
	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	MRS	Model
1	4	A-1a	4	A-2	5	A-4	4	B-6	4	A-2	94	A-2
2	8	B-1a	10	A-3	7	A-5	8	B-7	8	A-3	133	A-5
3	13	A-2a	11	A-4	14	A-2	12	A-5a	12	A-4	141	A-3
4	16	B-7a	15	A-1	16	A-1	18	B-4	17	A-3a	146	A-3a
5	19	B-6a	20	A-5	19	A-5a	20	A-2	21	A-5	147	A-5a
6	24	A-4a	26	B-1	25	A-3	23	A-5	22	A-1	152	A-4
7	28	B-5a	28	B-4	26	A-3a	27	B-2	30	A-4a	184	A-1
8	32	A-3a	31	B-5	32	A-4a	32	A-3a	31	A-5a	200	A-2a
9	39	B-2a	37	B-1a	36	B-7a	36	A-2a	39	A-2a	205	A-4a
10	41	B-4a	39	A-3a	42	A-2a	41	A-3	42	B-5	228	B-4
11	42	A-5a	43	A-5a	42	B-4	43	B-5	42	B-5a	248	B-5
12	46	B-2	49	A-4a	48	B-7	48	B-2a	47	B-6	259	B-7
13	52	A-2	55	A-1a	52	B-4a	53	B-1	50	B-7	261	A-1a
14	57	A-3	57	B-2	56	B-6a	56	A-4	55	B-1	263	B-6
15	61	B-5	57	B-6	62	A-1a	59	A-1	60	A-1a	276	B-1a
16	62	A-5	63	B-5a	66	B-1a	65	B-4a	64	B-4	292	B-1
17	68	A-4	69	B-7	67	B-6	70	A-4a	70	B-4a	297	B-7a
18	72	A-1	70	A-2a	71	B-5	71	B-3	71	B-7a	303	B-2
19	76	B-4	77	B-4a	76	B-5a	74	B-6a	75	B-1a	305	B-4a
20	80	B-1	79	B-2a	78	B-1	80	A-1a	82	B-2a	305	B-5a
21	84	B-7	84	B-6a	86	B-3	84	B-7a	84	B-2	319	B-6a
22	88	B-6	90	B-3	89	B-2	90	B-1a	86	B-6a	338	B-2a
23	92	B-3	90	B-7a	90	B-2a	90	B-3a	92	B-3	431	B-3
24	96	B-3a	96	B-3a	95	B-3a	96	B-5a	96	B-3a	473	B-3a

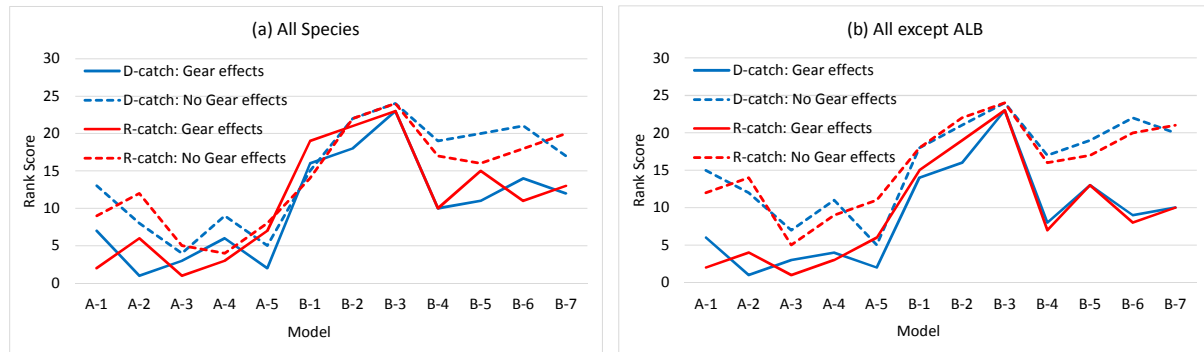
Figure 6.22: Comparison of the Species Ranked Scores (SRS) for each species and model for the 1D-Trial analyses. The mean and standard deviation of the SRS across all five species is also shown.



cluster-based B-Group models, the *kmeans*-untransformed cluster by set model (B-5) was the best, followed closely by the *kmeans*-transformed cluster by set model (B-7).

Third, for the ALB analyses, and except for Model B-3a, all the models without the gear effects included in the linear predictor (the 'a' models) performed better than those with the gear effects included (c.f. Table 6.2). Indeed, the two Base models without gear effects (Models A-1a and B-1a) are the best performing models for this species. This is unlike the results for the other species, where the performance of these two sets of models is more mixed with the models with gear effects included in most instances outperforming those without gear effects (c.f. Figure 6.23). The out-performance of the models with gear effects included is seen to be

Figure 6.23: Comparison of the rank of models which include the two gear effects in the fitted GLM with those models which do not include these effects. The comparison is shown for both the 1-D and 1-R analyses for (a) all species and (b) the four species other than ALB.



most pronounced for the analyses conducted on the SWO catches, where for example, the Error-2 results are less than 0.5% for when the gear effects are included, but greater than 3% when they are not included (Figure 6.20b). The reason for the poorer performance of the models with gear effects for the ALB-catches remains uncertain, but is somewhat surprising as it was the introduction of the 'deep' longlining technique (associated with the use of HPF settings greater than 20) in 2006 that is usually associated with the shift to targeting the deeper species such as Albacore Tuna that the simulated ALB was supposed to mimic (c.f. Figure 6.2).

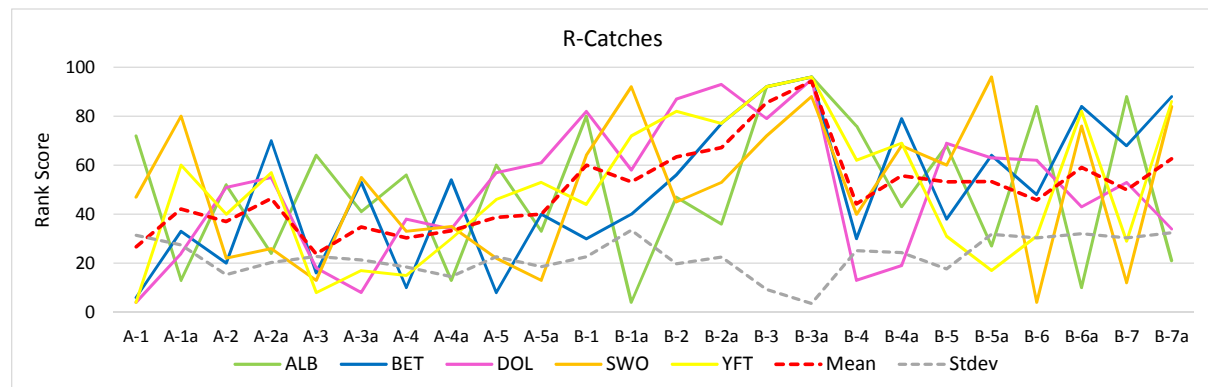
Finally, it is interesting to compare only the relative performance of the models without the gear effects included. This is because the gear effects are likely to have the larger influence on identifying the targeting strategy associated with a particular set, than that identified by the other (secondary and assumed) targeting effect included in the model based on clustering the catch. Secondly, many fisheries do not collect information on the particular gear configuration associated with any given set, with the consequence that the available data is generally limited to information on the catch alone. In these situations the only models available for standardising the CPUE would be the models without the gear effects.

Of the Group-A models which did not include gear effects, Model A-3a (Total cluster by trip model) was the best performing model, closely followed by Model A-5a (Bi-monthly cluster by set model), while Model A-1a (which is the only model which does not include a targeting effect) was the worst performing model (c.f. Table 6.2). This last result indicates that each of the models that includes a targeting effect performs better than models that exclude targeting effects, and that the use of catch-based clusters to identify 'targeting' has merit. It is also noteworthy that the cluster by trip model out-performs the cluster by set model when the gear effects are not included. This may indicate that without direct information on the targeting strategy associated with each set (as provided by gear information), the variability in species composition at the set level (due to the likely random nature of encounters between fishing gear and fish) may result in a misallocation of sets to specific fishing strategies, and that aggregating the data across trip reduces both this variability and the potential for misallocation. Of the Group-B models, and excluding the anomalous results for ALB, the B-4a model (PCA-Binomial) was the best performing, with the other models having a similar performance except for the B-3a model which was the worst performing.

Table 6.3: Species Ranked Scores (SRS) for each species and model and the total Model Ranked Score (MRS) for the 1R-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same SRS.

RANK	ALB		BET		DOL		SWO		YFT		All Species	
	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	MRS	Model
1	4	B-1a	6	A-1	4	A-1	4	B-6	4	A-1	119	A-3
2	10	B-6a	8	A-5	8	A-3a	12	B-7	8	A-3	133	A-1
3	13	A-1a	10	A-4	13	B-4	13	A-3	15	A-4	152	A-4
4	13	A-4a	16	A-3	18	A-3	13	A-5a	17	A-3a	166	A-4a
5	21	B-7a	20	A-2	19	B-4a	22	A-2	17	B-5a	174	A-3a
6	24	A-2a	30	B-1	24	A-1a	22	A-5	29	B-7	185	A-2
7	27	B-5a	30	B-4	34	A-4a	26	A-2a	30	A-4a	193	A-5
8	33	A-5a	33	A-1a	34	B-7a	33	A-4	31	B-5	200	A-5a
9	36	B-2a	38	B-5	38	A-4	35	A-4a	31	B-6	210	A-1a
10	41	A-3a	40	A-5a	43	B-6a	40	B-4	40	A-2	221	B-4
11	43	B-4a	40	B-1a	51	A-2	45	B-2	44	B-1	229	B-6
12	47	B-2	48	B-6	53	B-7	47	A-1	46	A-5	232	A-2a
13	52	A-2	53	A-3a	55	A-2a	53	B-2a	53	A-5a	250	B-7
14	56	A-4	54	A-4a	57	A-5	55	A-3a	57	A-2a	266	B-1a
15	60	A-5	56	B-2	58	B-1a	60	B-5	60	A-1a	266	B-5
16	64	A-3	64	B-5a	61	A-5a	64	B-1	62	B-4	267	B-5a
17	68	B-5	68	B-7	62	B-6	68	B-4a	69	B-4a	278	B-4a
18	72	A-1	70	A-2a	63	B-5a	72	B-3	72	B-1a	295	B-6a
19	76	B-4	77	B-2a	69	B-5	76	B-6a	77	B-2a	300	B-1
20	80	B-1	79	B-4a	79	B-3	80	A-1a	82	B-2	313	B-7a
21	84	B-6	84	B-6a	82	B-1	84	B-7a	82	B-6a	317	B-2
22	88	B-7	88	B-7a	87	B-2	88	B-3a	86	B-7a	336	B-2a
23	92	B-3	92	B-3	93	B-2a	92	B-1a	92	B-3	427	B-3
24	96	B-3a	96	B-3a	95	B-3a	96	B-5a	96	B-3a	471	B-3a

Figure 6.24: Comparison of the Species Ranked Scores (SRS) for each species and model for the 1R-Trial analyses. The mean SRS across all five species is also shown.



For the analyses fitted to the R-catches the Species Ranked Score (SRS) for each species and model is shown in Table 6.3, together with the total Model Ranked Score (MRS) across all species. The SRS for each species and the mean across all species is displayed in Figure 6.24. Again, several results are worth noting. First, a comparison how accurately the standardised indices match the assumed relative abundances, as measured by the mean of the Type 1 and 2 errors across all species, for both the D-catch and R-catch analyses is shown in Figure 6.26, indicating that in general the model fits are better for the D-catch analyses (both sets with and without the gear effects included). This relative out-performance is greater for the Group-A models but is quite small for the three PCA-based models (B-2, B-3 and B-4). Given the random element added to the R-catches, these differences are not unexpected and likely linked to the higher proportion of zeroes in the R-catches (c.f. Figure 5.20).

Figure 6.25: Comparison of the model ranks (1-24) between the 1-D and 1-R analyses by species.



Figure 6.26: Comparison of the mean Error-1 and Error-2 results across all species for models fitted to the D-catches and the R-catches. The comparison is also shown for both sets of analyses which include the two gear effects in the fitted GLM with those models which do not include these effects.

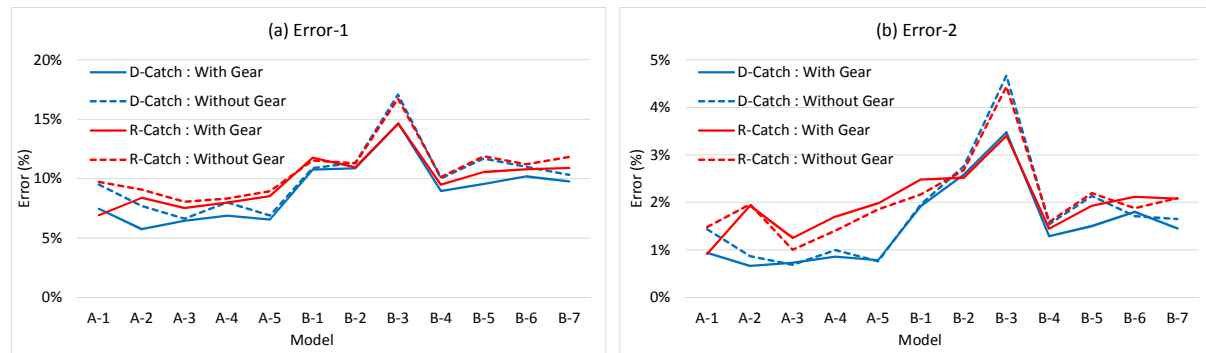


Table 6.4: Pearson correlation, R , between the model ranks (1-24) for the 1-D and 1-R analyses.

Species	ALB	BET	DOL	SWO	YFT	All
R	97.5%	96.1%	67.6%	87.4%	86.5%	92.5%

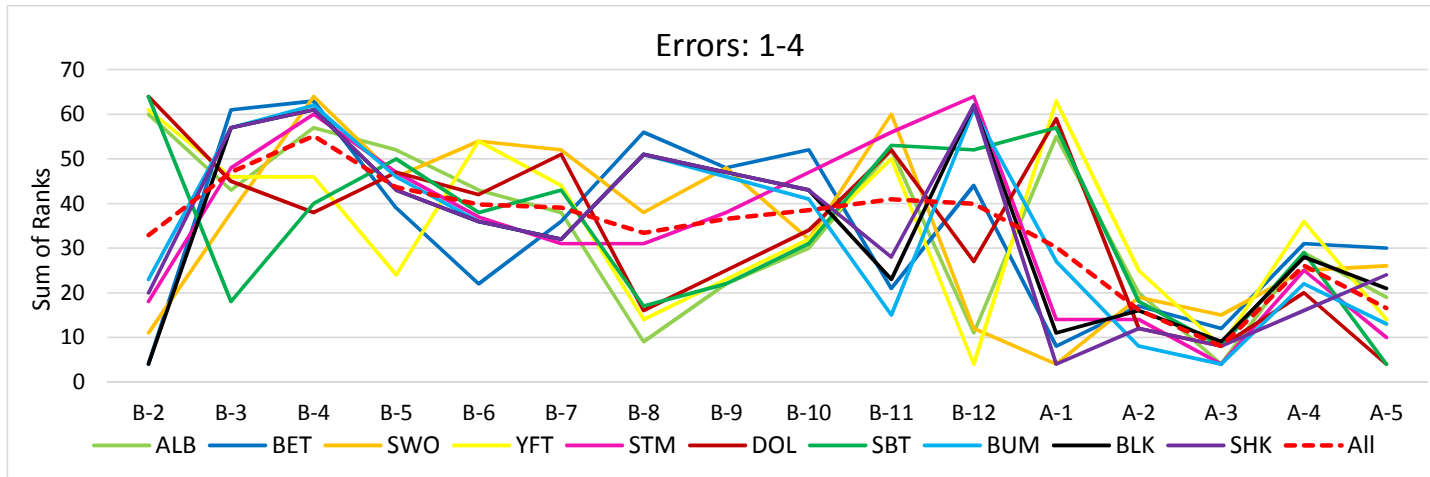
Second, the high variability in the performance of each model across the five species is similar to that noted for the D-catch results, again indicating that no single model is seen as being the best performer across all species. Also, the out-performance of the models without the gear effects included for the ALB analyses is retained. Third, a comparison of the overall Rank of each model fitted to both the D- and R-catches for each species and across all species is shown in Figure 6.25, and despite some differences indicates a relatively high degree of similarity between the two sets of analyses for most species. This similarity is highest for the ALB and BET analyses and lowest for the DOL analyses (c.f. Table 6.4).

While the A-1 (Base delta model) is the best performing across three species for the R-catch analyses (and is ranked 2nd across all species), the best performance of this model was 4th for the D-catch analyses (where it was ranked 7th across all species). The higher performance of the A-1 model fitted to the R-catches is most likely due to the fact that the influence of the cluster-based targeting effects has been reduced, due to the higher variability in the catches (with the random component added) and the greater influence of the gear effects alone. The best performing model overall for the R-catch analyses (model A-3: Total cluster by trip) was not the best model for any of the individual species, being 2nd, 3rd, 4th and 4th best for the YFT, SWO, DOL and BET analyses (respectively) and only 16th best for the ALB analysis. It was ranked 3rd overall for the D-based analyses but was the best performing model based on the sum of the *MRS* for both sets of analyses. On the other hand, the best performing model overall for the D-catch analyses (model A-2: annual cluster by set) was ranked 6th overall for the R-based analyses, but was the 2nd best performing model across both sets of analyses. The third best performing model across both catch types was model A-4 (Monthly cluster by set), but was ranked 6th and 3rd for the D- and R-based analyses respectively. Of the Group-B models, the B-4 model (PCA-Binomial) was again the best performing while the two other PCA-based models (B-2 and B-3) were the two worst performing models. Of the Group-B cluster-based models, model B-6 (*clara* clustering) was the best performing, unlike that for the D-catch analyses where the *kmeans*-based cluster models were better performing. Interestingly, the B-6 model was the best performing for the SWO analyses across both sets of analyses. Finally, of the models that did not include the gear effects, and again excluding the anomalous result seen for ALB, the A-3a model was the best performer.

Table 6.5: Species Ranked Scores (SRS) for each species and model and the total Model Ranked Score (MRS) for the 2R-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same SRS.

RANK	ALB		BET		SWO		YFT		STM		DOL		SBT		BUM		BLK		SHK		TOTAL	
	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	MRS	Model
1	4	A-3	4	B-2	4	A-1	4	B-12	4	A-3	4	A-5	4	A-5	4	A-3	4	B-2	4	A-1	80	A-3
2	9	B-8	8	A-1	11	B-2	8	A-3	10	A-5	8	A-3	8	A-3	8	A-2	9	A-3	8	A-3	161	A-2
3	11	B-12	12	A-3	12	B-12	14	B-8	14	A-1	12	A-2	17	B-8	13	A-5	11	A-1	12	A-2	165	A-5
4	19	A-5	17	A-2	15	A-3	14	A-5	14	A-2	16	B-8	18	B-3	15	B-11	16	A-2	16	A-4	261	A-4
5	20	A-2	21	B-11	19	A-2	23	B-9	18	B-2	20	A-4	18	A-2	22	A-4	21	A-5	20	B-2	302	A-1
6	22	B-9	22	B-6	25	A-4	24	B-5	25	A-4	25	B-9	22	B-9	23	B-2	23	B-11	24	A-5	329	B-2
7	29	A-4	30	A-5	26	A-5	25	A-2	31	B-7	27	B-12	29	A-4	27	A-1	28	A-4	28	B-11	334	B-8
8	30	B-10	31	A-4	32	B-10	32	B-10	31	B-10	34	B-10	31	B-10	32	B-7	32	B-7	32	B-7	366	B-9
9	38	B-7	36	B-7	38	B-3	36	A-4	37	B-6	38	B-4	38	B-6	36	B-6	36	B-6	36	B-6	385	B-10
10	43	B-3	39	B-5	38	B-6	44	B-7	38	B-9	42	B-6	40	B-4	41	B-10	43	B-5	43	B-5	391	B-7
11	43	B-8	44	B-12	46	B-5	46	B-3	47	B-5	45	B-3	43	B-7	46	B-5	43	B-10	43	B-10	398	B-6
12	52	B-5	48	B-9	48	B-9	46	B-4	47	B-9	47	B-5	50	B-5	46	B-7	47	B-9	47	B-9	399	B-12
13	52	B-11	52	B-10	52	B-7	50	B-11	48	B-3	51	B-7	52	B-12	51	B-8	51	B-8	51	B-8	410	B-11
14	55	A-1	56	B-8	54	B-6	54	B-6	56	B-11	52	B-11	53	B-11	57	B-3	57	B-3	57	B-3	437	B-5
15	57	B-4	61	B-3	60	B-11	61	B-2	60	B-4	59	A-1	57	A-1	61	B-12	61	B-4	61	B-4	470	B-3
16	60	B-2	63	B-4	64	B-4	63	A-1	64	B-12	64	B-2	64	B-2	62	B-4	62	B-12	62	B-12	552	B-4

Figure 6.27: Comparison of the Species Ranked Scores (SRS) for each species and model for the 2R-Trial analyses. The mean SRS across all ten species is also shown.

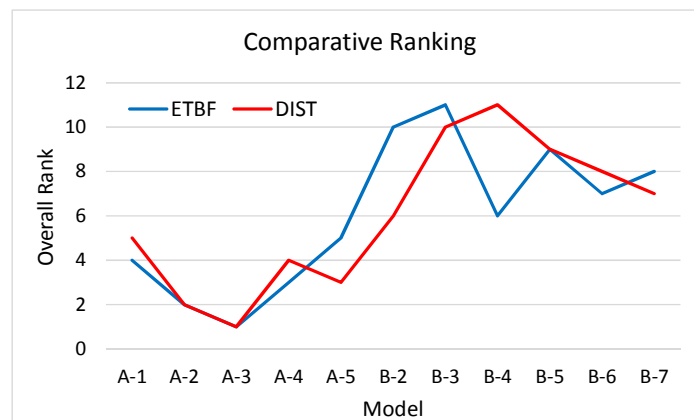


6.4.2. 2R-Trial: Group A and B models fitted to the DIST simulated data-set using the random R-catches.

For the analyses fitted to the R-catches, the Species Ranked Score (*SRS*) for each species and model is shown in Table 6.5, together with the total Model Ranked Score (*MRS*) across all species, while the *SRS* for each species and the mean across all species is displayed in Figure 6.27. Several results can be noted. First, similar to the 1D-Trial and 1R-Trial results there is high variability among model performance across the species, with no model performing best across all species. While model A-3 is the best performing for three of the ten species, and best performing overall, model B-12 is the best performing for YFT but is ranked only 12th overall. The variability in performance can be highly bi-modal across species for some models, for example models A-1 and B-12.

Second, similar to the 1D-Trial and 1R-Trial results, across all species each of the Group-A models out-perform all of the Group-B models (c.f. Figure 6.28). Of the Group-A models, model A-3 is the best performing and model A-2 is the second best performing. This result is consistent with the 1R-Trial results, as these two models were also the best performing for these analyses. Of the Group-B models, model B-2 (the PCA-ns model) is the best performing, but this result is different from that found for the 1R-Trial analyses where of the Group-B models the B-4 (PCA-binomial) model performed best. Interestingly, of the cluster-based models within Group-B, the *Ward*-cluster model (B-8) was the second best performing in the 2R-Trial and this cluster method is the same as that used in the Group-A models.

Figure 6.28: Comparison of the overall rank of the eleven models used in both the 1R-Trial (ETBF) and 2R-Trial (DIST) analyses.



6.4.3 3D-Trial: Group A and C models fitted to the N3-version of the ETBF simulated data-set using the deterministic D-catches.

Finally, for the analyses fitted to the D-catches pertaining to the smaller set (N3) of the ETBF-simulated data, the Error-1 and Error-2 results are shown in Figure 6.29. Species Ranked Score (*SRS*) for each species and model is shown in Table 6.6, together with the total Model Ranked Score (*MRS*) across all species. Finally, the *SRS* for each species and the mean across all species is displayed in Figure 6.30.

These results, which can be compared with the 1D-Trial results (c.f. Table 6.1), show similar variability in model performance across species. Furthermore, as noted previously, the models including gear effects generally out-perform those without gear effect. However, while this

Figure 6.29: Error-1 and Error-2 results for the Group A (blue) and C models (green) fitted to the streamlined (N3) ETBF simulated data-set. Results are shown for each species. Note, y-axis values differ between panels.

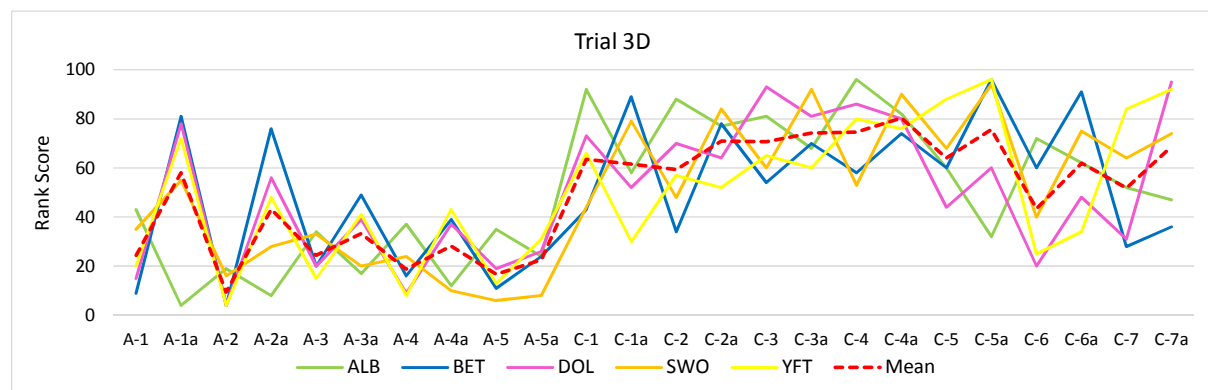


difference in performance is seen for all the A-Group models, this relative out-performance is only apparent for several of the C-Group models (e.g. C-5, C-6 and C-7). For three of the five species the best performing model is the same as that observed in the 1D-Trial analyses, with model A-2 relacing model A-4 for DOL and model A-5 replacing model B-6 for SWO. Over all species, the best performing model was model A-2, while the second best performing model was model A-5 – these are the same results as for the 1D-Trial analyses. The best performing of the C-Group was model C-6, which is the INLA discrete and continuous spatio-temporal model, and the most complex model within this group. While this model only out-performed one of the A-Group models (model A-1a) its performance was substantially better than most of the other models in the C-Group.

Table 6.6: Species Ranked Scores (SRS) for each species and model and the total Model Ranked Score (MRS) for the 3D-Trial analyses. All results are ordered from lowest to highest rank. Shading indicates models with the same SRS.

RANK	ALB		BET		DOL		SWO		YFT		TOTAL	
	SRS	Model	SRS	Model	SRS	Model	SRS	Model	SRS	Model	MRS	Model
1	4	A-1a	4	A-2	4	A-2	6	A-5	4	A-2	47	A-2
2	8	A-2a	9	A-1	9	A-4	8	A-5a	8	A-4	84	A-5
3	12	A-4a	11	A-5	15	A-1	10	A-4a	13	A-5	94	A-4
4	17	A-3a	16	A-4	19	A-5	16	A-2	15	A-3	113	A-5a
5	19	A-2	20	A-3	20	A-3	20	A-3a	20	A-1	122	A-1
6	24	A-5a	24	A-5a	20	C-6	24	A-4	25	C-6	122	A-3
7	32	C-5a	28	C-7	26	A-5a	28	A-2a	30	C-1a	141	A-4a
8	34	A-3	34	C-2	31	C-7	33	A-3	31	A-5a	166	A-3a
9	35	A-5	36	C-7a	37	A-4a	35	A-1	34	C-6a	216	A-2a
10	37	A-4	39	A-4a	39	A-3a	40	C-6	41	A-3a	217	C-6
11	43	A-1	43	C-1	44	C-5	44	C-1	43	A-4a	259	C-7
12	47	C-7a	49	A-3a	48	C-6a	48	C-2	48	A-2a	290	A-1a
13	52	C-7	54	C-3	52	C-1a	53	C-4	52	C-2a	297	C-2
14	58	C-1a	58	C-4	56	A-2a	55	A-1a	57	C-2	308	C-1a
15	60	C-5	60	C-5	60	C-5a	60	C-3	60	C-3a	310	C-6a
16	62	C-6a	60	C-6	64	C-2a	64	C-7	65	C-3	318	C-1
17	68	C-3a	70	C-3a	70	C-2	68	C-5	66	C-1	320	C-5
18	72	C-6	74	C-4a	73	C-1	74	C-7a	72	A-1a	344	C-7a
19	77	C-2a	76	A-2a	78	A-1a	75	C-6a	76	C-4a	353	C-3
20	81	C-3	78	C-2a	80	C-4a	79	C-1a	80	C-4	355	C-2a
21	82	C-4a	81	A-1a	81	C-3a	84	C-2a	84	C-7	371	C-3a
22	88	C-2	89	C-1a	86	C-4	90	C-4a	88	C-5	373	C-4
23	92	C-1	91	C-6a	93	C-3	92	C-3a	92	C-7a	378	C-5a
24	96	C-4	96	C-5a	95	C-7a	94	C-5a	96	C-7	402	C-4a

Figure 6.30: Comparison of the Species Ranked Scores (SRS) for each species and model for the 3D-Trial analyses. The mean SRS across all five species is also shown.



6.5 Further Analyses

The results presented in the previous Section raise a number of questions as to the reason behind the relative performance of each model. While a full investigation of these differences is beyond the scope of this project, in this Section we focus on the following three issues:

- 1) the relative out-performance of the Group-A of models,
- 2) the out-performance of the models without gear-effects for the ALB analyses, and
- 3) the selection of the GLM *Area* effects.

6.5.1 Group-A Models Out-Performance

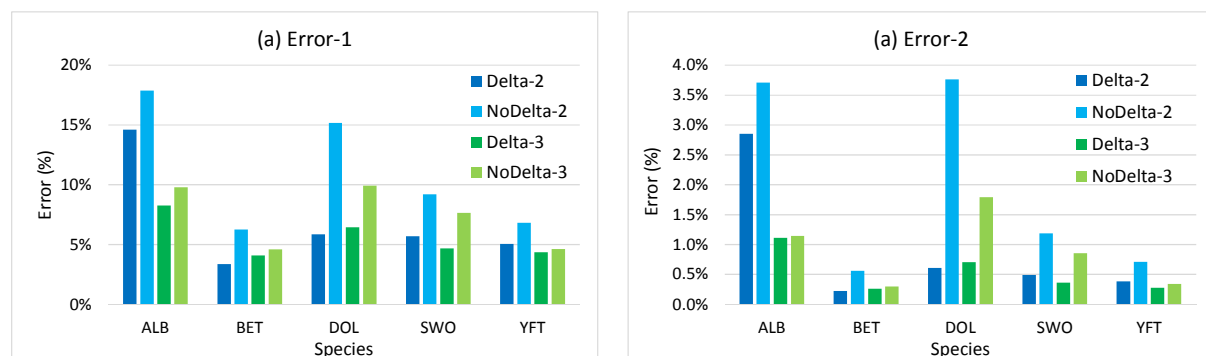
Apart from differences on how the 'target' effect was added to the linear predictor in both the Group-A and Group-B models, the other main difference was that the Group-A models utilised a two-stage analysis, with the first stage analysing the distribution of zero catches in the data and the second stage analysing the distribution of non-zero catches. As noted previously, this approach is often used because the number of zeroes in the catch data is usually 'inflated', and is assumed to result in a distribution of errors used in a single-stage analysis that will not be appropriate (i.e. it will be unable to adequately explain the large proportion of zero catches).

To explore the implications using a two-stage versus a single-stage analysis, the following set of analyses were undertaken. First, the A-1 model was fitted to the 100 sets of R-catches for the ETBF simulated data set as previously undertaken in the 1R-Trial. Second, the first stage Binomial model was dropped and instead all the data were fitted just to the Negative-Binomial model. These will be called the Delta-2 and NoDelta-2 analyses respectively. The analysis was also extended to investigate the relative performance of using a 3-way interaction between the *Year*, *Quarter* and *Area* effects, instead of just the two 2-way interactions used in all previous analyses. For this purpose, the above two analyses were repeated using the following linear-effects model:

$$I + Year*Qtr*Area + HPF + STIME$$

Following the notation used above, these will be called the Delta-3 and NoDelta-3 analyses respectively. Finally, the annual index of abundance was calculated for all four fitted models together with the corresponding deviations (Errors 1-4) from the true underlying index. A comparison of these errors for the four models is shown in Figure 6.31.

Figure 6.31: Comparison of the four index-based errors after fitting the following four variates of the A-1 GLM to the 100 sets of R-catches for the ETBF simulated data set: (i) a two-stage delta- (Bin-NegBin) model with either the two *Year*Qtr* and *Qtr*Area* 2-way interactions (Delta-2) or the single 3-way *Year*Qtr*Area* interaction (Delta-3) and (ii) a single-stage-NegBin-only model with either the two *Year*Qtr* and *Qtr*Area* 2-way interactions (NoDelta-2) or the single 3-way *Year*Qtr*Area* interaction (NoDelta-3).



Several results can be noted. First, for both error types and for all species, the Delta-2 model out-performs the NoDelta-2 model. The improvement in performance between these two models can be relatively substantial. For example, the type-1 error for DOL decreases by 61% (from around 15% to 6%), and by 38% averaged across all species. Similarly, the Delta-3 model out-performs the NoDelta-3 model for all species. The out-performance of the two-stage Delta models, compared to the single-stage non-Delta approach, helps to explain the out-performance observed in the previous Trials by the Group-A set of models in comparison to the Group-B and Group-C models. However, what proportion of the out-performance is explained by this single factor remains uncertain but it could be substantive, especially for some species.

Second, comparison of the results for those models including the two 2-way interactions with those models which include the single 3-way interaction, indicates in most instances the latter model out-performs the former. Indeed, for both error types and for all species the NoDelta-3 model out-performs the NoDelta-2 model, while for only three of the five species (ALB, SWO and YFT) does the Delta-3 model out-perform the Delta-2 model. The reasons for this differential relative performance across the different species remains unknown (perhaps based on temporal changes in the spatial distribution of the zero catches) but should be investigated.

6.5.2 Anomalous ALB Result

As noted in the previous Section, and unlike the results for the other species, the results of the 1-Trial analyses for ALB indicated that all models (except one) without the gear effects included in the linear predictor out-performed the models with the gear effects included (c.f. Tables 6.2 & 6.3). This result was seen as surprising, as it was the introduction of the 'deep' longlining technique that is usually associated with the pronounced shift to targeting of ALB that was undertaken in 2006 (c.f. Figure 6.2).

Interactions with gear effects

The cluster analyses of the ETBF catch data undertaken in Section 6.3 indicated strong seasonal changes in species composition of the major clusters identified, but for most clusters there was a relatively wide use of gears configurations. Furthermore, for similar cluster types there was a difference in the mix of gears between the seasonal clusters. For example, the mix of gear types associated with the fishing operations included with the summer and winter YFT and ALB clusters identified in Section 6.3 is shown in Figure 6.32. The large differences in the mix of gears associated with different clusters indicates that there is a strong *gear-by-cluster* interaction and the differences in the mix of gears associated with similar clusters at different seasons indicates that for some clusters there may also be a *gear-by-season* (e.g. *gear-by-quarter*) interaction.

In order to investigate whether the inclusion of either of these two interactions improves the performance of the GLMs used to standardise the CPUE of the simulated catch and effort data, especially for ALB, two alternative sets of analyses were undertaken where for both the A-3 and A-5 models described in Section 5.7 the linear predictor was updated to become:

$$I + \text{Year} * \text{Qtr} + \text{Qtr} * \text{Area} + \text{HPF} * \text{CLUSTER} + \text{STIME} * \text{CLUSTER}$$

The updated A-3 model investigates the inclusion of only the *gear-by-cluster* interaction, while the updated A-5 model also allows the *gear-by-cluster* effects to have different relative effects within each of the bi-monthly periods (as explained previously). Both models were fitted to the 100 simulated ETBF data-sets with the R-catches and again the Mean Annual Error (MAE) was

Figure 6.32: Gear types associated with fishing operations included in the summer and winter YFT and ALB clusters identified in Section 6.3 (c.f. Figure 6.17)

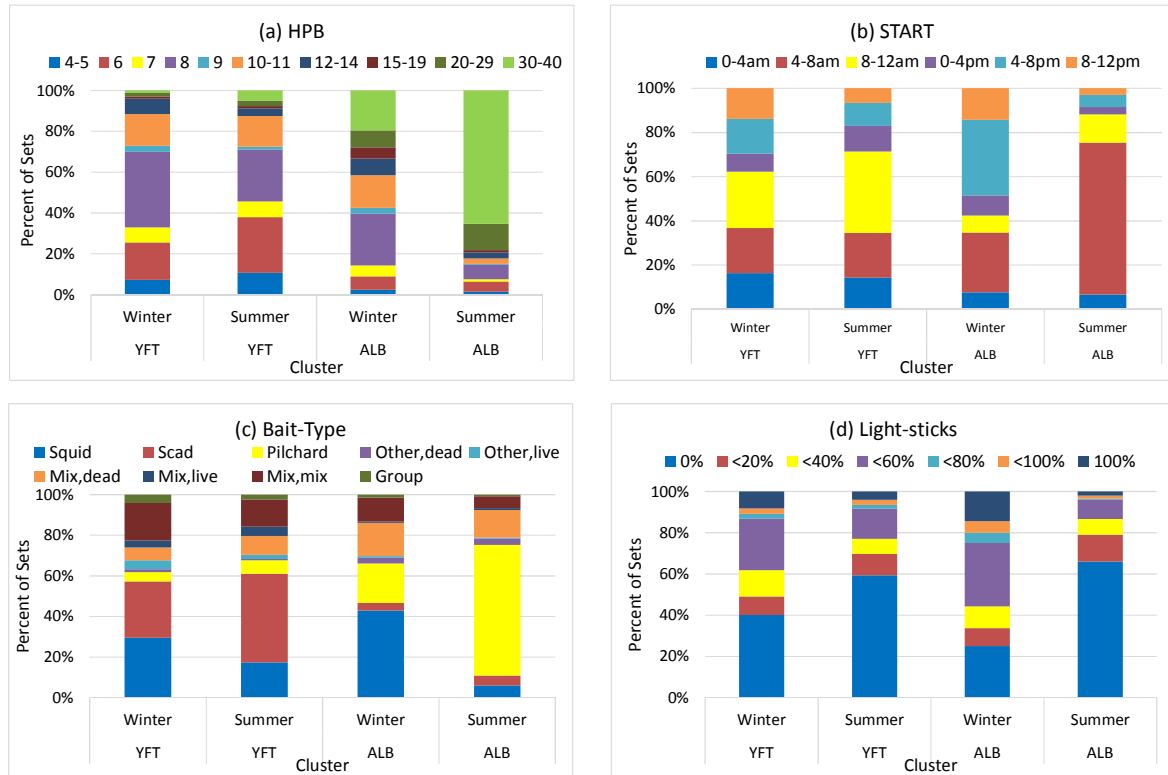


Figure 6.33: (a) Comparison of the MAE associated with the fitting the A-3 model to the ETBF data-set where the linear predictor includes; (i) no gear effects, i.e. cluster effects only, (ii) both gear and cluster effects, or (iii) the gear-by-cluster interactions.

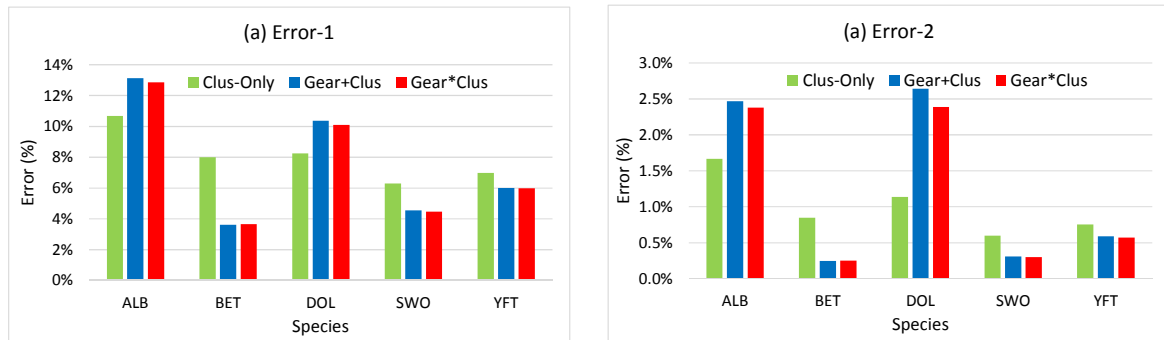
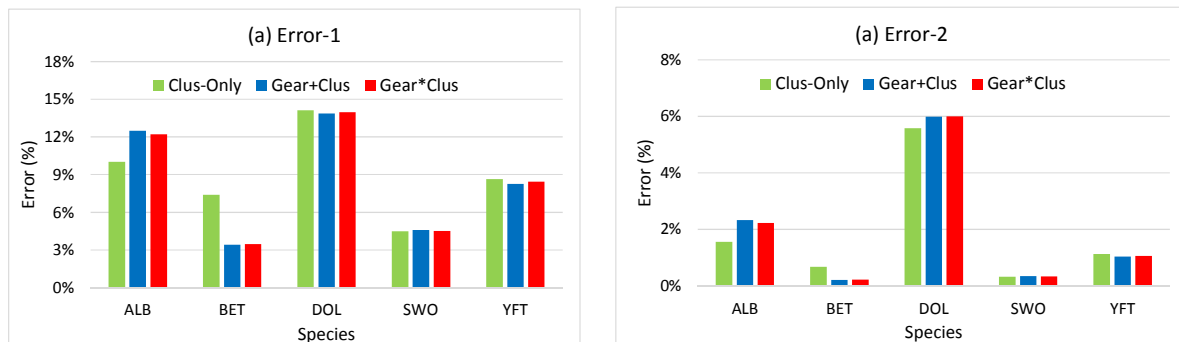


Figure 6.33: (b) Comparison of the MAE associated with the fitting the A-5 model to the ETBF data-set where the linear predictor includes; (i) no gear effects, i.e. cluster effects only, (ii) both gear and cluster effects, or (iii) the gear-by-cluster interactions.



calculated for each species. The results are shown in Figures 6.33a&b for both the Type-1 and Type-2 errors where they are also compared with the results from the previously fitted models where there was either (i) no gear effects included (i.e. only a cluster effect), or (ii) only the gear and cluster effects alone (i.e. no interaction).

Both sets of results indicate that the inclusion of the *gear*cluster* interaction terms in the fitted models makes little overall difference, if any, to the performance of the models. Whether the simulated ETBF data contains *gear*cluster* interaction effects like those observed in the real ETBF data remains uncertain. However, a more likely explanation for these results is that there is a substantial random element in the relationship between the deployment of the fishing gears and the resulting catch and its composition so that the relationship between these two is weak. Indeed, this is consistent with the results of the analyses presented in Section 6.3 which did not find particularly strong relationships between the catch-composition of the clusters identified in the ETBF data and the gear settings. Furthermore, relationships that do exist may not be consistent over time (implying an interaction with year).

The factors that influence the size and composition of the catch of any fishing operation are likely to be complex and subject to a reasonable degree of natural variability due to the random nature of fish interacting with longline hooks. No doubt further research is required to identify and improve our understanding of the factors used by fishers to 'target' the effort deployed at specific species and how this is reflected in the composition of the catch for individual fishing operations, in particular those factors that account for the variability seen in the composition of the catch given similar gear settings. These factors are likely to include spatial features in the ocean such as temperature fronts and eddies that are likely to influence the distribution of fish in the oceans and which are taken as cues by fishers for setting their gears.

Spatio-temporal changes

In lieu of the gear effects not providing the necessary explanatory power for improving the estimation of the ALB index, we explored possible spatio-temporal changes in the distribution of the catch rates in the simulated data. For this purpose, we calculated the Pearson correlation, R , between the annual distributions of nominal CPUE between years. The analysis was undertaken at two spatial scales: (i) at the 1x1-degree level, and (ii) at the GLM-area level. Furthermore, in order to identify any distinct pattern in the ALB data that is different from that associated with the other species, the analysis was undertaken across all five species. The results are displayed in Figure 6.34.

The pattern of annual correlations for ALB based on the distribution of 1-degree spatial CPUE shows a distinct change after 2005. The correlation between any year up to 2005 and any other year is generally low, with R mainly between 0.25 and 0.40. However, the correlation between any year after 2005 and any other year is generally high, with R mainly between 0.7 and 0.9. This pattern indicates a shift in the spatial distribution of catch rates after 2005 and is apparent in the annual distributions of CPUE displayed in Figure 6.35. Up until 2005 the location of squares with high CPUE was spread around the periphery of the fishery, but from 2006, and persisting across all remaining years, there is a distinct area of high CPUE (>12) located in the north-east of the fishery. The pattern of annual correlations for ALB based on the distribution of GLM-area CPUE also displays a distinct change after 2005, but also indicates another change after 2010. These changes in the distribution of the ALB resource between years infers a large *Year*Area* interaction between these two effects.

Figure 6.34: Pearson correlation between the annual distributions of nominal CPUE between years at: (a) the 1x1-degree level, and (b) the GLM-area level.

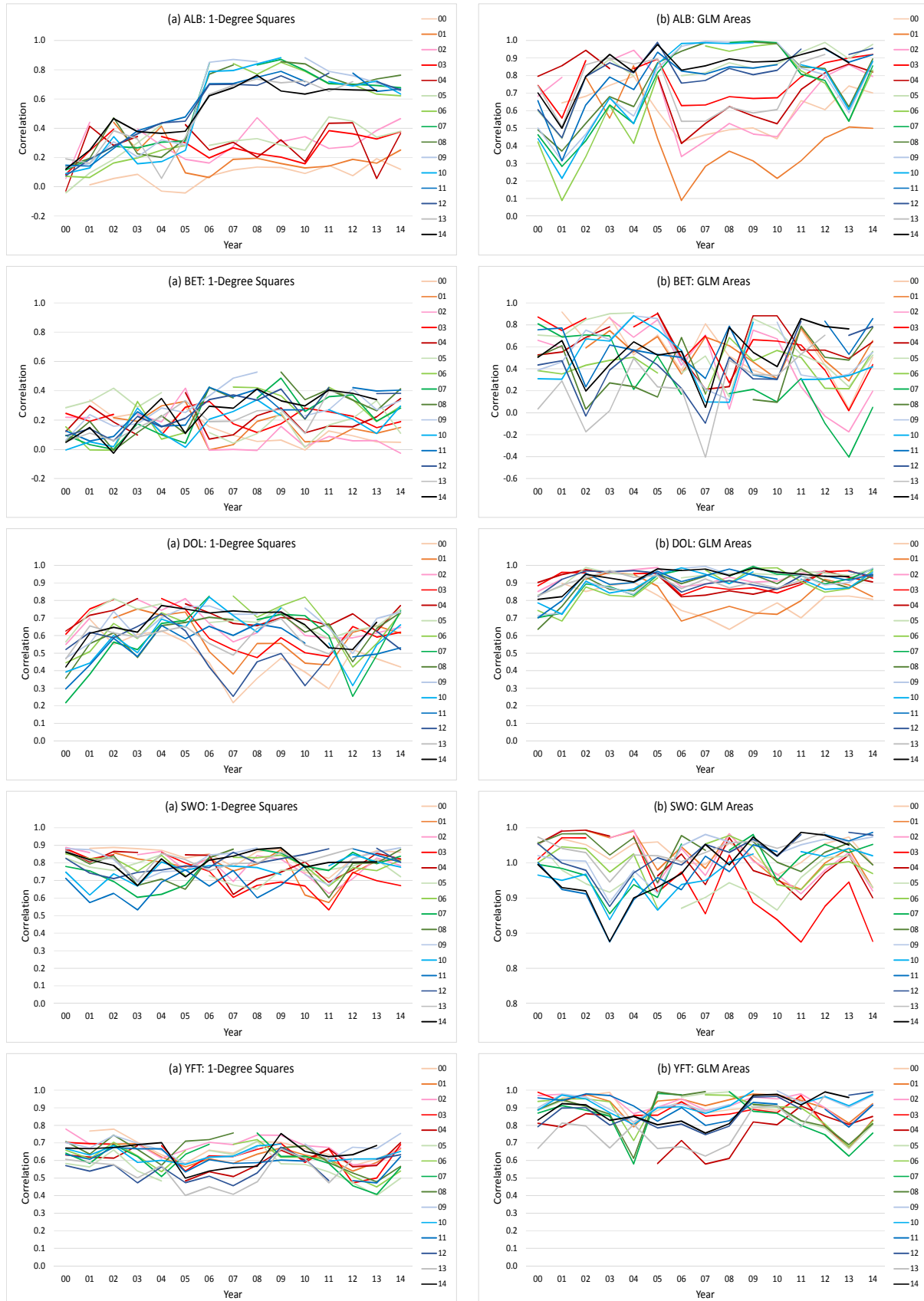


Figure 6.35: Annual distributions of nominal ALB CPUE for the simulation ETBF data-set. The colours indicate the following range of CPU: (i) yellow, <2.0, (ii) orange, <4.0, (iii) green, <6.0, (iv) light blue, <8.0, (v) dark blue, <10.0, (vi) red, <12.0, and (vii) purple, >12.

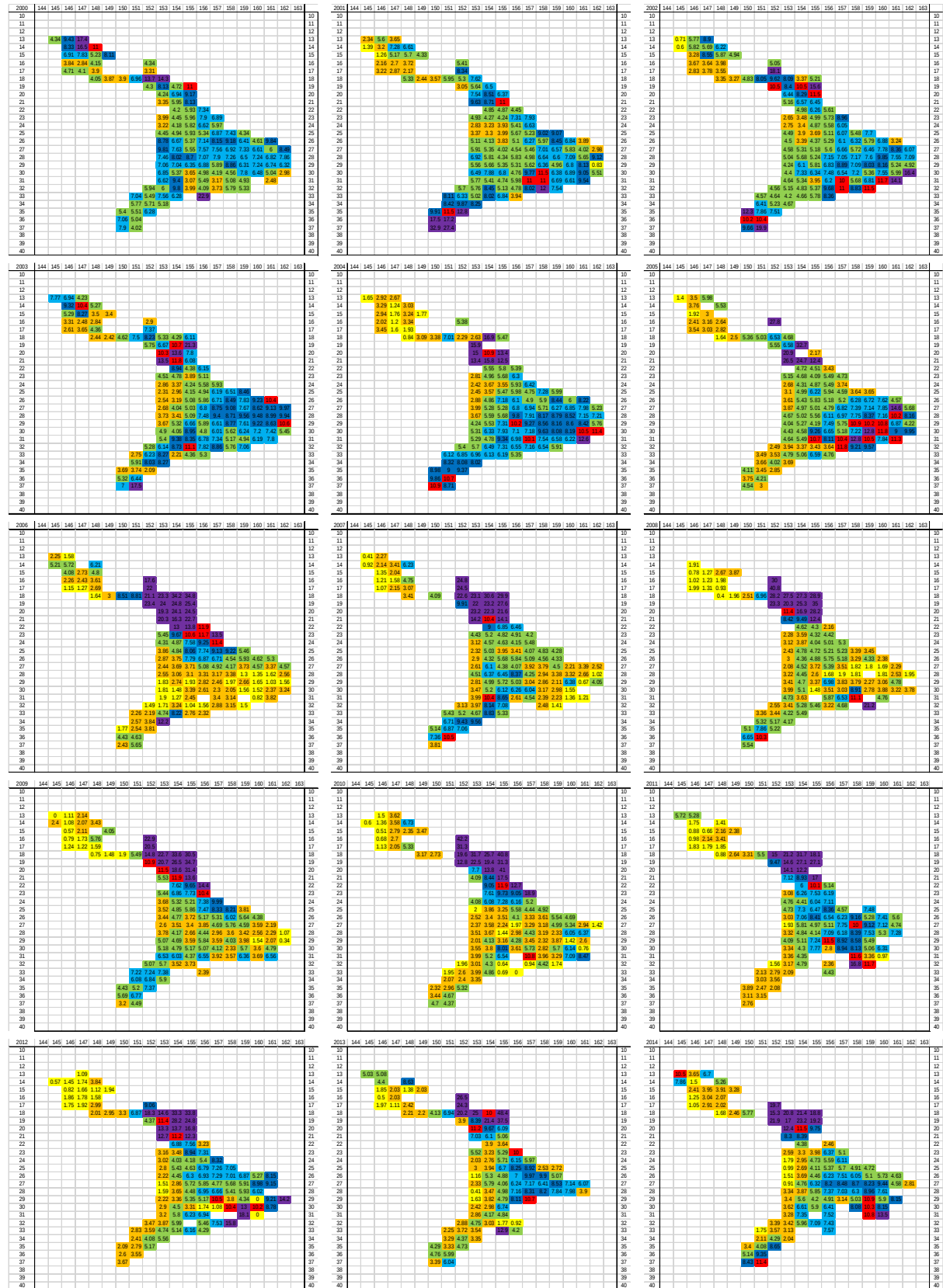
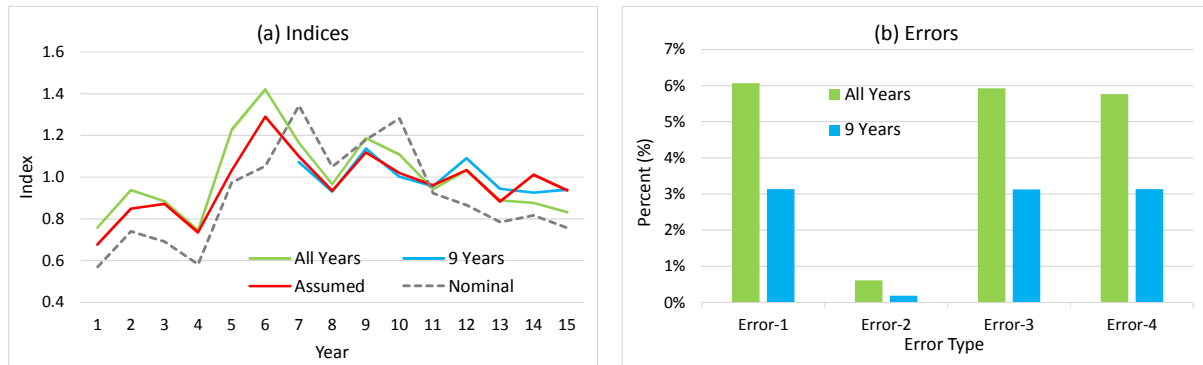


Table 6.7: Species Ranked Scores (SRS) for each model for the D9-Trial analyses. Models having equal rank scores are shaded.

Rank	1	2	3	4	5	6	7	8	9	10
SRS	4	8	14	14	20	25	27	32	36	40
Model	A-3	A-1	A-2	A-4	A-3a	A-4a	A-1a	A-5	A-2a	A-5a

Figure 6.36: Comparison of (a) the standardised CPUE indices and (b) the index errors based on the A-1a model for all years and the A-3 model for the last 9 years. The nominal CPUE and the assumed abundance indices are also shown.



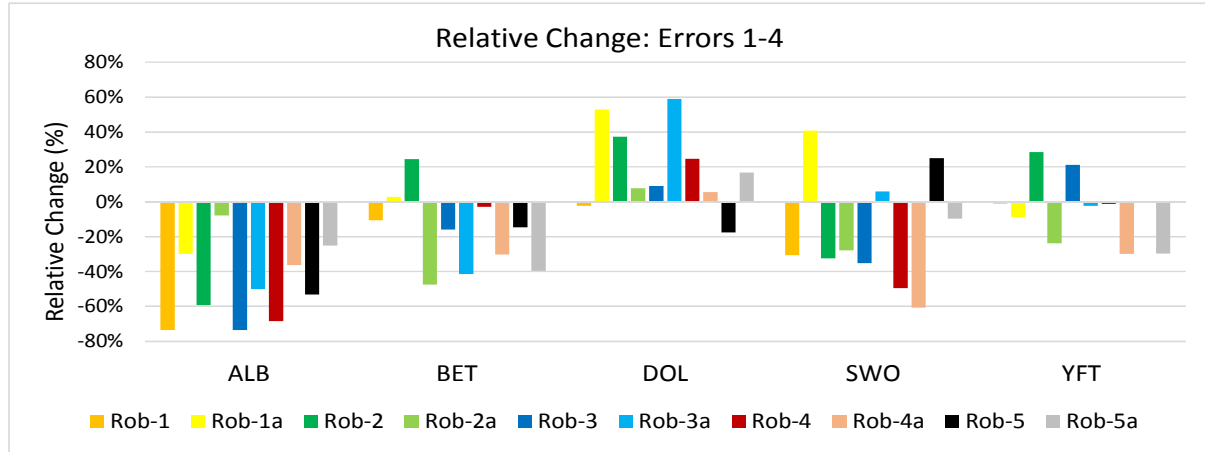
The pattern of annual correlations for the other four species, shown in Figure 6.34, display differing levels of inter-annual variability (e.g. high for BET and low for DOL), but do not display the substantive shifts noted in the results for ALB. This suggests that the distinctive changes noted in the distribution of ALB CPUE may be having an impact on the performance of the models fitted to this species.

To investigate this possibility, a further set of analyses (known as the D9-Trial) were undertaken similar to the 1D-Trial, but limited to the A-Group of models fitted to the ALB data for the last nine years of data only (i.e. it excludes the data before the substantive changes noted in the spatial correlations for ALB). The errors and species-ranked-scores were calculated as before for each model and the results are shown in Table 6.7. Unlike the previous 1D-Trial result for ALB, where the models without the gear effects out-performed the models including gear effects and the A-1a model was ranked first, when applied to the data for the last 9 years the models with gear effects perform best, and the A-3 model is ranked first. This last result is consistent with the previous results for the 1R-Trial and 2R-Trial analyses.

A further comparison of the D9-Trial and 1D-Trial results is shown in Figure 6.36. First, the annual abundance index for each Trial is displayed and compared with both the nominal index and the assumed annual index. All indices have been scaled such that the average of the last 9 years is equal to 1. In general, the D9 index better tracks the true index, and corrects the negative trend over the last 9 years seen in the 1D-Trial index. Second, a comparison of the Mean Annual Error (MAE) for the four errors calculated over the last 9 years, but with each index retaining its original scale, is shown in Figure 6.36b. Again, a substantive improvement in each of the four errors is seen, with around a 70% reduction in the Type-2 error and a 50% reduction in the other errors.

In order to investigate whether the shift noted in the fishing patterns for Albacore Tuna may have influenced the model fits to the other species, the above analysis of fitting the Group-A

Figure 6.37: Comparison of the overall model rank for the five Group-A models based in the D9-Trial analyses with and without the two gear effects included in the linear predictor.



models to the last nine years of data only was also undertaken for the other four species. The following average relative change in the MAE for each error i was calculated for each model m and species s :

$$\Delta MAE_{m,s} = \frac{1}{4} \sum_{i=1}^4 \frac{(MAE_{m,s,i,9y} - MAE_{m,s,i,all})}{MAE_{m,s,i,all}}$$

where the subscripts *all* and *9y* refer to the analyses over all years and the last 9 years respectively. The results are shown in Figure 6.37, where positive and negative changes indicate a larger and smaller MAE respectively (by the percentage shown) for the analysis over the last 9 years compared to all years. Consistent with the previous results above for ALB, the ΔMAE is negative for all models, indicating a generally substantial improvement in the estimation of the abundance index. However, this result is not consistent across all species, with some improvement in the estimation of the index over the last 9 years seen for both BET and SWO, little if any change seen for YFT, and a poorer estimation of the index seen for DOL. The reasons for these differences remain unclear, but are a further indication that there are factors specific to each species that influence the fit of any model to the data and as such the estimation of the annual abundance index.

6.5.3 Construction of GLM Area effects

For all models fitted to the ETBF simulated data-sets the *Area* effects correspond to the seven GLM-Areas shown in Figure 5.23. As previously noted, these *Areas* effects were determined externally from the model by visual inspection of the distributions of nominal CPUE across the 1-degree cells in the data. However, a disadvantage of this approach is that the construction of the *Area* effects has to be undertaken manually (or empirically) and separately for each species. This can be time-consuming if there are a large number of analyses (or species) to be completed. Furthermore, there is no guarantee that this rather *ad hoc* process, or the number of *Area* effects selected, is optimal. It would be preferable if the selection of *Area* effects could be based on some algorithm which can be incorporated into the analysing model. In this Section we compare the performance of the A-1 model using a suite of different *Area* effects while keeping all other model inputs constant.

The following six criteria were used for selecting the *Area* effects used in the model fitted to the data for each of the five species:

- Type-1 These are the seven GLM-Areas shown in Figure 5.23, and used in all previous analyses –where the number of *Area* effects for each species is seven;
- Type -2 These are based on the same empirical approach as Type-1 except the number of GLM-Areas for each species was increased to be between 9 and 11;
- Type -3 These are based on the same empirical approach as Type-1 except the number of GLM-Areas for each species was increased to be between 13 and 14;
- Type 4 The Type-1 GLM- Areas for YFT were used for all species.
- Type -5 The total area of the fishery was divided into seven GLM-Areas based on the following latitudinal strips <18S, <22S, <25S, <28, <31S, <34S, and ≥34S;
- Type 6 For each data-set seven GLM-Areas were identified using the algorithm used for analysis of the DIST data-sets and previously described in Section 5.7.2. First, for each year the nominal CPUE was calculated across all one-degree cells in the simulated data-set. Second, for each cell the mean nominal CPUE across all years was calculated. Third, the distribution of nominal CPUE across all cells was ranked from highest (rank=1) to lowest (rank=139). Each cell was then allocated to one of seven GLM-Areas using the following formula:

$$Area = 1 + floor \left[\frac{rank_i - 1}{20} \right]$$

where $rank_i$ is the ranking of the nominal CPUE of cell i . as determined in the third step above. This approach makes each GLM-Area a composite of 20 individual cells (so each Area is the same size), though the approach does not guarantee that each Area is composed of 20 cells that are contiguous.

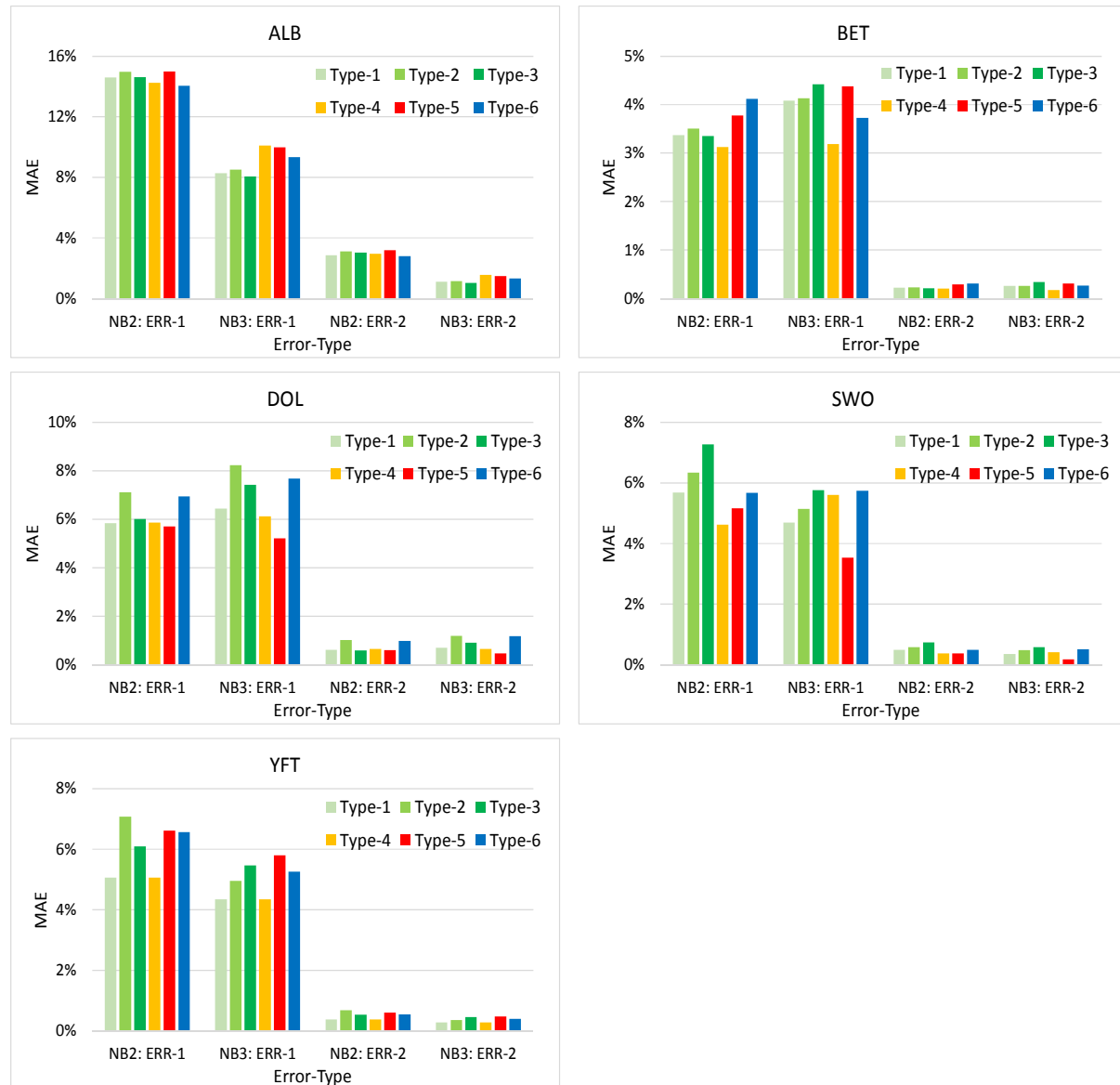
The A-1 model was fitted to the 100 simulated ETBF data-sets using the random R-catches and the Mean Annual Errors (MAE) calculated as described previously. The model was also fitted using either the two 2-way interactions between the *Year*, *Quarter* and *Area* effects (the NEG-2 model) or the single 3-way interaction (the NEG-3 model). For each species and error type, the performance of the model was ranked across the six types of *Area* effects (1=Best, 6=Worst) and the mean rank was then taken across the four model-error combinations. For each species, and for each type of *Area* effects, the Type-1 and Type-2 MAEs for each model are shown in Figure 6.38, while the mean ranks are shown in Table 6.8. Except for ALB, the size of the errors for the NEG-2 and NEG-3 models were similar for each species, indicating that there was little benefit in fitting the more parameter intense 3-way interaction.

The results for the three species-specific empirical *Area* effects (Types 1 to 3) indicate that varying the number of areas from 7 to 14 does not always improve the model fit, with the Type 1 model with only 7 Areas effects performing best for all species. On the other hand, the relative performance of the Type 2 and Type 3 models was mixed across the five species. Of these three *Area* effects, Type-1 is found to have the best overall fit (mean rank of 2.45, c.f. Table 6.8) and this result supports the initial decision to limit the number of GLM-Area effects to seven. Interestingly, use of the Type-4 *Area* effects (equivalent to the Type-1 GLM-Areas for YFT) has a better performance for BET than using the species-specific Type-1 *Area* effects, and are slightly better for SWO. However, the Type-4 model performs poorly for ALB. This implies that identification of appropriate *Area* effects is not a simple task and most likely needs to take account of temporal variations in the distributions of CPUE across the years as observed for ALB.

Table 6.8: Mean rank (1=Best, 6=Worst) of the performance of the A-1 model fitted to the 100 simulated ETBF data-sets using the R-catches and the six types of Area effects described in the text.

Area-Effect	Construction	ALB	BET	DOL	SWO	YFT	Mean
Type-1	Empirical	2.25	2.875	2.75	2.875	1.5	2.45
Type-2	Empirical	4	3.625	6	4.25	4.5	4.475
Type-3	Empirical	2.5	4	3.25	6	4	3.95
Type-4	YFT Area-1	4.25	1	2.75	2.5	1.5	2.4
Type-5	Strips of Longitude	5.5	5	1.25	1.25	5.5	3.7
Type-6	Ranked 1-degree CPUE	2.5	4.5	5	4.125	4	4.025

Figure 6.38: The Mean Annual Error (MAE) for the A-1 model fitted to the 100 simulated ETBF data-sets using the R-catches and the six types of Area effects described in the text. Note: NB-2 refers to the NEG-2 model, NB-3 refers to the NEG-3 model while ERR-1 refers to Type-1 Error and ERR_2 refers to the Type-2 Error.



Of the two algorithmic-based *Area* effects, the simple latitudinal stratification of the fishery (Type-5) performs best for both DOL and SWO, while ranking the nominal 1-degree cells (Type-6) outperforms the Type-5 model for the other three species. While the Type-1 models out-perform the Type-6 models for all species, the Type-5 models out-performs the Type-1 models for both DOL and SWO, but performs poorly for the other three species relative to the respective Type-1 model results. This result may be influenced by any specific cline in the north-south distribution for a given species.

These results indicate that, of the six approaches tested, there is no specific (or simple) approach to specifying the *Area* effects that performs best across all species. Analysts are therefore encouraged to investigate various approaches for identifying appropriate *Area* effects where necessary, especially where the data indicates that there are substantial shifts in the spatio-temporal distributions of CPUE over time (as observed in the ALB data). Auxiliary information, such as that which may help identify differences in preferred habitat types or demarcate management zones, is likely to be useful in this regard.

6.6 Implications for the ETBF Harvest Strategy

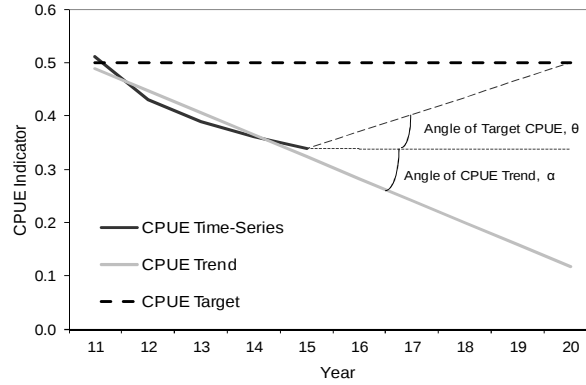
The management framework for the ETBF is predicated on setting an annual Total Allowable Commercial Catch (TACC) for each of the five principal target species. The TACC for Broadbill Swordfish and Striped Marlin is informed by the Recommended Biological Commercial Catch (RBCC) as determined by the harvest strategy adopted for this fishery (Campbell *et al.* 2007; Davies *et al.* 2008).

The harvest strategy is based on a decision-tree which utilises a number of empirical based indicators - the standardised CPUE of small, prime and large-sized fish (where small, prime and large fish are those within the lower 25th, middle 50th and upper-25th -percentile of the total size distribution of fish retained and landed in the fishery, respectively) and the proportion of these size classes in the total catch. The primary control rule for determining the RBCC in any year involves using the formula:

$$RBCC_{t+1} = TACC_t * (1 + \beta * S_{TARG}) \quad (6.6.1)$$

where S_{TARG} is the slope based on the angle subtended by the standardised prime-sized CPUE trend line and the line joining the present CPUE value and a target CPUE value a specified number of years, N_{TARGET} , in the future. Such a situation is shown in Figure 6.39 where it is assumed $N_{TARGET}=5$ years. The control parameter β is referred to as the feedback gain factor and defines how responsive the RBCC change is to changes in CPUE. MSE testing of the ETBF harvest strategy (Kolody *et al.* 2010) noted that the parameters β and N_{TARGET} were confounded and recommended a value of $\beta = 1$ for the ETBF harvest strategy. Furthermore, in order to smooth out the high inter-annual variability seen in the CPUE indices for some species the standardised CPUE indices are first smoothed using a LOWESS (locally-weighted scatterplot smoother) algorithm (Cleveland 1979; Cleveland and Devlin 1988). This is done to ascertain the underlying trend (the signal) in the abundance / availability of a species to the fishery (using the CPUE time-series as a proxy), but which can be hidden to some extent by shorter-term trends (noise) due to the imprecision in which the CPUE is measured and standardised and is a true index of underlying abundance.

Figure 6.39: Conceptual example of how the slope-to-target parameter used in the primary control rule is derived.



If an index is available that accurately tracks the annual size of the resource being managed then a measure of the error in the RBCC based on using the standardised CPUE can be calculated. For example, if S_{TRUE} is the slope-to-target of the true index over the last 5 years then the difference in the RBCC based on this true index and the RBCC based on the standardised CPUE is as follows:

$$\begin{aligned}\Delta RBCC &= RBCC_{CPUE} - RBCC_{TRUE} \\ &= TACC * (1 + S_{CPUE}) - TACC * (1 + S_{TRUE}) \\ &= TACC * (S_{CPUE} - S_{TRUE})\end{aligned}$$

This can be stated as a percentage of the true RBCC:

$$\frac{\Delta RBCC}{RBCC_{TRUE}} = \frac{S_{CPUE} - S_{TRUE}}{(1 + S_{TRUE})} * 100 \quad (6.6.2)$$

As noted in Figure 6.39, the value of S_{CPUE} (and S_{TRUE}) depend on both the angle subtended by the standardised CPUE (the true resource abundance) trend line and the line joining the present CPUE value and a target CPUE. If $A = \tan(\alpha)$ is the slope of the linear regression of CPUE over the past 5 years, and $B = \tan(\theta)$ is the slope to the target CPUE, then $S_{CPUE} = \tan(\alpha + \theta)$ and after accounting for the different configurations of A and B it can be shown that:

$$S_{TARG} = \begin{cases} \tan[\tan^{-1}(A) - \tan^{-1}(B)] & A > 0, B \geq 0, A \geq B \\ -\tan[\tan^{-1}(B) - \tan^{-1}(A)] & A > 0, B \geq 0, A < B \\ \tan[\tan^{-1}(A) + \tan^{-1}(-B)] & A > 0, B < 0 \\ -\tan[\tan^{-1}(-A) + \tan^{-1}(B)] & A < 0, B \geq 0 \\ \tan[\tan^{-1}(-B) - \tan^{-1}(-A)] & A \leq 0, B < 0, -A < -B \\ -\tan[\tan^{-1}(-A) - \tan^{-1}(-B)] & A \leq 0, B < 0, -A \geq -B \end{cases}$$

In order to simplify the calculation of both S_{CPUE} and S_{TRUE} in the following analysis it is assumed that that $B=0$ (i.e. $\theta=0$) so that $S_{TARG}=A=\tan(\alpha)$ is simply the slope of the standardised CPUE (true abundance) trend line. Also, the LOWESS smoothing of the standardised CPUE index is not used.

Based on the assumed abundance trend for each species used to generate the simulated ETBF data (c.f. Figure 5.21), and the annual standardised CPUE indices based on each of the models used to analyse these data, the absolute value of the percentage error in the RBCC given by

Figure 6.40: For the four simulation trials undertaken based on the ETBF data, comparison of the Mean Annual Error in the RBCC by species and model.

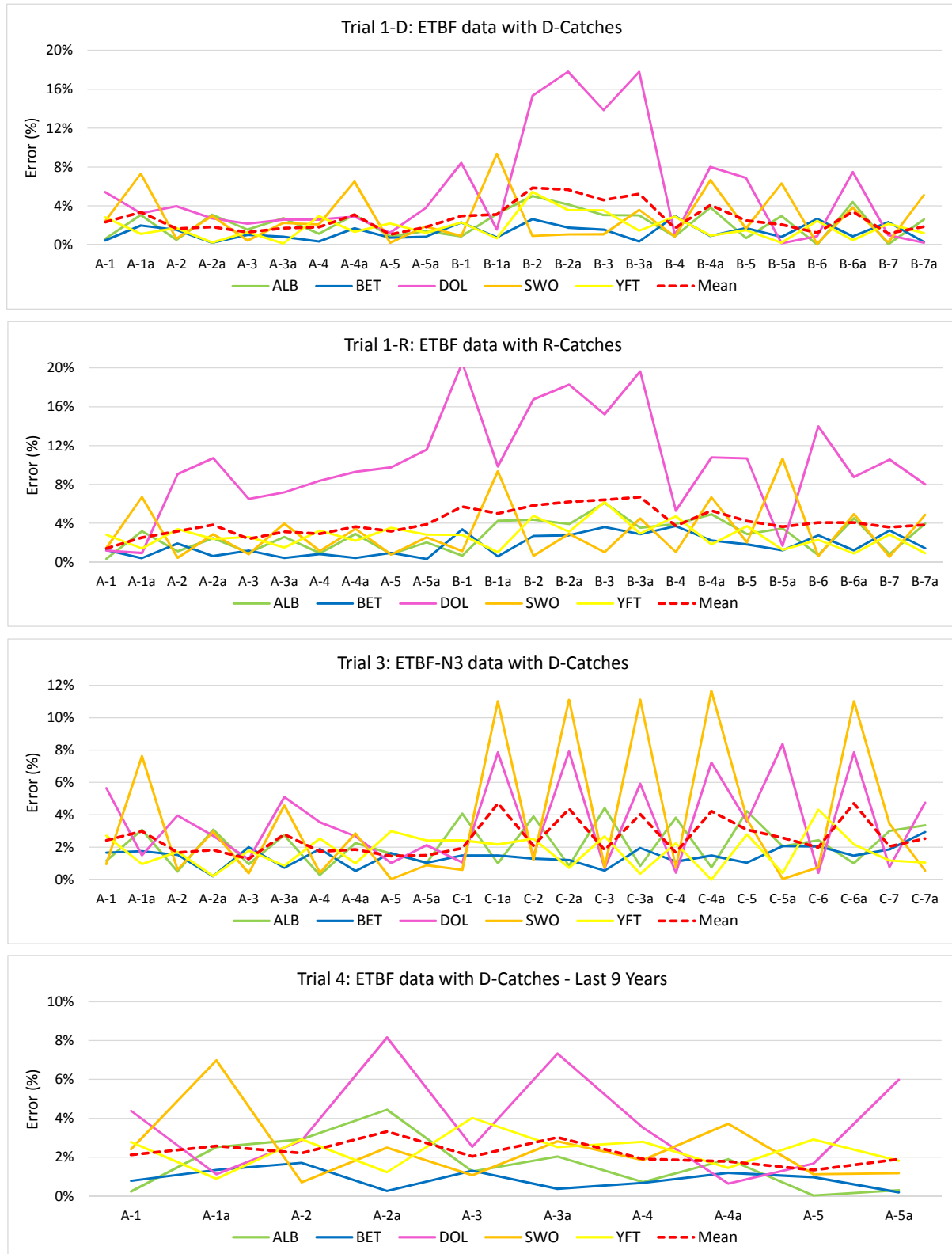
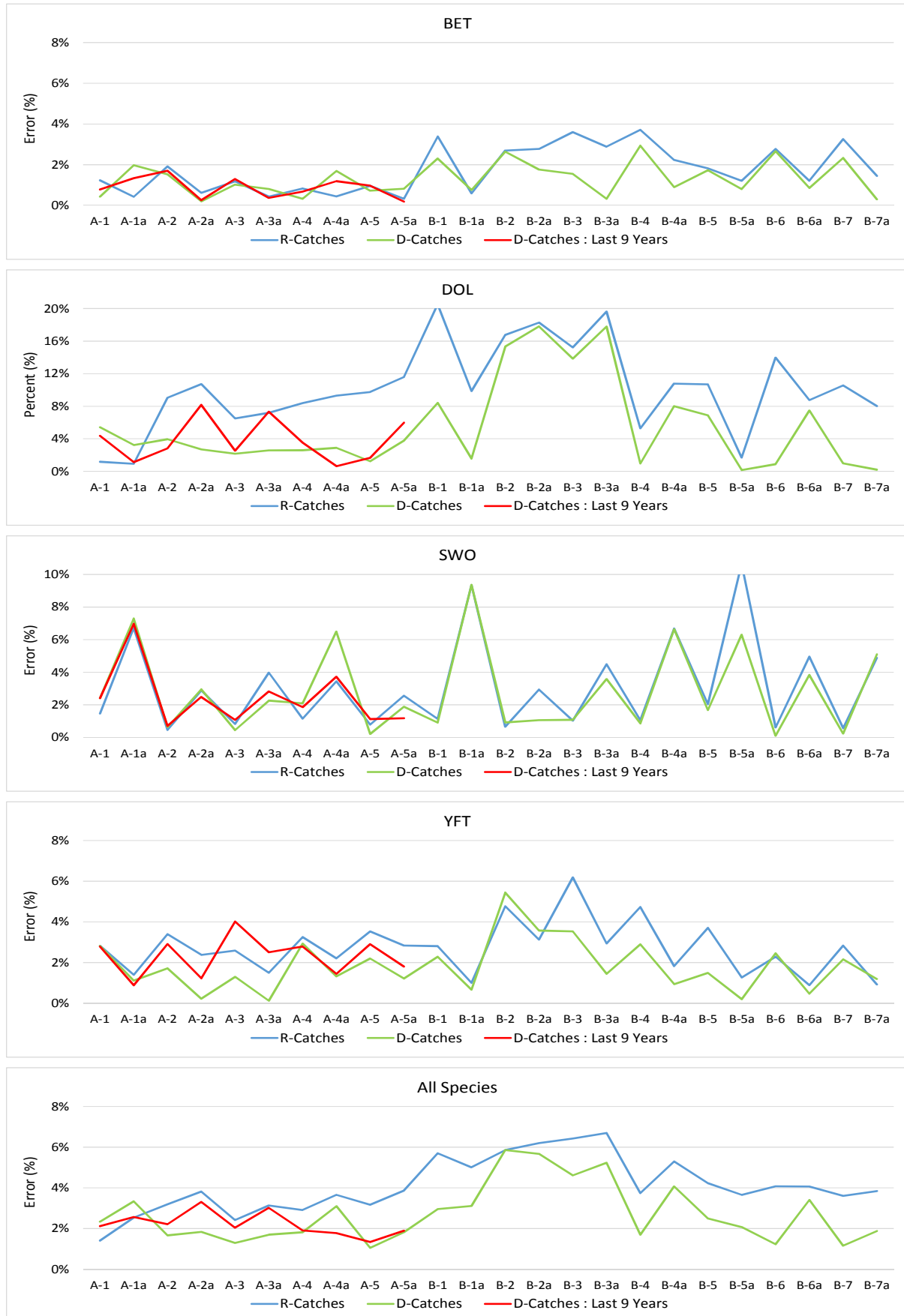


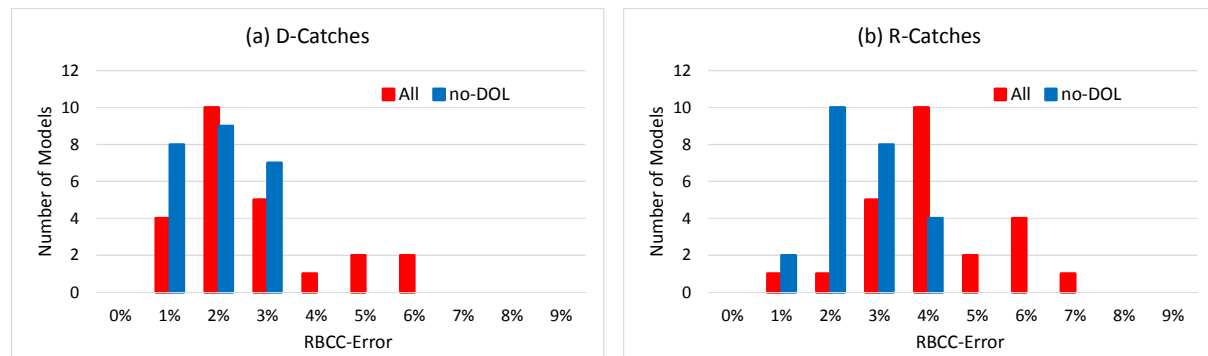
Figure 6.41: For each species and model, comparison of the Mean Annual Error in the RBCC across the three simulation trials undertaken based on the ETBF data.



Eqn. (6.6.2) above was calculated for each of the trials previously described. For the D-catch analyses the results are based on the single standardised CPUE index generated, while for the R-based analyses the results were the mean error over all 100 simulations. As with the four previous Index-based errors, the Mean Annual Error (MAE) based on the RBCC-Error was calculated for each species and model and the results are shown in Figures 6.40 and 6.41.

Several features are note-worthy. First, as with the previous Index-based errors, the size of the error in the RBCC varies across the different species and models such that no single model minimises this error. Second, apart from the results for DOL the mean error is generally less than 4% (and often less than 2%) and is similar for a number of models (c.f. Figure 6.42). Third, again the models not including the gear effects (the 'a' models) generally perform less well than those including the gear effects. This is particularly noticeable for the Group-C models for DOL and YFT, though the 'a' models in this Group generally perform better for YFT and ALB. The reasons for the higher errors observed for DOL remains unknown and should be investigated further. Fourth, across the three model Groups, the Group-A models perform best while the two PCA-based models (B-2 and B-3) perform worst. Fifth, across all species the mean error is generally higher for the analyses based on the R-catches compared to the analyses based on the D-catches, though this difference is largely due to the variation seen in the DOL results (and to a smaller degree in the YFT results). The analyses based on the last 9-years of data also generally align with the results based on the analyses across all years (c.f. Figure 6.41).

Figure 6.42: Distribution of the mean RBCC-Error (rounded to the nearest percent) across the 24 models for the two 1D-Trial and 1R-Trial analyses. Results are shown for both the mean across all five species and across the four species excluding DOL.

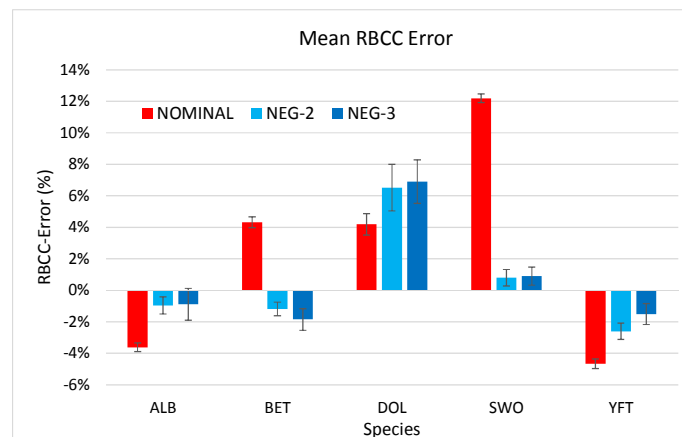


Finally, the Model Ranked Score (*MRS*) was tabulated for each model using the approach adopted previously and the over-all rank of each model from best-to-worst performing is shown in Table 6.9 where results are shown for the two 1-Trial analyses. For the D-catch analyses, the A-5 model was the best performer followed by a group of six models each having a similar *MRS*. Given that these results are based on only a single analysis, and noting that the slope of the standardised index can vary based on small differences in the index in a single year (which would be minimised to some extent had the LOWESS smoother been used), it remains uncertain as to how robust this result would be for other simulations using the D-catch approach. On the other hand, for the R-catch analyses based on 100 random-based simulations, the two models A-3 and A-1 were the best performers (*MRS* of 38 and 39 respectively), followed by their 'a' model counter-parts with a slightly higher *MRS* (43 and 44 respectively). The A-3 model was also the best-performing model for estimating the abundance index based on the R-catches.

Table 6.9: Over-all ranking and Model Rank Score (MRS) for each model based on ranking the mean RBCC-Error across all species for the two 1-Trial analyses. Best=1, Worst=24.

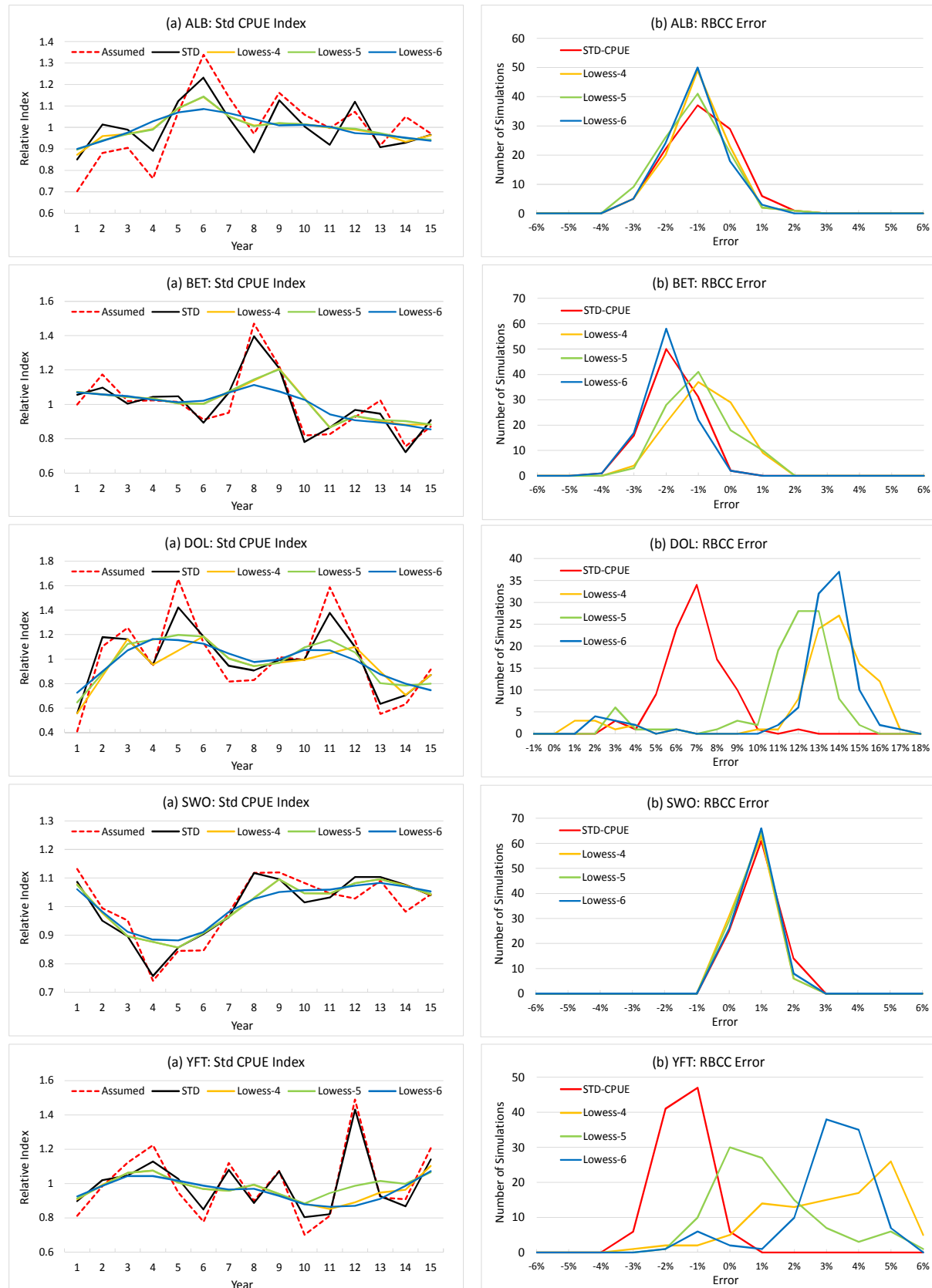
	RANK	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
D-Catch	MRS	38	43	45	46	46	46	46	50	51	55	57	59	60	63	63	70	71	74	78	79	80	84	96	100
	Model	A-5	B-7a	B-5a	A-3	A-3a	B-6	B-7	A-2a	A-2	A-5a	A-4	A-1	B-4	B-1a	B-5	B-3a	B-1	B-6a	A-4a	A-1a	B-4a	B-3	B-2a	B-2
R-Catch	MRS	38	39	43	44	49	50	52	53	53	55	56	56	56	60	63	64	69	74	76	84	86	89	93	98
	Model	A-3	A-1	A-3a	A-1a	A-4	A-5	A-4a	A-2	B-6	B-7	A-2a	A-5a	B-5a	B-7a	B-6a	B-1a	B-1	B-5	B-4	B-4a	B-2	B-2a	B-3a	B-3

While the above results give estimates of the size of the errors in the RBCC calculations using the different models, in order to understand the direction of the error, the nominal values (i.e. non absolute values) of the RBCC-error need to be investigated. For this exercise, we used the best-performing A-3 model identified above, and calculated the mean and standard deviation of the distribution of the nominal RBCC-error values across the 100 simulations conducted using the R-catches. The results for each species are shown in Figure 6.43. Results are also shown when the slope-to-target trend is based on using either the nominal CPUE or the standardised CPUE based on a GLM incorporating either the two 2-way interactions (NEG-2) or a single 3-way interaction (NEG-3) between the *Year*, *Quarter* and *Area* effects. For both ALB, BET and YFT the mean errors (based on the NEG-2 and NEG-3 models) are negative, indicating that the RBCC is conservative in that it under-estimates the true RBCC, while for the other two species the mean errors are positive and the RBCC is over-estimated. Apart from the results for DOL, the RBCCs based on the standardised CPUE indices are substantially more accurate than that based on the nominal CPUE. This is particularly noted for SWO where the error is reduced from around 12% to less than 1%. The standard deviations of the related distributions are also relatively small indicating that for four of the simulated species the estimated RBCCs are most likely (i.e. with 95% confidence) within 3% of the true RBCC. Finally, the RBCC-errors based on the NEG-2 or NEG-3 models are seen to be similar, with no particular model consistently performing better than the other.

Figure 6.43: For each species, comparison of the mean and standard deviation of the distribution of the nominal values of the RBCC-error across the 100 simulations.


Finally, in order to investigate the utility of the LOWESS smoother on the calculation of the RBCC, the previous analysis was repeated but the standardised CPUE index was smoothed before calculation of the corresponding trend over the past 5 years. A user-specified input to the procedure called the "*bandwidth*" or "*smoothing parameter, S*" determines what proportion of the data is used to fit each local polynomial incorporated into the smoothing function. Large values of *S* produce the smoothest functions that wiggle the least in response to fluctuations in

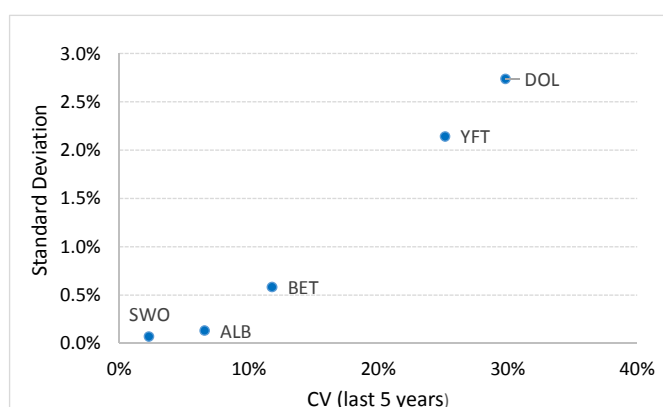
Figure 6.44: For each species (a) example comparison of assumed abundance index, standardised CPUE index and three LOWESS smoothed indices for a single analysis and (b) the distribution of the errors in the RBCCs calculated using either the standardised CPUE or the LOWESS smoothed indices across 100 simulations.



the data. The smaller S is the closer the regression function will conform to the data. Using too small a value of the smoothing parameter is not desirable, however, since the regression function will eventually start to capture the random error in the data. Useful values of the smoothing parameter typically lie in the range 0.25 to 0.5 for most LOWESS applications. In the following we used values of S of 0.27, 0.33 and 0.40, being equivalent to using 4, 5 and 6 of the 15 annual data points.

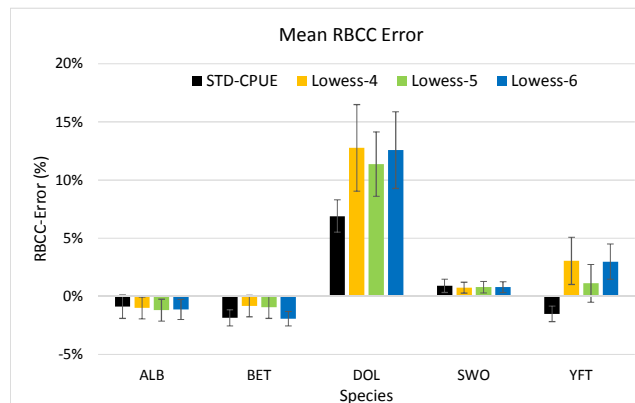
For each species, comparison of the assumed annual abundance index, the standardised CPUE index calculated using the A-3 NEG-3 model, and the corresponding three LOWESS smoothed indices for a single simulation analysis is shown in Figure 6.44a. Variations in the trajectories of the LOWESS smoothed indices can be seen, based on both the value of the smoothing parameter used, and the inherent inter-annual variability in the standardised CPUE being smoothed. The corresponding distribution of the error in the RBCC based on use of each index in the primary control rule across all the 100 simulated data-sets analysed is shown in Figure 6.44b. The four distributions shown for SWO are very similar and this is likely linked with the fact that the annual standardised CPUE for this species is relatively stable over the last 5 years over which the CPUE trend is calculated. A high degree of similarity between the four error distributions is also seen for ALB. On the other hand, there is a large divergence seen in the results for DOL and YFT. Again, the annual standardised CPUE for these two species are quite variable over the last 5 years. Using the CV of the values of the annual standardised CPUE over the last 5 years as a measure of the level of 'smoothness' or variability in the index, the values of ALB, BET, DOL, SWO and YFT are 6.6%, 11.8%, 19.8%, 2.3% and 25.2% respectively. Indeed, there appears to be a relationship between the CV of the last 5 data points of the standardised CPUE and the standard deviation of the mean of the four RBCC-errors calculated above (c.f. Figure 6.45). This implies that the higher the CV of the last 5 data points in the standardised CPUE, there is likely to be a commensurate higher variability in the calculated RBCC given the use of different smoothing parameters used with the LOWESS. Indeed, some care should be taken in selecting which parameter to use and the criteria for identifying the most appropriate parameter requires further investigation.

Figure 6.45: For each species plot of the standard deviation of the mean of the four RBCC-errors shown in Figure 6.44a versus the CV of the last 5 data points of the standardised CPUE.



Finally, the comparison of the mean and standard-deviation of the distributions of the four RBCC-errors across the 100 simulations discussed above is shown in Figure 6.46. The errors for the LOWESS-based RBCCs for DOL are substantially greater than those associated with the RBCC based on the non-smoothed CPUE. As discussed above, this is likely to be due to the high CV in the CPUE time-series. On the other hand, the errors for two of the LOWESS-

Figure 6.46: Comparison of the mean and standard-deviation of the distribution of the RBCC-error across the 100 simulations discussed in the text.



based RBCCs for BET are around half that associated with the RBCC based on the non-smoothed CPUE. The error for the LOWESS-based index is also smaller for at least one of the indices used for SWO and YFT, though the smoothing parameter associated with the best performing index is not always the same. While the variation in the errors between the results based on the different smoothing parameters is usually small, again some care should be taken in selecting the value of this parameter if the CPUE index is to be smoothed before being used in the harvest strategy.

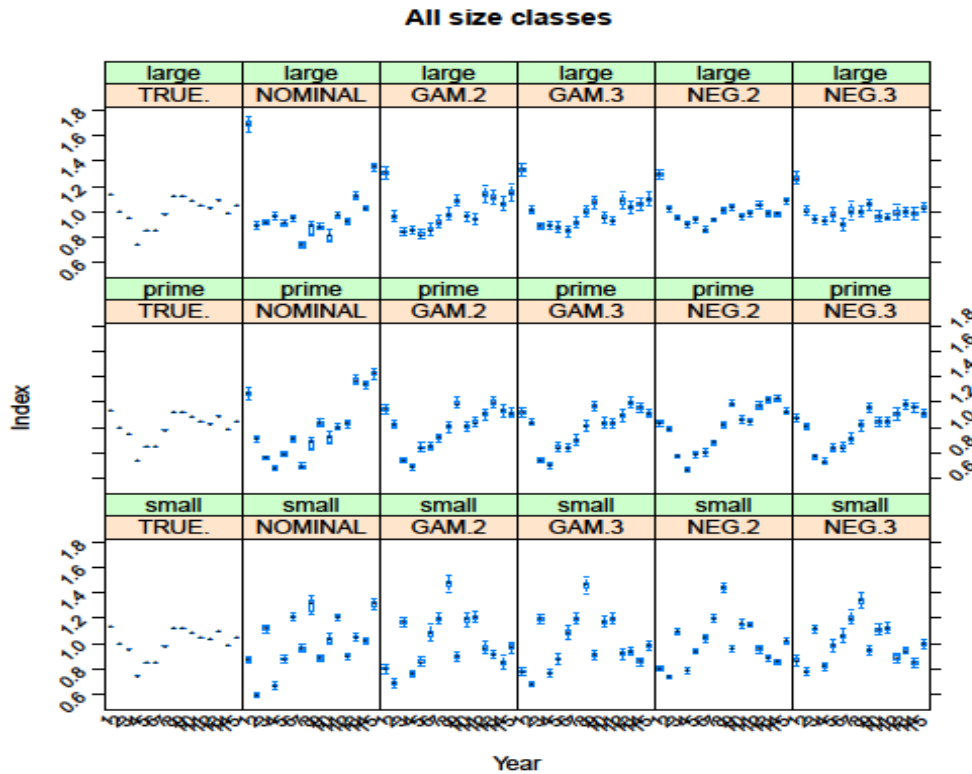
6.7 Further Implications for the ETBF Harvest Strategy in a Stock Assessment Context

This final Section focusses on two important, but essentially flow-on issues of the multispecies standardisation work: (i) how do alternate series affect the running of the current harvest strategy (HS) in the ETBF; and (ii) what are the stock assessment implications of producing overtly multispecies CPUE series. The first issue is obviously directly linked to the actual management implications of alternate methods of standardising CPUE, given these indices drive the current HS. The second issue is both a more subtle and wide-ranging one. Current single-species assessments, by construction and assumption, treat the series independently with the only commonality the effort series. However, the series themselves may well contain correlation structures in both observation and process errors that would have to be dealt with if one wanted to explore a truly more multispecies assessment approach - especially in terms of concepts like multispecies 'optimal' effort and yield (be it MSY or MEY in nature).

6.7.1 Harvest Strategy Implications

The current ETBF harvest strategy is only implemented for the two billfish species: Swordfish and Striped Marlin. The main reason for this is an apparent lack of feedback for the two tropical tuna species Bigeye Tuna and Yellowfin Tuna, stemming from the potentially minimal impact of the current ETBF catches on those stocks (Hillary *et al.* 2016). The four GLM approaches (GAM-2, GAM-3, NEG-2, NEG-3) outlined in Section 5.6 were run for the Swordfish example (100 simulations for 15 years of data) and it is on these simulated series that we focus the HS implications work on. There are six series in total: the 'true' (simulated) series; the nominal (unstandardised) series, two types of Gamma GLM-derived series; and finally two types of Negative-Binomial GLM-derived series. Figure 6.47 summarises all six series in a boxplot for all the three size classes (small, prime, and large).

Figure 6.47: Boxplot summary of the Swordfish CPUE for all size classes (small, prime, large) and for each of the six series (true, nominal, the two Gamma and two Negative-Binomial models).



The ETBF HS adjusts the most recent agreed TACC by a certain amount each year, driven by a hierarchical decision tree rule (Kolody *et al.* 2010). The primary level is the trend in prime CPUE, with subsequent levels in the tree making adjustments given inferred trends in recruitment and likely spawners from the CPUE of the other size classes small and large. A recent development was the use of a LOWESS smoother applied to the series for each size class to reduce the often quite high levels of variability in the indices over time.

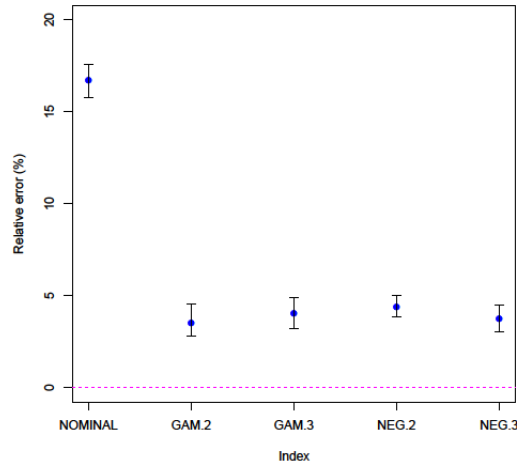
To analyse the impact of the suite of alternative standardisation models we consider two main summary statistics:

1. Relative error: what are the bias properties (if any) relative to the true final-year predicted TACC for each of the other four indices
2. Average annual variation (AAV): for a quasi-retrospective analysis (removing increasing numbers of years and calculating resultant TACCs) what are the levels of percent-age variation from year to year for all the indices (including the true index)

In lieu of a full reworking of the previous MSE work (Hillary *et al.* 2016; Kolody *et al.* 2010), with the CPUE simulation and standardisation models embedded within augmented Operating Models (OMs), it is difficult to assess the wider impact of alternative CPUE series. The two simple approaches we explore here can, however, quickly indicate whether key features (i.e. are we tracking the 'true' dynamics in terms of TACCs and the inter-annual variability in TACCs?) are being unduly affected by the alternative CPUE approaches.

The relative error statistic is a very simple but useful bias indication. We have the true TACC predicted using the full 15 years of CPUE data, for each of the 100 simulations. We also have the TACCs as predicted by the HS when using the nominal and four candidate standardised

Figure 6.48: Relative errors (median and 80% CI) in the TACC predicted using the whole data series (15 years) for each of the 100 simulations. The percentage error is expressed relative to the true TACC for each of the other candidate series.



indices: $TACC_y^i$, where i denotes the TACC arising from using the nominal, GAM or NB-GLM indices as inputs to the HS. The relative error, Δ_y^i , is simply given by the following:

$$\Delta_y^i = \left(1 - \frac{TACC_y^{true}}{TACC_y^i}\right) \times 100 \quad (6.71)$$

with positive/negative values indicative of the type of bias (i.e. over/under estimation, respectively).

Figure 6.48 summarises the relative error properties of the four series and there are a number of inferences that can be made from it. Focussing on the nominal series *versus* all the other standardised series, it is notably more biased than all four standardised series, with a positive bias (over estimation) of *ca.* 16-17%. For the four standardised series they all display very similar relative error properties: small positive bias for all (*ca.* 3-5%) and with no obvious bias trend across standardisation method (GAM or NEG).

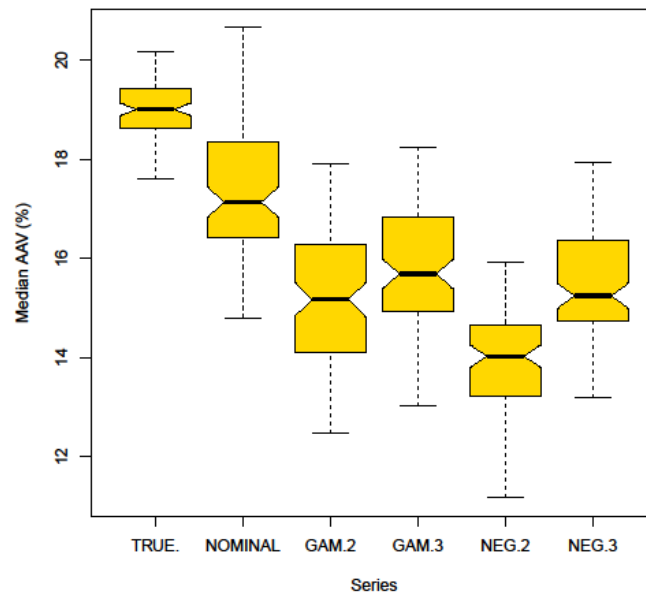
The AAV statistic is very common in MSE work that succinctly summarises how much the TACC (or effort in input control strategies) is changing from year to year (Rademeyer *et al.* 2007). The base statistic for each year is calculated as follows:

$$AAV_y^i = \left| \frac{TACC_{y+1}^i - TACC_y^i}{TACC_y^i} \right| \times 100 \quad (6.72)$$

with the median over years being used as the final summary statistic for each series denoted by the index i .

Figure 6.49 shows a boxplot summary of the AAV statistics for each of the CPUE series for a 5-year quasi-retrospective analysis. We use the phrase ‘quasi-retrospective’ because it is not a true retrospective analysis, where data are successively removed and the standardisation models re-fitted. It is quasi in the sense that we simply remove a year of the index at a time (for five years) and estimate the TACC that *would* have been predicted by the HS given those increasingly truncated data.

Figure 6.49: Boxplot summaries of the AAV statistics for each of the six CPUE series, for a five year quasi-retrospective, and across all 100 simulations.



The results are interesting, and in some cases might not have been expected. Firstly, the true index has the highest AAV across all the series - including the nominal series - but more predictably, shows the least variability around its central tendency. At first this might seem somewhat counter-intuitive, as the true index has no innate observation error or structural uncertainty issues like any model would have. Also, process error affects the other indices as much as the true index. The primary reason why the true index has a higher AAV than all others is that over the five year retrospective period the trend in the true index is downwards, whereas it is up for the nominal and very mixed for the other indices. This downwards trend in the true index drives successive decreases in TACC over this period, whereas it generally goes up slowly for the nominal series, and slightly down for the other series. This also adds to the increased AAV for the true index, as it is acting faster to decrease the TACC than any of the other series. So, in this sense, the lower AAV for the other indices is not necessarily a good thing - particularly for the nominal series - given the true exploitable abundance is going down slightly.

6.7.2 Multispecies assessment implications

At the most general level there is a fundamental linkage between all the species in the ETBF fishery: effort. At the annual level, the effort series that one might expect to have *some* kind of a relationship with fishing mortality (as is assumed in the WCPO assessments) is the same for all species. This is a common issue across all 'true' multispecies fisheries (i.e. multiple target species of direct economic importance caught by a common gear type). In the single-species stock assessment context this is not necessarily an issue, but it *is* an issue when it comes to setting reference points or management measures that relate in any way to fishing mortality - with F_{msy} being the most obvious example. If there is some kind of relationship between effort, E_y , and annual fishing mortality, F_y - and for argument's sake make it the simplest: $F_{s,y} = q_s E_y$ for species s in year y - then F_{msy} only makes sense at the fishery level (i.e. across all species) for $q_s = q$ for all species with identical life-histories and selectivities,. Alternatively, and less likely, they might all be different but still result in a single effort level that could produce F_{msy} for each species simultaneously. If the catchability parameters, q_s , and/or life-history and

selectivities are different, then what might be E_{msy} for one species will surely not be for another. The outcome being in cases such as these that there is no uniquely definable level of F_{msy} than can be attained for all species concurrently. There is no *a priori* reason as to why catchability would be the same across species - especially for say a tropical tuna and a billfish with very different life-history dynamics and feeding strategies. Also, while we tend to treat catchability parameters as nuisance scaling parameters that are rarely reported, they are indeed often very different across even similar species caught in the same gear and that includes the WCPO assessments.

The concept of multispecies maximum sustainable yield (MS-MSY) has been explored in a number of settings (ICES, multiple-species RFMOs), particularly where the single-species idea of MSY is the driving force behind the implemented management approach. At the fishery level, across all target species, the idea is to estimate the optimal effort level that maximises total yield, or some weighted sum of total species yield with a weighting given for various factors, such as economic importance for example. The most basic link between all such approaches is this: we need some kind of understanding of the across-species relationship between effort, catchability and F .

In this Section, we explore some simple state-space models that outline some of the things that *could* be explored in the multispecies CPUE context when considering inter-species correlative effects, differing catchabilities and so on, in the stock assessment paradigm. The current WCPO stock assessment models are very complex, spatially explicit integrated assessment models that do not lend themselves easily to these kinds of augmentations or explorations. Indeed, a number of the examples we explore here are, while very simple, actual multispecies models that would - at present - require significant development work to include in the MULTIFAN-CL assessment package favoured by the WCPFC. There are many reasons for the level of complexity required in the current WCPO assessments, so this is not a criticism; it just means that alternative approaches are often easier to do in much simpler models initially.

Single and multispecies state-space models

State-space models are ideal for the kind of explorations given they require an explicit treatment of the observation and process models separately. Issues common to multispecies fisheries occur in both the observation and process models and so this separation simply makes the process clearer. This first model is very simple: a random walk state-space model for the log-scale relative abundance of the LL exploitable population. We have the annual index we actually model: $I_y = \log CPUE_y$, and the following probability model is assumed:

$$\begin{aligned} n_{y+1} &\sim N(n_y, \sigma_n^2), \\ I_y &\sim N(n_y, \sigma_I^2), \end{aligned} \tag{6.73}$$

with initial state n_0 . Estimated parameters are the hidden states/random variables n_y , and the variance hyper-parameter σ_n . We will call this model M_0 for brevity. The next level model, M_E , includes an additional effort covariate model and is different to M_0 only in the population model

$$n_{y+1} \sim N(n_y - qE_y, \sigma_n^2). \tag{6.7.4}$$

where the additional estimated parameter is $q > 0$. Given we model log-scale relative abundance, and use the traditional rate-based interpretation of fishing mortality - i.e. proportion of actual abundance lost to fishing is $\exp(qE_y)$, then it is qE_y that is the term which appears in the process model for n_y in Eqn. (6.7.4). An interesting mathematical aside of this is model is that if E is mean-standardised, then q becomes the *de facto* mean F .

The first truly multispecies step (model M_{0s}) would then include the across-species observation and process error covariance structure into the simplest model. Now $\mathbf{n}_y$ is a vector, with each element the relative abundance of the particular species. As before, we have the initial states, $\mathbf{n}_0$, with the following process model:

$$\mathbf{n}_{y+1} \sim MVN(\mathbf{n}_y, \Sigma_n), \quad (6.7.5)$$

and observation model:

$$\mathbf{I}_y \sim MVN(\mathbf{n}_y, \Sigma_I), \quad (6.7.6)$$

where $MVN(\cdot)$ is the multivariate normal distribution, and Σ_n and Σ_I are the process and observation covariance matrices, respectively. The effort covariate extension to M_{0s} , model M_{Es} , would have the following process model:

$$\mathbf{n}_{y+1} \sim MVN(\mathbf{n}_y - \mathbf{q}E_y, \Sigma_n), \quad (6.7.7)$$

and $\mathbf{q}$ is the vector of species-specific catchability parameters, with the important feature that annual effort is shared across species. In all the multispecies extensions, the estimated hyperparameters include not just the species-specific variance parameters (diagonal elements of Σ_n) but also the correlation parameters that help define the off-diagonal elements of Σ_n .

These four putative models (M_0 , M_E , M_{0s} , and M_{Es}) cover a lot of possibilities and, at least at the multispecies end, require some sophisticated software and statistical concepts to sensibly estimate and choose between. What is worth covering off on first is what some of these somewhat abstract parameters may mean in the real world. Without effort data, these models lump together in one random component basically all factors that could change the relative abundance index from year to year: recruitment, mortality, catchability changes etc. With the effort covariate, it is then covering off on everything not linked to the specific fishery effort level E_y . One can increasingly add more covariates and/or model augmentations to account for these various factors explicitly, but in the models we explore here they cover a number of possible drivers. In the multispecies context the correlation parameters also have an ambiguous interpretation. For example: correlation in process errors could mean either catchability changes over time are correlated across species *and/or* actual changes in abundance in species are correlated. In the catchability space, a positive correlation might mean that you generally tend to get better catches of both species concurrently in the absence of actual abundance changes. In the abundance space, a negative correlation might indicate some kind of competitive effect, where higher abundances of one species tend to occur with lower abundances of another.

The main point is this: when dealing with the multispecies side of the problem in the generation of the abundance indices, there are flow-on model augmentations we can consider to the standard single-species assessment framework that might be of real interest. At the very least, a model that attempted a rigorous estimate of MS-MSY would, at *some* level, have to consider an observation and process model similar to those outlined in Eqns. (6.7.5) and (6.7.6), respectively. The whole point of the indices in the end is to estimate the trend in the exploitable abundance and, in the ETBF case, actually base TACC recommendations on those trends. Models that can accommodate the correlative side of both the observation and process sides of the problem are, generally speaking, in a better position to estimate real trends rather than spurious ones.

Choosing between these models is a little more complex than in the GLM space. These class of models fall into the random effect classification (the hidden/unobserved states n_y with

estimated variance/correlation parameters). This makes standard model-selection tools like the Akaike (AIC) or Schwartz (BIC) information criteria unusable. This is because the concept of the number of degrees-of-freedom (DOF) in these models is quite fluid. The *effective* DOF will be strongly linked to the actual values of the variance parameters (σ_n/Σ_n), so one cannot define a simple integer value of the additional DOF of a more complex model (like M_E in comparison with M_0) and use either the AIC/BIC or likelihood ratio test to choose between them.

To get around this, we use a fairly straightforward model-selection tool: leave-one-out cross-validation (LOOCV). What LOOCV does is to randomly remove a single point in the data series, in year y^* say, and fit the model to the remaining data. If we denote $\widehat{I}_{y^*}$ as the fitted index when considering all the data, we are interested in developing a useful statistic that summarises how well the model (with one less data point) predicts the removed data point, $\widetilde{I}_{y^*}$, relative to the model with all the data included, $\widehat{I}_{y^*}$. The root mean squared error (RMSE) for the assumed model $M_\bullet$ is a good option in this case:

$$RMSE(M_\bullet) = \sqrt{\mathbb{E} \left([\widehat{I}_{y^*} - \widetilde{I}_{y^*}]^2 \right)} \quad (6.7.8)$$

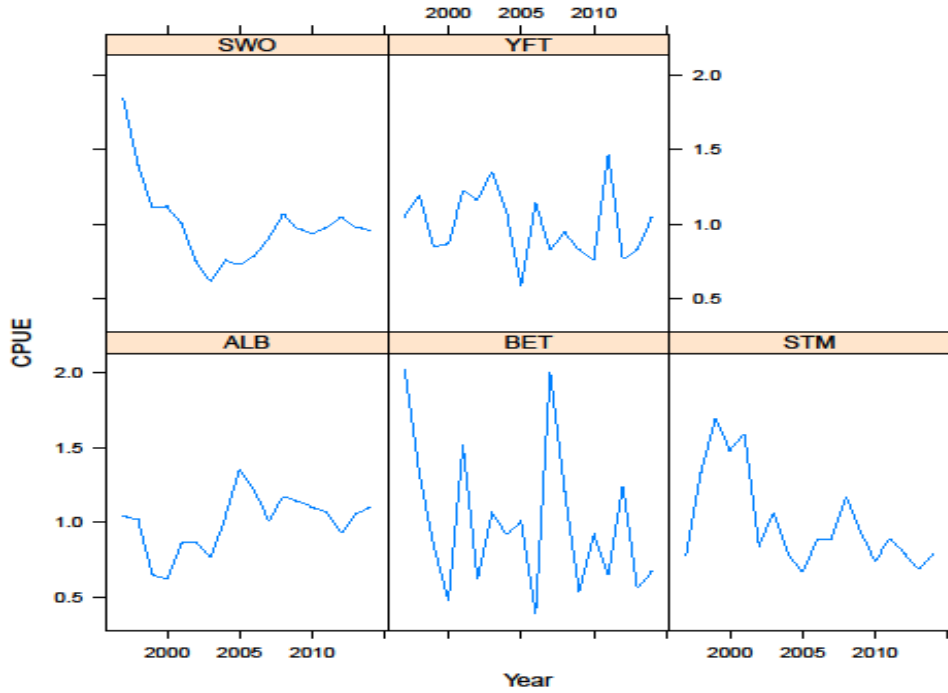
where the expectation is taken over the randomly removed data points y^* . The main idea is that, averaged over all the removed data points, the "best" model will produce predictions with **lower** values of the RMSE, relative to the comparison model(s). Models with too little freedom will tend to predict more poorly relative to ones with more flexibility (under-fitting), but as we increase their flexibility there tends to be some point where the predictions become poorer again (over-fitting). Models which minimise the RMSE are usually judged to be the most parsimonious, although we still have to be satisfied that they are actually fitting to the data well enough. The useful thing about using LOOCV is that we don't need to know about actual or even effective DOF to make it work, which is very useful in random effect models such as these. In all cases 500 iterations of the algorithm are used to compute the RMSE for each model.

6.6.3 Application to ETBF CPUE series

Figure 6.50 shows the (mean standardised) CPUE indices from 1997 to 2014 for each of the five main target species in the ETBF: Bigeye Tuna, Yellowfin Tuna, Swordfish, Striped Marlin, and Albacore Tuna. There are, given the number of species and models, a large number of potential species-model combinations we can explore. For example: some species show correlations with one another, whereas others do not; some species show an effort effect, whereas others do not. The LOOCV approach is applicable to all the various model combinations, with the only caveat being that we are comparing models with the same underlying observations inside them, so we can use this to get to what seem the 'best' overall species-model groupings. Detailing all the various steps, selections, and actual LOOCV statistics required in making this selection would take a lot of text and tables, and given we are trying to convey a more general message from the results, we focus on the outcomes not the intermediate steps.

The results are, perhaps, not surprising. Yellowfin Tuna and Bigeye Tuna group together in model M_s , inferring that there is no effort effect (it is estimated very close to zero and is deselected by the LOOCV process) but with process correlation between the two species. Swordfish and Striped Marlin group together in model M_E , so an apparent effort effect but with no apparent process correlation between them. Albacore Tuna sits alone in model M_0 with no apparent effort effect and showing no process correlation with any of the other species. We detail the results more thoroughly below, going by model-species grouping.

Figure 6.50: CPUE indices (mean standardised) from 1997 to 2014 for each of the five main target species in the ETBF.



Bigeye Tuna and Yellowfin Tuna model: M_s

To parameterise the process $S \times S$ covariance matrix, Σ_n , in the most general sense we need a total of S variance parameters, σ_s^2 , and $S(S-1)/2$ correlation parameters, $\rho_{s,t}$. To maintain positiveness the log-scale variance parameters are estimated, and we use the following transformation to model the correlation parameters

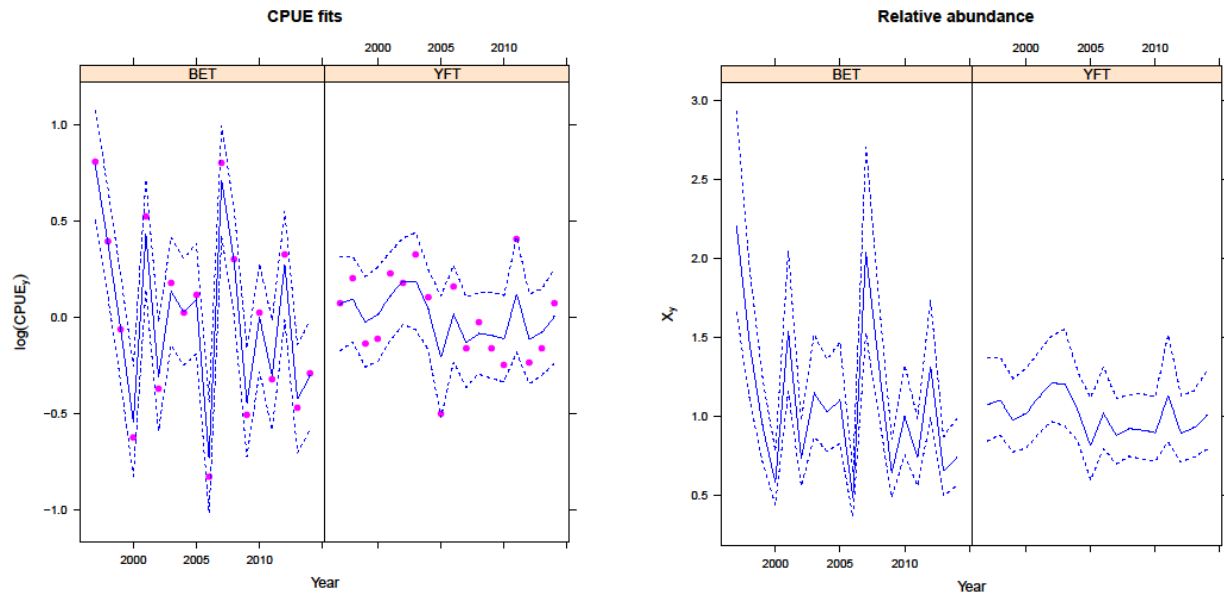
$$\rho_* = \frac{2}{1+e^{-\eta_*}} - 1, \quad (6.7.9)$$

where η_* is the estimated parameter, and by construction ensures that $\rho_* \in (-1,1)$. For the observation error matrix we do not have annual covariance estimates (in terms of total variation or correlation), so a time-independent diagonal observation error matrix with a CV of 0.15 in both cases was assumed.

The models were all estimated using the Template Model Builder (**TMB**) package in **R**, the latest and most efficient to date software for estimating the parameters of random-effect models such as the general class of models we consider here. The **TMB** package uses the Laplace approximation method and the expectation-maximisation algorithm to estimate the parameters, random effects and associated hyper-parameters. The random effects (herein, the unobserved relative abundance parameters n_y) are estimated first at putative values of the parameters (like q) and hyper-parameters (variance and correlation parameters). The Laplace approximation is used to obtain the marginal likelihood of the parameters/hyper-parameters which is then maximised to obtain their MLEs (the expectation-maximisation algorithm part). Standard asymptotic approaches and the delta method are used to obtain the standard errors of the parameters and random effects.

In terms of parameters values, the estimates of the standard deviation of the covariance matrix were 0.66 and 0.19 for the Bigeye and Yellowfin Tuna, respectively. The correlation MLE (and approximate 80%CI) was $\rho = -0.31$ (-0.71-0.15). Figure 6.51 summarises the fits to the data

Figure 6.51: Predicted log-scale CPUE (left) for the Bigeye and Yellowfin Tuna data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the right the predicted relative abundance, again in terms of the MLE and approximate 95% CI.



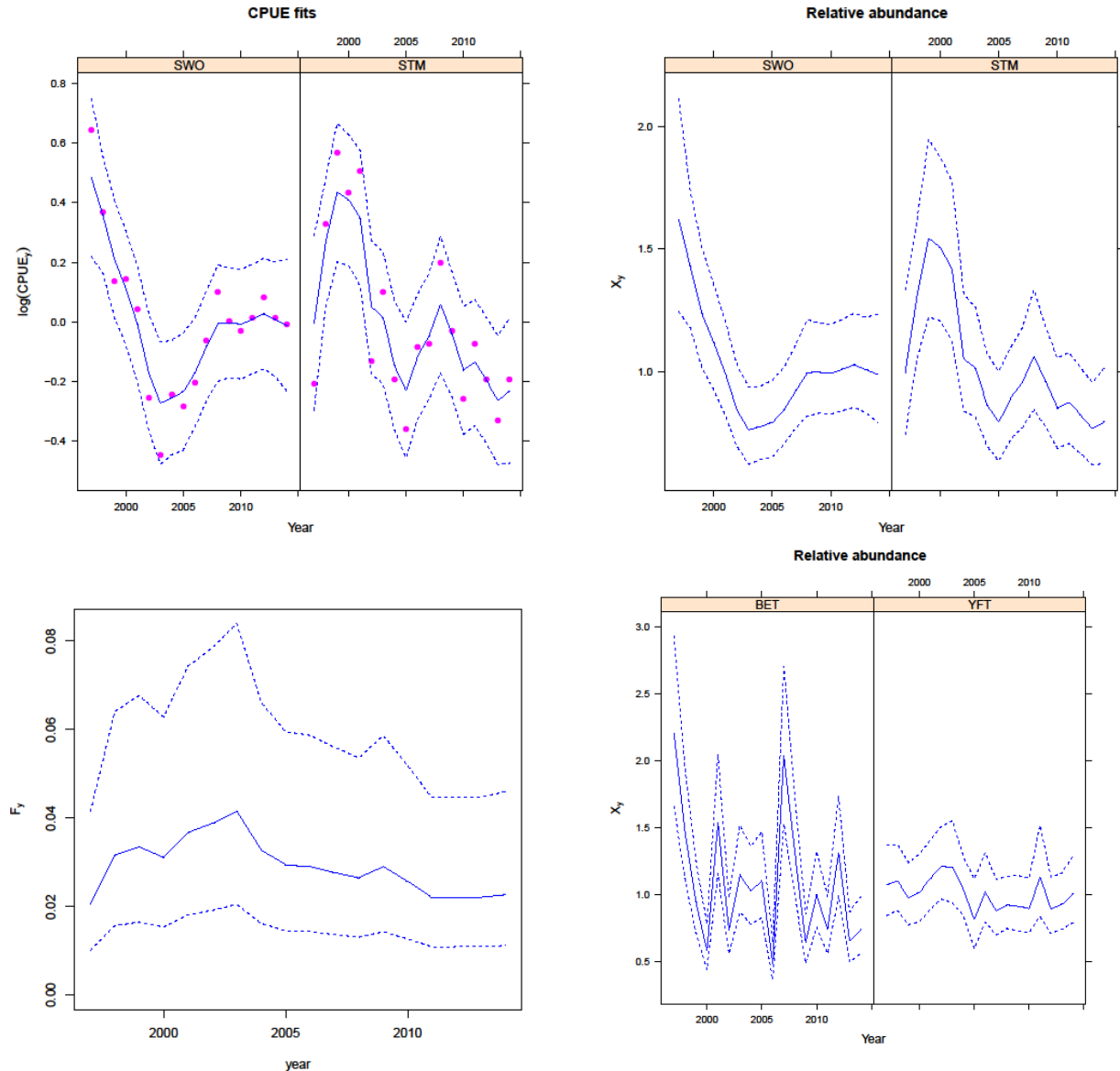
and the actual real-space abundance trends for both species. The model fits very well to the Bigeye Tuna data (interpreting the high variability as essentially process error), and less closely though relatively well to the Yellowfin Tuna data, just smoothing out the variability while maintaining the same marginal downward trend. The correlation estimates are not extremely precise given the length of the series (18 years in total), but they are coming out as negative, suggesting that higher/lower CPUE of one species *tends* to be somewhat weakly associated with lower/higher CPUE of the other. We cannot ascribe this to something specific, but obvious candidates would be a relationship in catchability or a population dynamic effect, where more of one species tends to result in less of the other (competition for resources/space).

Swordfish and Striped Marlin model: M_E

When treated separately, both these species exhibited a relationship between effort and relative abundance trend (as opposed to both Bigeye and Yellowfin Tuna). They also showed no process correlation linkage with either each other, or the other three tuna species. As a result, the model structure that best explained the Swordfish and Striped Marlin data was the M_E model. On closer inspection, both species had *very* similar estimated catchability parameters (Swordfish was higher by around 8%), but given their respective standard errors, were not in any way statistically significantly different. When compared (again via LOOCV) with a model where the two species share the *same* catchability (i.e. $\mathbf{q} = q$), the model with a single catchability was actually preferred.

Figure 6.52 summarises the fits, relative abundance trends, and the ETBF-only annual F estimates for Swordfish and Striped Marlin. Both data sets are fitted well. In terms of parameter estimates, the process variance MLEs were more similar than for the tropical tuna example: 0.13 for Swordfish and 0.18 for Striped Marlin. The correlation parameter was estimated to be zero, with a confidence interval evenly spread across the interval $(-1, 1)$, hence the model selection process removing this factor from the final model. Estimates of the ETBF-only F term were quite low but also very variable given the uncertainty in the estimates of q : peaking at around 0.02-0.08 in the early 2000s, decreasing towards the 0.015-0.04 more recently.

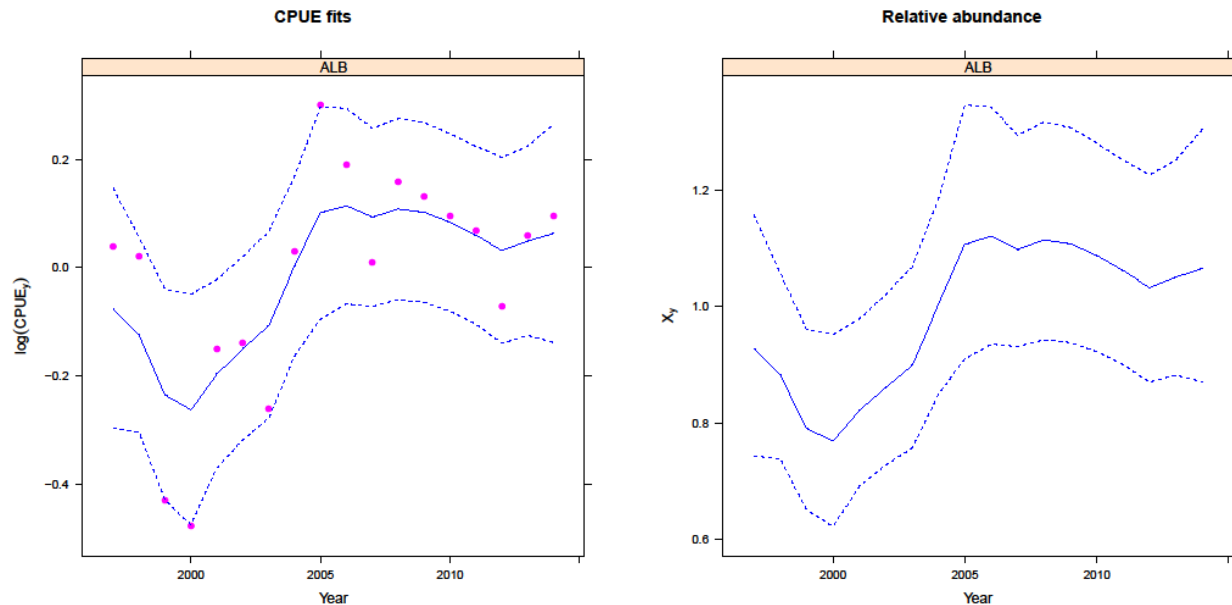
Figure 6.52: Predicted log-scale CPUE (top left) for the Swordfish and Striped Marlin data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the top right is the predicted relative abundance, again in terms of the MLE and approximate 95% CI. On the bottom left is the MLE and approximate 95% CI of the inferred F_y terms, given the effort time-series and the species-shared catchability term.



Albacore Tuna model: M_0

As mentioned at the start of this particular subSection, the Albacore Tuna stood alone in the sense that they showed neither an apparent effort effect, nor do they correlate in a process error sense with the other target species. As such, the simplest model (M_0) was selected by the LOOCV process. Figure 6.53 details the fits to the data and the relative abundance trend estimated in this case. The process error SD estimate was 0.13, with the data generally well fitted apart from the extremes in the series (both highs and lows) and an abundance trend showing an initial dip in the late 1990s, followed by an increase in the early 2000s and plateauing at the current level for the last decade almost.

Figure 6.53: Predicted log-scale CPUE (left) for the Albacore data (magenta circles) in terms of the MLE (thick blue line) and approximate 95% CI (dashed blue lines). On the right is the predicted relative abundance, again in terms of the MLE and approximate 95% CI.



6.7.4 Final Remarks

No obvious performance issues have been identified with respect to the harvest strategy for the Swordfish case. For all of the species considered in this Section the standardised models perform demonstrably better than just using the nominal series (which is always a comforting result). Additionally they don't seem to be increasing the noise relative to the true index and suggesting different TACC levels (5% +ve bias tops), resulting in AAV statistics that appear acceptable relative to other previous work. For example, Radermeyer *et al.* (2007) mentions 10-20% as something of a desirable range after lots of examples around the world especially in South Africa. These findings are consistent with the management plan for Southern Bluefin Tuna.

Some simple multispecies assessment explorations were undertaken for all the five major target species using statistically rigorous state-space models of relative abundance. Correlative and effort-related relationships were explored in terms of drivers of changes in relative abundance across the species. Bigeye Tuna and Yellowfin Tuna grouped together showing little to no relationship of abundance with changing fishing effort but showed a clear negative correlation – this could be abundance or catchability related. Swordfish and Striped Marlin grouped together both showing no process correlation but a clear relationship between abundance and fishing effort, and with (statistically speaking) the same overall catchability coefficient. Albacore formed a group on its own showing neither correlation with any of the other species nor any apparent effort-abundance relationship either. These kinds of analyses and statistical methods can serve to outline the complexity involved in implementing a general management framework like the Commonwealth Harvest Strategy Policy in the multispecies context. Clearly a single MSY/MEY type approach is not likely to be workable in the ETBF context, providing additional support to the more MSE-driven approach that the ETBF is currently being managed under which can take account of the various species-specific issues and trade-offs encountered.

7. Discussion and Conclusions

Australia's two tropical tuna fisheries (the Eastern Tuna and Billfish Fishery, ETBF, and the Western Tuna and Billfish Fishery, WTBF) are both multispecies fisheries that target a range of large pelagic fish. A major constraint for assessing multispecies fisheries, however, is a lack of reliable abundance indices that are a pre-requisite for the accompanying stock assessments, which have flow-on impacts to the results (i.e. TACCs) obtained from harvest strategies. Unlike single species fisheries where all effort is directed at the target species, in multispecies fisheries the effort is directed at a range of species. Consequently, the fishing effort needs to be standardised so that the 'effective' effort directed at any single species of interest can be ascertained. If this is not undertaken correctly then the resulting index of resource abundance is likely to be biased and unreliable. Although there are methods available that are currently used to standardise effort in multispecies fisheries, it is believed that new techniques need to be developed so that the resulting indices of resource abundance can be made more reliable.

In this project we have reviewed existing methods for standardising multispecies CPUE and a number of methods were selected for further development. While most models are based on the GLM framework developed for single species analyses, they included a number of extensions aimed at standardising the fishing effort to account for the differential targeting of species within a multispecies fishery. One set of extensions follows the method of He *et al.* (1997), who used cluster analysis to identify different fishing strategies based on the species composition of the catches from all fishing operations within a fishery. The utility of various clustering techniques (*Ward*, *clara*, *kmeans*) was investigated, together with clustering across different levels of data aggregation (i.e. at the set or trip level), as well as undertaking separate analyses at different temporal periods (e.g. monthly) to account for seasonal changes in species availability and assemblages. A second set of models followed the Principal Components Analysis initially developed by Winker *et al.* (2013, 2014) to identify targeted species groups in the data. Finally, a third, and new, set of Bayesian spatial-temporal models utilising the relatively new tool Integrated Nested Laplace Approximation (INLA) were developed. In total, the performance of twenty-four different model variations were investigated.

Two simulators, which are useful for comparing the performances of different methods at estimating known quantities, were also developed for testing the comparative performance of the standardising models. To help limit potential bias if the data structures incorporated into the simulator to match the analytical framework of one (or more) of the methods to be tested. The first simulator utilised an empirical approach, based on a framework generally known as the habitat-based-standardisation method. Data requirements include: the total number of hooks deployed by each longline set, the density of each species within each simulated spatial area during each month, the proportion of hooks for each longline set within each simulated depth stratum, and the proportion of fish for each species during each month and hour within in each simulated depth stratum. These were all based on direct observations made in the ETBF. For example, observations of hook depths were based on the 2050 individual time-depth recordings obtained from vessels deploying longlines in the ETBF during a previous FRDC funded project (Campbell and Young, 2010), while fish depth profiles were obtained from archival tags deployed on fish caught and tagged within the ETBF during several projects. On the other hand, the spatial-temporal density surfaces for each of the five species included in the simulator were modelled on the monthly distributions of nominal CPUE observed in the ETBF, while the spatial-temporal distribution of fishing effort across the fishery was taken to be the same as that observed in the ETBF during the years 2000 to 2014. Finally, in order to simulate

the stochastic characteristic of any catch, the final simulated catch was selected from a negative binomial distribution, and the probability of success for each species was selected so that the proportion of zeros in the distribution of simulated catches for each species was similar to that in the distribution of observed catches. To the extent that the simulator has made use of direct observations from the fishery, and has been tuned to replicate the distributions of catches in the fishery, it is hoped that the modelled catch and effort data generated by the simulator for analysis has many of the same characteristics of the data coming from the real fishery.

For the second simulator, catch and effort data were modelled using a more generic, flexible and individual vessel-based approach and was designed to generate catch rate data that reflected species abundance, targeting practises, and (unlike the ETBF simulator) individual vessel efficiency to capture the fact that different vessels have their own characteristic catchability and fishing behaviour. There is also spatial and seasonal variation in catch rates, which differed among species. While the first simulator was designed to generate data that mimics a specific fishery, the second simulator was designed to generate data for a more generic fishery. This allowed the standardising models to be tested across a wider range of fishery types than just that for the ETBF.

Targeting Practices in the ETBF

Before testing the ability of the selected standardising model to estimate the assumed annual abundance indices used to generate the simulated catch and effort data, it is important to understand the factors that influence the variability observed in the catch and effort data from the fishery being assessed. In particular, in a multispecies fishery such as the ETBF, it is important to understand and identify how fishers attempt to target the different species by alternating the manner in which they deploy their fishing gears. For this purpose, the information on the gear settings recorded in both logbooks and by observers in the ETBF were examined. Results indicate that there can be considerable variability in the manner that longline gears are deployed, both between years and within years on individual vessels. A substantive change was noted in 2006 when vessels began deploying more than 25 hooks-per-float (HPF). Commensurate with this change were shifts in the bait usage (more pilchard), light-stick usage (fewer) and set start-time (earlier in the day). The introduction of this new 'deep-longline' technique was due to a significant change in fishing strategy to incorporate the direct targeting of Albacore Tuna in this fishery.

Fortunately, each of these gear settings are recorded in the ETBF logbook and so these changes can be taken into account when standardising CPUE for changes in targeting and associated fishing strategies. On the other hand, there have also been changes in other gear settings which are not recorded in the logbook and which therefore cannot be accounted for in the CPUE standardisation. For example, there has been an increase in the mean length of the buoy-line deployed from around 10m to 15m during the 2000s, while the mean length of the branch-lines has decreased. The weights placed on the branch-lines (to assist in sinking the hooks to the fishing depth) have also increased over time. There was also a significant increase in the use of circle hooks after 2007 (and a corresponding decrease of J-hooks) due to management measures introduced to reduce turtle bycatch. Due to the lack of information on these gear types it is difficult to estimate how changes in the use of these gears impacts on catches, though the results of previously published research can provide some guidance. For example, Watson *et al.* (2005) found that Blue Shark catch rates were 8–9% higher on circle hooks compared to J hooks, while Ward *et al.* (2008) found that catch rates of Bigeye Tuna were higher on nylon than on wire leaders, after the latter were banned in ETBF in 2005. Whether or not there is sufficient information in the observer data for assessing how these and other changes in gear

usage not recorded in the logbook impacts on catches remains uncertain, but perhaps some consideration should be given to expanding the range of information on gear usage recorded in the logbook used in the ETBF (and other multispecies fisheries).

Analysis of the gear setting practices within a fishing trip indicate variable degrees of consistency in the deployment of each gear type. For example, similar hooks-per-float settings are used for 60-80% of all sets during a trip, while generally similar start-times are used for less than 20% of sets. Observer data also indicates that the recorded primary target species is not always the same for all sets deployed during a trip and suggests that multiple targeting strategies can be utilised within individual trips.

Examination of the particular gear settings associated with particular target species recorded by observers indicates a broad range of settings for each gear-type. While several characteristics of how the gears are deployed (e.g. buoy-line length and line-setting speed) do not display any appreciable variation between the different recorded target species, one can nevertheless discern particular groups of gear settings that have a higher usage when targeting particular species. For example, more hooks and hooks-per-float are deployed when targeting Albacore Tuna in comparison to the other species. Albacore targeting also generally uses fewer light-sticks, a shorter distance between branch-lines and longer buoy-lines, early morning sets, circle hooks and dead pilchard bait. On the other hand, when targeting Swordfish there is a preference to deploy the gear during late afternoon, use a high percentage of light-stick, a greater distance between branch-lines and dead squid bait. When targeting Striped Marlin, there is a preference for shallow sets (i.e. few hooks-per-float) with fewer hooks, minimum use of light-sticks, live mackerel bait and the use of a line-shooter. These observations indicate that while a range of gear settings are utilised when targeting particular species, there are combinations of gears that are more commonly used. The reasons for the wide-range of settings associated with targeted species remains uncertain, but may be due to differences in the costs (and ready availability) associated with different baits and the use of light-sticks, as well as differences in the depth range of species given differences in oceanographic characteristics at different times and locations. The weather is also likely to be an important factor influencing gear selection and deployment methods.

Examination of the catch composition of observed sets also found that for each of the main target species, the proportional catch of that species is highest when that species is recorded as the primary targeted species. This demonstrates that the vessel (skipper) has some ability to target and catch a desired species, and no is doubt correlated with the groups of gear settings associated with respective targeted species note above. Nevertheless, it is also clear that many other species are also usually caught, indicating that it is not possible to just target and catch a single species. Indeed, the target species sometimes is not the dominant catch. For example, on average more Yellowfin Tuna are caught than Striped Marlin when targeting the latter.

Based on these results it would seem obvious to simply recommend that the target species be recorded on the logbook by the vessel skipper. This would negate the need to use indirect methods (i.e. cluster analysis) to infer the fishing tactic or target species for use in the model used to standardise CPUE. However, such a recommendation is not needed as there is already a field in the logbook for the skipper to record the target species. While this field is usually completed, nevertheless, what seems simple in theory does not always translate to simplicity in practice. Indeed, comments from industry members on the Tropical Tuna Resource Assessment Group (TTRAG) indicate that the information recorded as the target species on the logbooks is likely to be unreliable, as the logbook is completed after the catch is landed on the

vessel on very often the recorded target species is just taken to be the species making up the largest proportion of the catch. A further complication is that often more than one species is recorded. As a consequence, TTRAG has recommended against using this information in the models used to standardise CPUE.

Analysis of the observer data also indicates that the proportion of fishing operations targeting particular species changes during the year. While Yellowfin Tuna is the dominant species targeted overall, there are periods during the year when the other species are the dominant target species. The targeting of Southern Bluefin Tuna over winter, together with the increased targeting of Bigeye Tuna in late autumn, of Yellowfin Tuna in spring and Swordfish over summer also correlate the changes in relative availability of these species throughout the year (as inferred from the changes in CPUE throughout the year). This suggests that the seasonal availability of the principal target species changes throughout the year and the fishers are able to change their targeting practices to avail themselves of these changes in relative abundance. Changes in the temporal availability of species is likely linked with the movement of fish associated with changes in the oceanographic conditions within the ETBF. Fish are not randomly distributed across the fishing grounds but tend to have distinct species-specific habitat preferences, which cause catch rates to vary according to the habitat that is targeted. These habitats are also likely to change their spatial size and location in response to prevailing currents and seasonal changes in oceanographic conditions. As a consequence, there is likely to be a strong seasonal component in both the distributions of fish species and related fishing and targeting strategies across any fishery.

The persistence of the spatial distributions for each species was investigated by calculating the Pearson correlation, R , between the four quarterly distributions of nominal CPUE. The spatial distribution of CPUE (assumed to be a proxy for biomass) was found to be most consistent for Broadbill Swordfish, where for the six quarterly comparisons R was between 0.84 and 0.95, and reasonable high for Yellowfin Tuna (R between 0.62 and 0.95). However, correlation between the quarterly distributions was much smaller for both Bigeye and Albacore Tuna (R generally less than 0.4). The same method was used to investigate the seasonal persistence of species associations and the results suggest that if associations between species do exist, due possibly to co-habitation within defined habitats, then these associations show generally weak persistence throughout the year (or across years). This has possible consequences for the ability to categorise the targeting practices of individual fishing sets using cluster analysis based on similarities in the catch composition.

Finally, the catch composition of observed sets also indicates that the proportional catch of a given species is generally highest when that species is recorded as the primary targeted species. This suggests that the vessel (skipper) has some ability to target and catch a desired species. Nevertheless, it is also clear that each of the other principal species are also usually caught, indicating that it is not possible to just target and catch a single species. Indeed, the target species sometimes is not the dominant catch. For example, on average more Yellowfin Tuna are caught than Striped Marlin when targeting the latter. Cluster analyses of the ETBF catch data indicate strong seasonal changes in species composition of the major clusters identified, which is likely to be related to seasonal differences in the co-occurrence of species reported above. Furthermore, while the fishing operations associated with some clusters display distinct differences in the gear configurations, most clusters were characterised by a broad mix of gear configurations, indicating that the relationship between the composition of the catch and the configuration of the gear is not strong. It was also noted that the mix of gear types associated with similar cluster types can vary on a seasonal basis (i.e. the dominant Yellowfin Tuna

cluster in summer and winter). These results indicate that there are substantial cluster by gear interactions and season-by-gear-by-cluster interactions. This suggests that 'targeting' and the consequent composition of the catch is likely to depend on more factors than just the configuration of the fishing gear, and exploring the nature of the relationship between the catch and these other factors (e.g. availability effects associated with the spatial location such as water temperatures, sea-mounts, eddy features) is encouraged.

The results of the cluster analysis also indicated that there can be substantive differences in the characteristics of each identified cluster, dependent upon which species are included in the analysis. This indicates that some thought needs to be given to how best to cluster the catch data for classifying sets and possibly identifying suites of species associated with different targeting practices. For example, should cluster analyses used to identify 'targeting' practices be based only on the main target species in a fishery or should such analyses also include by-product and by-catch species (even though these species are not targeted). The sensitivity of abundance indices on using the different cluster-types identified by these different types of analyses in procedures to standardise CPUE should also be investigated.

Comparative Performance of Standardisation Models

Due to practical reasons associated with the analyses being undertaken by three analysts independently (with each at a different location), the comparative analyses of the performance of the twenty-four different standardisation models selected for testing were broken into a number of different trials. The main results from these trials were as follows:

- The relative performance of each model varied considerably between the analyses conducted on the different species included in the two simulated data-sets. As a consequence, there was no one best performing model across all species. For example, the overall best performing model across all species in one set of trials was not the best model for any of the individual species analysed.
- Based on the analyses conducted on the ETBF simulated data, the mean Absolute Relative-Error ($ARE = \text{abs}[(\hat{I}_{s,t} - \hat{T}_{s,t})/\hat{T}_{s,t}]$, where $\hat{I}_{s,t}$ and $\hat{T}_{s,t}$ are the estimated and assumed abundance, respectively, for species s and year t after rescaling to have a mean of one) across all years and species for each of the 24 models ranged between 3.3% and 20% and averaged 9.5% over all models. This indicates that the difference between the estimated and assumed abundance index averages around 3% for the best performing model, and around 20% for the worst performing model, demonstrating that there were substantial differences in overall model performance.
- The overall best performing model fitted to the deterministic catches of the ETBF-simulated data-set was the model where the clustering was undertaken at the set level. The mean ARE for this model varied between 3.1% and 11.3% across the five species include in this data-set. On the other hand, the overall best performing model fitted to the catches randomly sampled from a negative-binominal distribution was the model where the clustering was undertaken at the trip level. For this model the mean ARE varied between 3.6% and 12.6% across the five species. The higher variability in species composition at the set level with the use of the randomly sampled catches, may be leading to a higher misallocation of sets using different fishing strategies, and aggregating the data across trips may help to reduce this variability, and therefore the misallocation of sets.
- Models which included a two-stage delta-GLM approach performed considerably better than models which only used a single staged approach. Analyses based on the current

model used to standardise CPUE in the ETBF indicate that the two-stage delta-GLM approach reduces the *ARE* of the estimated abundance indices by around 36% on average.

- Models which included the two gear-effects included in the simulated data also generally out-performed models which did not include these effects, i.e. the inclusion of gear effects in the standardising models improved the ability of such models to recover the assumed relative abundance indices than the inclusion of derived effects such as those based on catch-composition derived clusters. This result should not be unexpected, as it is the actual configuration of the fishing gear that is under the control of the fisher and which can be directly modified to 'target' different species. On the other hand, the use of a method such as a cluster analysis of the catch composition of a set to infer the target species is not only in-direct, but based on a result realized at the end of the set, and which is highly likely to be influenced by a number of other factors.
- For the subset of analyses with tested models which did not include the gear effects, but which were based on the two-stage delta-GLM approach, the model where the derived targeting effect was based on clustering undertaken at the trip level again displayed the best overall performance (with the mean *ARE* across the five species ranging between 6.3% and 12.2%). On the other hand, the model which did not include any targeting effect was the worst performing model. This last result indicates that the use of catch-based clusters to identify 'targeting' has merit, and inclusion of such derived effects, in general, improves the performance of the standardising model.
- Of the three PCA-based models tested, the model where the values of the principal components variables were transformed to zeroes and ones was the best performing, while the two other PCA-based models generally performed poorly.
- The results of fitting the models to the second simulated data-set were consistent with the results for the ETBF-based data-set, with the cluster-by-trip model being the best overall performer, followed by the cluster-by-set model. However, of the three PCA models tested, the model where the three principal components are fitted as cubic splines generally performed best.
- The best performing of the INLA-based models was also the most complex, which included both discrete and continuous spatio-continuous effects. Unfortunately, due to the computational demands of this model it could only be tested on a sub-set of the ETBF-data-set (and using the CSIRO high-performance computer), and this limited our ability to fully compare its performance with the other sets of the models. While the computational demands of this model may reduce its utility, nevertheless, this model displays considerable merit and warrants further comparative testing in the future.
- The current GLM used in the ETBF to standardise CPUE is performing reasonably well, no doubt due to the fact that this model incorporates several of the features (e.g. two-stage analysis, inclusion of gear effects) that were found to perform well in the simulations undertaken.

Despite the above results, a consistent observation from all the trials conducted was the variability in the performance of each model across the species included in the simulated data-sets. An example of this variability was demonstrated by the result that for one of the species analysed in the ETBF simulated data-set (ALB), models that did not include the gear effects out-performed the models that included these effects. This was clearly a different outcome than for the other four species analysed. The reasons for anomalous result remain uncertain, but provide an indication that there are factors specific to each species that influence the fit of any model to the data and as such the estimation of the annual abundance index.

One possible explanation may be the manner in which the ALB catch data was simulated, as the depth-profile for ALB used in the simulation model was not based on archival tag data (as was the done for the other four species) but was based on a created hybrid profile. Furthermore, in making this hybrid profile substantially different (i.e. deeper) than the observation-based depth-profiles for the other species (c.f. Figure 5.16), perhaps it is unrepresentative of the main species caught in the ETBF (and, in particular, the depth profile of ALB that it was supposed to represent). This may have resulted in a degree of mismatch between the depth profile simulated for ALB and the depths simulated for the fishing gears. For example, while the majority (85%) of hooks in the simulated data occur at depths less than 100m, only 23% of ALB, on average, are found within this depth range (c.f. Figure 5.16), with the consequence that the relation between gear settings and the subsequent catch of ALB is not as strong as for the other species.

Putting this possibility aside, some further analyses were undertaken to explore some of the other factors that may be influencing the anomalous outcome for ALB. First, analysis of the ETBF logbook data demonstrated that for catch-composition based clusters, the gear configurations varied across the different cluster-types and also varied seasonally within a year for similar cluster-types. However, the inclusion of gear-related interaction terms in the standardising models, ostensibly to account for such effects, did not improve the estimation of the abundance indices. However, this result could just be an indication that the relationship between catch-composition cluster types identified in the data and associated gear configurations is weak. Indeed, such a conclusion is consistent with the analysis of the ETBF logbook data that did not find particularly strong relationships between the catch-composition of the clusters identified and the gear settings (c.f. Figure 6.17). The analysis of the observer data for the ETBF, mentioned in Section 6.2, found that the highest proportion of the catch of any species generally occurred in those sets where that species had been specified as the primary target species. Unfortunately it is not possible to investigate whether such a relationship exists between the catch composition of the main cluster types and the species targeted, as this latter information is not reliably recorded in the logbooks. As such, how differences in targeting practices (partially expressed by differences in how the fishing gear is deployed) are expressed as differences in the resulting composition of the catch, and how these are related to the cluster-types identified in the overall catch data for the fishery, remains difficult to explore. An alternative method for investigating such relationships, based on clustering the fishing operations by gear configurations (instead of catch composition), is provided in Appendix C.

Second, in lieu of the gear-related interaction terms alone not providing the necessary explanatory power for improving the estimation of the abundance index for the anomalous results for the simulated ALB catches mentioned above, further investigations indicated that there had been a substantive change in the spatial distribution of catch rates for this species after 2005. Re-fitting the models to the simulated ETBF data-set, but limiting the analysis to the years after 2005, gave a result consistent with the results for the other four species: the models with gear effects out-performing the models without gear effects and the cluster-by-trip model performing best. There was also a substantive improvement in the estimation of the abundance index with around 50% reduction in the Absolute Relative Error. This result is consistent with some preliminary analyses undertaken using the simulated data for the ETBF which indicated that the spatial structure of the data has a greater impact on the ability of the model to infer the true annual abundance of a species than gear related factors such as the depth profiles of the fish.

This last analysis suggests that temporal changes in the spatial patterns of fishing and associated catch rates observed across a fishery are likely to have a substantive impact on the performance of models used to standardise CPUE, and fishery scientists undertaking such analyses should investigate whether such changes have occurred in fisheries they are assessing. Whether such patterns explain some of the variability in the performance of the models across the other species remains unclear, though this result does demonstrate that there may be a wide range of factors specific to each species that influence the fit of any model to the data, and as such the estimation of the annual abundance index. Indeed, the factors that influence the size and composition of the catch of any fishing operation are likely to be complex, and subject to a reasonable degree of stochastic variation due to the random nature of fish interacting with a given set of fishing gear, especially gear such as pelagic longline hooks. No doubt further research is required to identify and improve our understanding of the factors used by fishers to 'target' the deployed effort, and which control the composition of the catch for individual fishing operations, in particular those factors that account for the variability seen in the composition of the catch given similar gear settings. These factors are likely to include spatial features in the ocean such as temperature fronts and eddies that fishers often take cues from for setting their gears.

Implications for the ETBF Harvest Strategy

The harvest strategy for the ETBF is based on a decision-tree which utilises a number of empirical based indicators - the standardised CPUE of small, prime and large-sized fish and the proportion of these size classes in the total catch. The primary control rule for determining the RBCC in any year involves using the formula:

$$RBCC_{t+1} = TACC_t(1 + \beta \cdot S_{TARG})$$

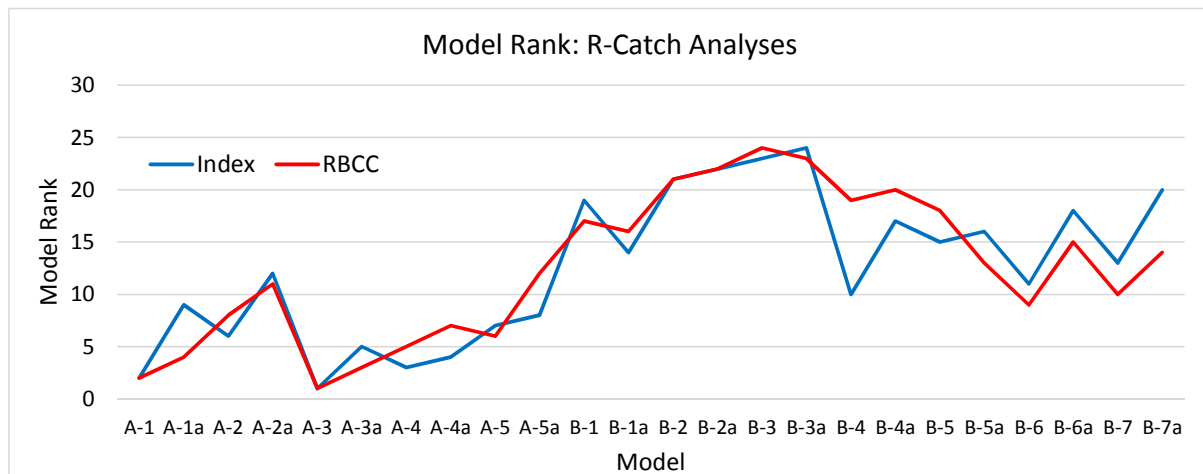
where S_{TARG} is the slope based on the angle subtended by the standardised prime-sized CPUE trend line over the previous five years and the line joining the present value of the CPUE and a target CPUE value a specified number of years, $N_{TARGET}=5$, in the future. The control parameter β is referred to as the feedback gain factor and for the ETBF harvest strategy $\beta=1$. Once the initial RBCC has been determined by the primary control rule it is then subject to review, and possible modification, based on the rules specified in the higher levels of the decision tree.

The performance of the harvest strategy in determining a 'correct' RBCC was tested using the same simulation framework for the ETBF as that used for testing the standardising models. Using the assumed abundance trend for each species used to generate the simulated ETBF data-set and the estimated standardised CPUE indices generated by each of the models used to analyse these data, the mean value of the Absolute Relative Error in the RBCC [$ARE = (RBCC_{STD-CPUE} - RBCC_{TRUE})/RBCC_{TRUE}$] was calculated across 100 data-sets where the catches were sampled from a random negative binomial distribution. As with the previous Index-based results, the size of the error in the RBCC was found to vary across the different species and models, such that no single model performed best across all species. For four of the species analysed, the mean error was generally less than 4% (and often less than 2%), but for one of the species the error was higher at around 6-7%. Results also indicate that the direction of the error was not consistent, being conservative and under-estimating the true RBCC for some species, while over-estimating the true RBCC for other species.

The best performing model was the delta-GLM cluster-by-trip model, closely followed by the base delta-GLM model currently being used in the ETBF. While the cluster-by-trip model was also the best-performing model for estimating the abundance index based on the same set of

analyses, it should be noted that the results are actually based on different time-intervals of the estimated abundance index. First, the results comparing the relative performance of each standardising model for estimating the assumed annual abundance time-series were based on the mean annual error over all 15 years in this time-series. On the other hand, the results comparing the relative performance of the each model for estimating the true RBCC were based on the errors in the trend of the estimated abundance index over only the last five years. Given these differences, and the similarity in the performance of several standardising models, it would not have been surprising to see different models perform differently across the two sets of results. However, in general there was a reasonable level of correlation between the performances of each model across the two sets of analyses (c.f. Figure 7.1). This is reassuring as it indicates that the same standardising model can be used for both sets of analyses.

Figure 7.1: Comparison of the overall rank achieved by each model (1=Best, 24=Worst) when used to estimate either the abundance index over all years or the RBCC. The analyses are based on the average errors calculated across 100 ETBF data-sets where the catches were sampled from a random negative binomial distribution.



The potential for the different performance of models dependent on the time-series of data being analysed, as noted above, also has implications for the question of over what time-period of data should an analysis be conducted taking into consideration the end-purpose for which the results are to be used. For example, if one is conducting a stock assessment then it is probably appropriate that the constructed abundance index be estimated over a time-period as long as possible (i.e. over the entire available data-set). On the other hand, if there is a need to use only a shorter time-period in the subsequent analysis (such as only using the temporal trend of the estimated abundance index over the last five years, as required for the EBF harvest strategy), then some consideration should be given as to whether the abundance index needs to be estimated over all the years that the data is available or some shorter period. This decision should be guided by investigation of both temporal changes in the characteristics of the data to be analysed that may influence (or bias) subsequent results, and the sensitivity of the constructed abundance index to changes in the time-series of data included in the analyses.

Finally, investigation of applying a LOWESS smoother to the annual abundance index before calculating the slope-to-target used in the ETBF harvest strategy, found that for some species the calculated RBCC varied to some extent, dependent on the smoothing parameter used. It was also found that the variability in the calculated RBCC increased as the variability of standardised CPUE over the last five years (used for calculating the slope used in the primary control rule) increased. This suggests that some care should be taken in selecting which

smoothing parameter to use and the criteria for identifying the most appropriate parameter requires further investigation.

Concluding Remark

Given the multitude of factors which are known to influence the catch of any fishing operation, the relationship between catch-rates and underlying stock abundance has long been a vexed one. This is especially the case with multispecies fisheries, where the issue of targeting is central to the relationship between the effectiveness of the fishing effort and the catch of any specific species. While there has been a long history of research directed at understanding the nature of these relationships, together with the development of methods aimed at standardising fishing effort so that the time-series of underlying stock abundance can be recovered, the complexity of the issue raises a number of challenges that prevent easy solutions. In this project we have investigated the nature of the relationship between the manner in which the fishing gears are deployed in the pelagic longline sector of the ETBF and the resulting species composition of the catch, and results indicate that such relations are often elusive. We have also selected, developed and tested using simulated data-sets, a suite of models to standardise catch rates and construct indices of relative stock abundance. While the results found that performance varied widely across the models tested, no single model was found to perform best across all the trials undertaken. Indeed, the variability of the performance of the models across the different species simulated was one of the most consistent characteristics of these results. While this project has investigated a number of factors likely to influence the composition of the catch and the ability of any model to reconstruct the true time-series of abundance from catch and effort data, these results do demonstrate that there may be a wide range of factors that influence the fit of any model to the data. Indeed, the factors that influence the size and composition of the catch of any fishing operation are likely to be complex, and subject to a reasonable degree of stochastic variation due to the random nature of fish interacting with a given set of fishing gear, especially gear such as pelagic longline hooks. No doubt continuing research is required to further identify and improve our understanding of the factors used by fishers to 'target' the deployed effort and which influence the composition of the catch for individual fishing operations, in particular those factors that account for the variability seen in the composition of the catch given similar gear settings. These factors are likely to include spatial features in the ocean such as temperature fronts and eddies that fishers often take cues from for setting their gears.

8. Implications

Standardised CPUE is a central input to stock assessments undertaken for many fisheries around the world. Within Australia, standardised CPUE is used in the assessments and harvest strategies for the two tropical pelagic fisheries (ETBF and WBTF) as well as many other fisheries, including the multispecies South-East Scalefish and Shark Fishery (SESSF). While the results of this project found there to be no single method which was best for standardising CPUE due to differences in the spatio-temporal distributions in the data for each species, the outcomes and associated recommendations (see Section 9) of this project will provide guidance on improving the methods used to standardise CPUE in these and other fisheries. In turn, this will ensure (i) more reliable and accurate stock abundance indices, (ii) improved inputs and to the harvest control rules dependent on standardised CPUE, (iii) improved outputs of harvest strategies, in particular the appropriateness of identified TACCs, and (iv) improvements to our ability to assess the resource status of non-target species, as required to achieve the management objective of ecologically sustainable fisheries.

The main end users of this research will be the assessment scientists developing indices of abundance based on the standardisation of CPUE, especially those within multispecies fisheries. As improvements in assessment results will have flow on effects for improved scientific advice and harvest strategy outcomes (e.g. improved TACCs), the managers of Australia's multispecies fisheries will also be direct beneficiaries. Improvements in management outcomes will lead to a reduction in the risk of over-or-under exploitation of the associated resources, benefiting the long-term sustainable management of these fisheries, the related fishery resources and related industries.

For fisheries such as the ETBF (valued at \$35m in 2014/15, ABARES 2015) and the SESSF (valued at \$67m in 2014/15, ABARES 2015), this reduction in risk will increase efficiencies (i.e. improve returns per vessel) and enhance the potential to achieve optimal exploitation levels, improving total value of these fisheries (potentially by many millions - far greater than the cost of the research). In this manner permit holders, the fishing industries, and more broadly the Australian community will be beneficiaries, through increased and sustained profits.

The outcomes of this project will also benefit fish stock assessments associated with the pelagic fisheries within the Western Central Pacific Ocean and Indian Ocean to which Australia's domestic tuna fisheries are connected.

9. Recommendations

Based on the outcomes of this project, the following recommendations are made:

1. The results of this project should be taken into consideration by fisheries scientists undertaking analyses of catch and effort data, and for selecting appropriate methods for standardising CPUE for stock assessment purposes.
2. The recording in logbooks of information on the characteristics of the fishing gears deployed by fishers at the set level should be encouraged and where this information is available it should be incorporated into the models used to standardise CPUE.
3. Logbooks should also record the target species for each set or shot. Where such a data field already exists in the logbook, effort should be made to improve both the reliability of this information and our understanding of exactly to what this data relates (e.g. is it just the most common species in the catch?).
4. The logbooks for Australia's two tropical tuna and billfish fisheries should be reviewed, to ensure that the appropriate information related to how the longline fishing gear is deployed can be recorded. Information considered for inclusion in the logbooks includes: hook-type, the length of the float and branch lines, and information related to the sag-ratio of the line.
5. Where gear information is not available, derived targeting effects should be considered for inclusion in models used to standardise CPUE. Based on the results of this study, the approach using a Cluster Analysis based on catch-composition to categorise sets is recommended over the use of the Principal Component Analysis approach. Where the appropriate information is available, and dependent on the nature of the fishing operations undertaken within a trip, consideration should be given to undertaking the cluster analysis at the trip level of data aggregation.
6. Before selecting a model to standardise CPUE for constructing an annual abundance index for a fishery, it is important that analysts attempt to understand the factors that influence the variability observed in the catch and effort data from the fishery being assessed. In particular, analysts should investigate changes in the spatio-temporal distribution of CPUE across a fishery and the factors that influence these changes.
7. Given the potential for the different performance of models dependent on the time-series of data being analysed, analysts should consider the appropriateness of the time-period of data analysed relative the end-purpose for which the results are to be used. This decision should be guided by investigation of both temporal changes in the characteristics of the data to be analysed that may influence (or bias) subsequent results, and the sensitivity of the constructed abundance index to changes in the time-series of data included in the analyses.
8. The investigation of temporal changes should also be guided by information that may not be available within the dataset, such as technology creep. For instance, if it is known that a fleet-wide adoption of a new technology resulted in a large suspected change in effective effort (e.g., GPS, colour sounders) during a defined period of time, then there is good reason to incorporate that temporal feature into the CPUE standardisation (especially if one has month of adoption data for each vessel).
9. Further work should be undertaken to continue the review, development and comparison of the performance of approaches to standardise catch rates for Australia's multispecies fisheries. For example, the simulation framework presented in this report could be extended to model other assumed abundance trends so that models used to standardise CPUE could be evaluated across a greater range of hypothetical (but

plausible) conditions. Also, further work is required to investigate of the performance of the INLA based spatio-temporal model developed by this project, as well as the performance of the spatial dynamic factor analysis model recently published by Thorson *et al.* (2016).

10. Finally, further research is required to identify and improve our understanding of the factors used by fishers to 'target' the deployed fishing effort to control the composition of the catch for individual fishing operations, in particular those factors that account for the variability seen in the composition of the catch given similar gear settings. These factors are likely to include spatial features in the ocean such as temperature fronts and eddies that fishers often take cues from for setting their gears.

10. Extension and Adoption

The outcomes of this project will be extended and communicated to a range of end-users in the following ways.

Resource Assessment Groups

Through their membership on a number of Resource Assessment Groups (RAGs) a number of the project team will be able to use the outputs of this project to inform the adoption of new and improved methods for future assessments. This will have direct benefits on improving the accuracy and reliability of the indices of resource abundance based on the standardised CPUE as well as the assessments of the related fish resources. This will have flow on benefits for the utility of the harvest strategies used to manage these fisheries, in particular the ETBF and WTBF and other multispecies fisheries such as the SESSF, which are highly dependent on the accuracy of the standardised CPUE.

The Principal Investigator (R.C.) is the principal assessment scientist for the ETBF (currently undertaking both the CPUE standardisation and the harvest strategy calculations for this fishery) and together with R.H. is a scientific representative on the Tropical Tuna Resource Assessment Group (TTRAG). Another member of the project team (M.H.) is a principal assessment scientist for the Southern and Eastern Scalefish and Shark Fishery (SESSF), which is a multi-sector, multispecies fishery that covers almost half of the Australian Fishing Zone, and a scientific representative on the associated Resource Assessment Group. Finally, S.Z. is a member of the Squid Fishery Resource Assessment Group. Through these multiple roles these team members will be able to consult widely with the relevant RAGs, AFMA, the fishing industry, and other scientists about the outcomes of this project and implement the results to improve the CPUE standardisations used in the stock assessments and the related inputs into the associated harvest strategies. In particular, the outcomes of this project will inform the review of the ETBF harvest strategy, and the associated methods to standardise CPUE, to be undertaken during 2017 and as part of this process the results of this project were presented to the meeting of TTRAG held 28-30 March 2017.

Regional Fisheries Management Organisations

This project will have direct input into scientific work and programs undertaken by, and on behalf of, the Western Central Pacific Fisheries Commission (WCPFC) and the Indian Ocean Tuna Commission (IOTC).

The Principal Investigator (R.C.) has attended the annual meetings of the Scientific Committee (and its predecessor) for WCPFC since the mid-1990s, and more recently the Pre-Assessment Workshops convened by the Scientific Service Provider (Secretariat of the Pacific Community, SPC) in April each year. He also contributes standardised CPUE indices to several assessments undertaken by SPC (e.g. southwest Pacific Swordfish and Striped Marlin) and has consulted widely with other scientists in this region on the need for improving CPUE standardisation for the associated assessments (e.g. Hoyle *et al.* 2014). As part of these ongoing contributions the results of this project were presented to the Workshop on the Analysis of CPUE for Stock Assessments, held 20-21 April at SPC in Noumea. This meeting was held as part of the Preparatory Workshop for the 2017 Stock Assessments for the Western Central Pacific Fisheries Commission.

Another member of the project team (S.H.) regularly undertakes stock assessments for tuna and billfish stocks in the Indian Ocean and attends the annual meetings of the Scientific Committee for the IOTC while S.Z. has consulted with IOTC on a project to undertake CPUE studies for neritic tunas. The outputs of this project will inform the adoption of new and improved methods for future assessments on the fish stocks in this region. A third member of the project team (R.H.) is a member of Scientific Committee for the Convention of the Conservation of Southern Bluefin Tuna (CCSBT) and will be able to convey the results of this project to this forum.

Broader Dissemination

Finally, the results of the project will be published in relevant peer-review journals to help disseminate the results to the broader scientific community.

11. References and Further Reading

- ABARES 2015. *Australian fisheries and aquaculture statistics, 2015*. Department of Agriculture and Water Resource, Canberra.
- Abbott, J.K., Haynie, A.C., and Reimer, M.N. 2015. *Hidden flexibility: institutions, incentives, and the margins of selectivity in fishing*. Land Econ. 91: 169–195.
- Allen, R.L., Punsley, R.G., 1984. *Catch rates as indices of abundance of Yellowfin Tuna, Thunnus albacares, in the Eastern Pacific Ocean*. Bull. Inter-Am. Trop. Tuna Comm. 18: 301–379.
- Basson, M., Hobday, A.J., Eveson, J.P. and Patterson, T.A. 2012. *Spatial interactions among juvenile southern bluefin tuna at the global scale: a large scale archival tag experiment*. Final Report for project 2003/002, Fisheries Research and Development Corporation. Canberra.
- Beguin, J., Martino, S., Rue, H., and Cumming, S.G. 2012. *Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation*. Methods in Ecol. Evol. 3: 921–929.
- Beverton, R.J.H., and Holt, S.J. 1957. *On the dynamics of exploited fish populations*. Chapman & Hall, London.
- Beverton, R.J.H., Parrish, B.B., 1956. *Commercial statistics in fish population studies*. Rapp. Proc. Verb. Reun. Cons. Int. Explor. Mer. 140: 58–66.
- Bigelow, K. A., C. H. Boggs, and He, X. 1999. *Environmental effects on Swordfish and blue shark catch rates in the, US North Pacific longline fishery*. Fish. Oceanogr. 8, 178-198.
- Bigelow, K., Hampton, J. and Miyabe, N. 2002. *Application of a habitat-based model to estimate effective longline fishing effect and relative abundance of Pacific Bigeye Tuna (Thunnus obesus)*. Fish. Oceanogr. 11: 143-155.
- Biseau, A. 1998. *Definiton of a directed fishing effort in a mixed-species trawl fishery, and its impact on stock assessments*. Aquat. Living Resour. 11: 119-136.
- Bishop, J., Venables, W. N., Dichmont, C. M., and Sterling, D. J. 2008. *Standardizing catch rates: is logbook information by itself enough?* ICES Journal of Marine Science, 65: 255-266.
- Blangiardo, M., Cameletti, M., Baio, G. and Rue, H. 2013. *Spatial and spatio-temporal models with R-INLA*. Spatial and Spatio-temporal Epidemiology 7: 39–55
- Blangiardo, M., Cameletti, M. 2015. *Spatial and Spatio-temporal Bayesian Models with R-INLA (First Edit)*. John Wiley & Sons.
- Boggs, C.H. 1992. *Depth, capture time, and hooked longevity of longline-caught pelagic fish: timing bites of fish with chips*. Fish. Bull. 90: 642-658.
- Brill, R. 1994. *A review of temperature and oxygen tolerance studies of tunas pertinent to fisheries oceanography, movement models and stock assessment*. Fish. Ocean. 3: 204-216.
- Campbell, R.A. 2004. *CPUE standardisation and the construction of indices of stock abundance in a spatially varying fishery using general linear models*. Fish. Res. 70: 209-227.
- Campbell, R.A. 2012. *Aggregate and size-based standardised CPUE indices for longline target species caught within the ETBF*. Working paper presented to the 5th meeting of the Tropical Tuna RAG, held 4-5 September 2012, Canberra.

- Campbell, R.A. 2015. *Constructing stock abundance indices from catch and effort data: Some nuts and bolts*. Fish. Res. 161: 109–130.
- Campbell, R. 2016a. *Aggregate and size-based standardised CPUE indices for longline target species caught within the ETBF – 2016 Update*. Working paper presented to the 15th meeting of the Tropical Tuna RAG, held 11-12 July 2016, Mooloolaba
- Campbell, R. 2016b. *Implementation of the ETBF Harvest Strategy and calculation of Recommended Biological Commercial Catches for Broadbill Swordfish and Striped Marlin for the 2017/18 Quota Year*. Working paper presented to the 16th meeting of the Tropical Tuna RAG, held 18-19 October 2016, Mooloolaba.
- Campbell, R. 2016c. *Fishery indicators for tropical tunas within the ETBF and the WCPO – 2016 update*. Working paper presented to the 16th meeting of the Tropical Tuna RAG, held 18-19 October 2016, Mooloolaba.
- Campbell, R. 2016d. *Annual catch by fleet and fishing method within the southwest Pacific – 2016 Update*. Working paper presented to the 16th meeting of the Tropical Tuna RAG, held 18-19 October 2016, Mooloolaba.
- Campbell, R.A. and Young, J. 2010. *Determination of Effective Longline Effort in the Eastern Tuna and Billfish Fishery*. Final report for project 2005/004, Fisheries Research and Development Corporation, Canberra.
- Cao, J., Chen, X., Chen, Y., Liu, B., Ma, J., & Li, S. 2011. *Generalized linear Bayesian models for standardizing CPUE: an application to a squid-jigging fishery in the northwest Pacific Ocean*. Scientia Marina 75: 679–689.
- Carruthers, T.R., Ahrens, R.N.M, McAllister, M.K. and Walters, C.J., 2011. *Integrating imputation and standardization of catch rate data in the calculation of relative abundance indices*. Fish. Res. 109: 157-167.
- Carruthers, T.R., McAllister, M.K. and Ahrens, R.N.M., 2010. *Simulating spatial dynamics to evaluate methods of deriving abundance indices for tropical tunas*. Can. J. Fish. Aquat. Sci. 67: 1409–1427.
- Carvalho, F.C., Murie, D.J., Hazin, F.H.V., Hazin, H.G., Leite-Mourato, B., Travassos, P. and Burgess, G.H., 2010. *Catch rates and size composition of blue sharks (Prionace glauca) caught by the Brazilian pelagic longline fleet in the south western Atlantic Ocean*. Aquat. Living Resour. 23: 373–385.
- Carvalho, F., Ahrens, R., Murie, D. and Ponciano, J.M., Aires-de-Silva, A., Maunder, M., Hazin, F. 2014. *Incorporating specific change points in catchability in fisheries stock assessment models: An alternative approach applied to the blue shark (Prionace glauca) stock in the south Atlantic Ocean*. Fish. Res. 154: 135-146.
- Davies, C, R. Campbell, J. Prince, N. Dowling, D. Kolody, M. Basson, K. McLoughlin, P. Ward, I Freeman and A. Bodsworth. 2008. *Development and Preliminary Testing of the Harvest Strategy Framework for the Eastern and Western Tuna and Billfish Fisheries*. Final report to AFMA, Canberra.
- Dunstan, P.K., Foster, S.D. and Darnell, R. 2011. *Model based grouping of species across environmental gradients*. Ecol. Mod. 222: 955-963.
- Diggle, P. and Ribeiro, P. J. 2007. *Model-Based Geostatistics*. Springer, New York. 242 pp.
- Dunstan, P.K., Foster, S.D. and Darnell, R. 2011. *Finite mixture of regression modeling for high-dimensional count and biomass data in ecology*. J. Agric. Biol. Env. Stat. 18: 357-375.

- Evans, K. 2010. *Investigation of local movement and regional migration behaviour of Broadbill Swordfish targeted by the ETBF*. Final report for project 2006/809, Australian Fisheries Management Authority, Canberra. 95pp.
- Evans K., Patterson, T. and Pedersen, M. 2011. *Movement patterns of Yellowfin Tuna in the Coral Sea region: during connectivity with stocks in the western Pacific Ocean region*. Final report for project 2008/804, Australian Fisheries Management Authority, Canberra, 65pp.
- Foster, S.D., Givens, G.H., Dornan, G.J., Dunstan, P.K. and Darnell, R. 2013. *Modelling biological regions from multispecies and environmental data*. *Environmetrics* 24: 489-499.
- Gelman, A., and Pardoe, I. 2006. *Bayesian measures of explained variance and pooling in multilevel (hierarchical) models*. *Technometrics* 48: 241–251.
- Glazer, J.P. and Butterworth, D.S. 2002. *GLM-based standardization of the catch per unit effort series for the South African west coast hake, focusing on adjustment for targeting other species*. *Afr. J. Med. Sci.* 24: 323–339.
- Gulland, J.A., 1956. *On the fishing effort in English demersal fisheries*. *Fish. Invest.* 20 (Series 2): 1–41.
- Gunn, J., Hampton, J., Evans, K., Clear, N., Patterson, T., Bigelow, K., Langley, A., Leroy, B., Williams, P., Miyabe, N., Sibert, J., Bestley, S. and Hartmann, K. 2005. *Migration and habitat preferences of Bigeye Tuna, Thunnus obesus, on the east coast of Australia*. Final report for project 1999/109, Fisheries Research and Development Corporation, Canberra, 204pp.
- Hampton, J., Bigelow K. and Labelle, M. 1998. *Effect of longline fishing depth, water temperature and dissolved oxygen on Bigeye Tuna (Thunnus obesus) abundance indices*. Working paper 17 presented to the 11th meeting of the Standing Committee on Tuna and Billfish, held 28 May-6 June 1998, Honolulu, Hawaii.
- Hanamoto, E. 1987. *Effect of oceanographic environment on Bigeye Tuna distribution*. *Bull. Jpn. Soc. Fish. Oceanogr.* 51: 203–216.
- Harley, S.J., Myers, R.A., and Dunn, A. 2001. *Is catch-per-unit-effort proportional to abundance?* *Can. J. Fish. Aquat. Sci.* 58: 1760–1772.
- Hartigan, J. A. and M. A. Wong 1979. Algorithm AS 136: A *k*-means clustering algorithm. *J. Royal Stat. Soc..Series C (Applied Statistics)* 28: 100-108.
- He, X., Bigelow, K. and Boggs, C. 1997. *Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery*. *Fish. Res.* 31: 147-158.
- Hilborn, R. 1985. *Fleet dynamics and individual variation: why some people catch more fish than others*. *Canadian Journal of Fisheries and Aquatic Sciences*, 42: 2-13
- Hillary, R. M., Preece, A., Kolody, D., Evans, K., and Davies, C. R. 2016. *Development of an approach to harvest strategy management of internationally managed multispecies fisheries*. Final report 2013/203, Fisheries Research and Development Corporation, Canberra.
- Hinton, M.G. and Nakano, H. 1996. *Standardizing catch and effort statistics using physiological, ecological, or behavioural constraints and environmental data, with an application to blue marlin (Makaira nigricans) catch and effort data from the Japanese longline fisheries in the Pacific*. *Inter-Am. Trop. Tuna Comm.* 21: 169-200.
- Hinton, M.G., Maunder, M.N. 2003. *Methods for standardising CPUE and how to select among them*. Working paper MWG-7 presented to the 16th meeting of the Standing Committee on Tuna and Billfish, held 9-16 July 2003, Mooloolaba, Australia.

- Holland, K., Brill, R.W. and Chang, R.K.C. 1990. *Horizontal and vertical movements of yellowfin and bigeye tuna associated with fish aggregating devices*. Fish. Bull. 88: 493-507.
- Hoyle, S. and Okamoto, H. 2013. *Target changes in the tropical WCPO Japanese longline fishery, and their effects on species composition*. Working paper SA-IP-04 presented to the 9th regular meeting of the Scientific Committee for the Western and Central Pacific Fisheries Commission, held 6-14 August 2013, Pohnpei, Federated States of Micronesia.
- Hoyle, S., Langley, A.D., Campbell, R.A. 2014. *Recommended approaches for standardising CPUE data from pelagic fisheries*. Information paper SA-IP-10 presented to the 10th Scientific Committee for the Western Central Pacific Fisheries Commission, held 6-14 August 2014, Majuro, Republic of the Marshall Islands.
- Hui, F.K.C., Warton, D.I., Forster, S.D. and Dunstan, P.K. 2013. *To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models*. Ecology. 94: 1913-1919.
- Kaufman, L. and P. J. Rousseeuw. 2009. *Finding groups in data: an introduction to cluster analysis*, John Wiley & Sons.
- Kimura, D. K. 1981. *Standardized measures of relative abundance based on modelling $\log(c.p.u.e)$, and their application to Pacific ocean perch (*Sebastes alutus*)*. Journal du conseil / Conseil international pour l'exploration de la mer, 39: 211-218
- Kimura, D. K. (1988). *Analysing relative abundance indices with log-linear models*. North American Journal of Fisheries Management, 8(2), 175-180.
- Klaer, N.L. and Smith, D.C. 2012. *Determining primary and companion species in a multispecies fishery: Implications for TAC setting*. Mar. Policy 26: 606-612.
- Kolody, D.S., A.L. Preece, C.R. Davies, J.R. Hartog and N.A. Dowling. 2010. *Integrated evaluation of management strategies for tropical multispecies long-line fisheries*. Final report for project: 2007/017, Fisheries Research and Development Corporation, Canberra.
- Langley, A., Bigelow, K. and Maunder, M. 2005. *Longline CPUE indices for yellowfin and bigeye in the Pacific Ocean using GLM and statistical habitat standardisation methods*, Working paper SA WP-8 presented to the 1st meeting of the Scientific Committee for the Western Central Pacific Fisheries Commission, held 8-19 August, Noumea, New Caledonia.
- Leaper, R., Dunstan, P.K., Forster, S.D., Barrett, N.S. and Edgar, G.J. 2014. *Do communities exist? Complex patterns of overlapping marine species distributions*. Ecology 95: 2016-2025.
- Lindgren, F., Rue, H. 2014. *Bayesian spatial modelling with R-INLA*. J. Statistical Software, 55: 1-26.
- Lo, N., Jacobson, L.D., Squire, J.L., 1992. *Indices of relative abundance from fish spotter data based on delta-log normal models*. Can. J. Fish. Aquat. Sci. 49: 2515-2526.
- Lynch, P.D., Shertzer, K.W., Latour, R.J. 2012. *Performance of methods used to estimate indices of abundance for highly migratory species*. Fish. Res. 125-126: 27-39.
- Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, M. Studer and P. Roudier 2014. *Package 'cluster'*.
- Maunder, M.N. and Punt, A.E. 2004. *Standardizing catch and effort data: a review of recent approaches*. Fish. Res. 70: 141-159.
- Maunder, M. and Hinton, M.J. 2006. *Estimating relative abundance from catch and effort data, using neural networks*. Special Report 15, Inter-Amer-Trop. Tuna Comm.

- Maunder, M., Hinton, M.J, Bigelow, K.A and Langley, A.D. 2006. *Developing indices of abundance using habitat data in a statistical framework*. Bull. Mar. Sci. 79: 545-559.
- Maunder, M., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P. and Harley, S.J., 2006. *Interpreting catch per unit effort data to assess the status of individual stocks and communities*. ICES J. Mar. Sci. 63: 1373-1385.
- McCullagh, P. and J.A. Nelder 1989. *Generalized Linear Models (second edition)*. Monographs on Statistics and Applied Probability 37, 511p.
- McKechnie, S., Hoyle, S. and Harley, S.J. 2013. *Longline CPUE series that account for changes in the spatial extent of fisheries*. Information paper SA-IP-05 presented to the 9th meeting of the Scientific Committee for the Western Central Pacific Fisheries Commission, held August 2013, Pohnpei, Federated States of Micronesia.
- Muñoz, F., Pennino, M. G., Conesa, D., López-Quílez, A. and Bellido, J. M. 2013. *Estimation and prediction of the spatial occurrence of fish species using Bayesian latent Gaussian models*. Stochastic Environmental Research and Risk Assessment 27: 1171–1180.
- Murtagh, F. and P. Legendre 2014. *Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion?* J. Classification 3: 274-295.
- Nelder, J. A., & Wedderburn, R. W. M. 1972. *Generalized linear models*. Journal of the Royal Statistical Society. Series A (General), 135: 370-384.
- Ono, K., Punt, A. E., & Hilborn, R. 2015. *How do marine closures affect the analysis of catch and effort data?* Canadian Journal of Fisheries and Aquatic Sciences, 72: 1177-1190
- Ovaskainen, O., Hottola, J. and Sittonen, J. 2010. *Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions*. Ecology 91: 2514-2521.
- Ovaskainen, O. and Soinnen, J. 2011. *Making more out of sparse data: hierarchical modeling of species communities*. Ecology 92: 289-295.
- Pelletier, D. and Ferraris, J., 2000. *A multivariate approach for defining fishing tactics from commercial catch and effort data*. Can. J. Fish. Aquat. Sci. 57: 51–65.
- Pennington, M, 1983. *Efficient estimators of abundance for fish and plankton surveys*. Biometrics 39: 28 1-286.
- Pennington, M. and Berrien, P. 1984. *Measuring the precision of estimates of total egg production based on plankton surveys*. J. Plankton Res. 6: 869-879.
- Pennington, M. 1996. *Estimating the mean and variance from highly skewed marine data*. Fish. Bull. 94: 498-505
- Pledger, S. and Arnold, R. 2014. *Multivariate methods using mixtures: Correspondence analysis, scaling and pattern-detection*. Computational Statistics and Data Analysis 71: 241-261.
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Veski, P.A., McCarthy, M.A. 2014. *Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM)*. Methods in Ecology and Evolution 5: 397-406.
- Preece, A.L., Campbell, R.A. and Hillary, R.M. 2013. *Investigation of possible changes in fishing strategies under quota management and implications for the ETBF harvest strategy*, Final report for project 2010/811, Australian Fisheries Management Authority, Canberra.
- Punt, A.E., Walker, T.I., Taylor, B.L. and Pribac, F. 2000. *Standardization of catch and effort in a spatially-structured shark fishery*. Fish. Res. 45: 129-145.

- Quirijns, F.J., Poos, J.J. and Rijnsdorp, A.D. 2008. *Standardising commercial CPUE data in monitoring stock dynamics: Accounting for targeting behaviour in mixed fisheries*. Fish. Res. 89: 1-8.
- R Core Team 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rademeyer, R.A., Plaganyi, E.E. and Butterworth, D.S. 2007. *Tips and tricks in designing management procedures*. ICES J. Mar. Sci. 64: 618-625.
- Raiche, G. 2010. *nFactors: an R package for parallel analysis and non-graphical solutions to the Cattell scree test*. R package version 2.3.3.
- Robson, D.S., 1966. *Estimation of relative fishing power of individual ships*. ICNAF Res. Bull. 2: 5-14.
- Rogers, J.B. and Pikitch, E.K. 1992. *Numerical definition of groundfish assemblages caught off the coasts of Oregon and Washington using commercial fishing strategies*. Can. J. Fish. Aquat. Sci. 49: 2648-2656.
- Rue, H., Martino, S. and Chopin, N. 2009. *Approximate Bayesian inference for latent Gaussian models by using INLA*. J. Royal Stat. Soc. B. 71: 319-392.
- Sanchirico, J.N., Holland, D., Quigley, K., and Fina, M. 2006. *Catch-quota balancing in multispecies individual fishing quotas*. Mar. Pol. 30: 767-785.
- SAS Institute Inc. 1999. *SAS/STAT® User's Guide, Version 8. Chapter 23. The CLUSTER Procedure*. Cary, NC: USA.
- SAS Institute Inc. 2008. *SAS/STAT® 9.2 User's Guide. The GENMOD Procedure*. Cary, NC: USA.
- Smith, T.D. 2007. *Scaling fisheries: the science of measuring the effects of fishing, 1855-1955*. Cambridge University Press, Cambridge, U.K.
- Snedecor, G.W. and Cochran, W.G., 1980. *Statistical Methods*, Seventh Edition. The Iowa State University Press, Ames, IA, USA, 507 pp.
- Stephens, A. and MacCall, A. 2004. *A multispecies approach to subsetting logbook data for purposes of estimating CPUE*. Fish. Res. 70: 299-310.
- Steven, S. and Die, D. 2006. *Using an agent based, object-oriented model to test the performance of catch per unit of effort as an estimator of fish abundance in small-scale, multispecies fisheries*. Proceedings of the 59th Gulf and Caribbean Fisheries Institute.
- Su, N.-J., Yeh, S.-Z., Sun, C.-L., Punt, A.E., Chen, Y. and Wang, S.-P. 2008. *Standardizing catch and effort data of the Taiwanese distant-water longline fishery in the western and central Pacific Ocean for Bigeye Tuna, Thunnus obesus*. Fish. Res. 90:235-246.
- Suzuki Z., Warashina, Y. and Kishida, M. 1977. *The comparison of catches by regular and deep tuna longline gears in the western and central equatorial Pacific*. Bull. Far Seas Fish. Res. Lab. 15: 51-89.
- Tascheri, R., Saavedra-Nievas, J.C. and Roa-Ureta, R. 2010. *Statistical models to standardise catch rates in the multispecies trawl fishery for Patagonian grenadier (Macruronus magellanicus) off Southern Chile*. Fish. Res. 105: 200-214.
- Thorson, J.T., Shelton, A.O., Ward, E.J. and Skaug, H.J. 2015. *Geostatistical delta-generalised linear mixed models improve precision for estimated abundance indices for West Coast groundfishes*. ICES J. Mar. Sci. 72: 1297-1310.
- Thorson, J.T., Fonner, R., Haltuch, M.A., Ono, K. and Winker, H. 2016. *Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data*. Can. J. Fish. Aquat. Sci. 73: 1-14.

- Turing, A.M. 1950. *Computer machinery and intelligence*. Mind LIX (236), 433-460.
- Venables, W.N. and Dichmont, C.M. 2004. *A generalised linear model for catch allocation: an example from Australia's Northern Prawn Fishery*. Fish. Res. 70: 409-426.
- Walters, C. 2003. *Folly and fantasy in the analysis of spatial catch rate data*. Can. J. Fish. Aquat. Sci. 60: 1433-1436.
- Ward, J.H., 1963. *Hierarchical grouping to optimize an objective function*. J. Am. Stat. Assoc. 58: 236-244.
- Ward, P.J. and Myers, R.A. 2006. *Do habitat models accurately predict the depth distribution of pelagic fishes?* Fish. Oceanogr. 15: 60-66.
- Ward, P., Lawrence, E., Darbyshire, R., Hindmarsh, S. 2008. *Large-scale experiment shows that nylon leaders reduce shark bycatch and benefit pelagic longline fishers*. Fish. Res. 90: 100-108.
- Watson, J.W., Epperly, S.P., Shah, A.K., Foster, D.G., 2005. *Fishing methods to reduce sea turtle mortality associated with pelagic longlines*. Can. J. Fish. Aquat. Sci. 62: 965-981.
- Watters, G. and Deriso, R. 2000. *Catch per unit of effort of Bigeye Tuna: a new analysis with regression trees and simulated annealing*. Bull. Inter-Amer. Trop. Tuna Comm. 21: 527-571.
- Winker, H., Kerwath, S.E. A.E. and Attwood, C.G. 2013. *Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery*. Fish. Res. 139: 118-131.
- Winker, H., Kerwath, S.E. and Attwood, C.G. 2014. *Proof of concept for a novel procedure to standardize multispecies catch and effort data*. Fish. Res. 155: 149-159.
- Wood, S. 2006. *Generalized Additive Models: an introduction in R*. CRC Press.
- Zhang, Z., & Holmes, J. 2009. *Generalized linear Bayesian models for standardization of CPUE with incorporation of spatial-temporal variations*. Working paper SA-WP-07 presented to the 5th Regular Session of the Scientific Committee for the Western and Central Pacific Fisheries Commission, held 10-21 August 2009, Port Vila, Vanuatu.
- Zhou, S., Vance, D.J., Dichmont, C.M., Burrridge, C. Y., and Toscas, P.J. 2008. *Estimating prawn abundance and catchability from catch-effort data: Comparison of fixed and random effects models using maximum likelihood and hierarchical Bayesian methods*. Mar. Fresh. Res. 59: 1-9.
- Zhou, S., Punt, A.E., Deng, R. and J. Bishop. 2011a. *Estimating multifleet catchability coefficients and natural mortality from fishery catch and effort data: comparison of Bayesian state-space and observation error models*. Can. J. Fish. Aquat. Sci. 68: 1171-118.
- Zhou, S., Punt, A.E., Deng, R. Kienzle, M. and Rochester, W. 2011b. *Bayesian fishable biomass dynamics models incorporating fished area and relative fish density*. Can. J. Fish. Aquat. Sci. 68: 1603-1614

Appendix A. Annual longline catch by species in the ETBF

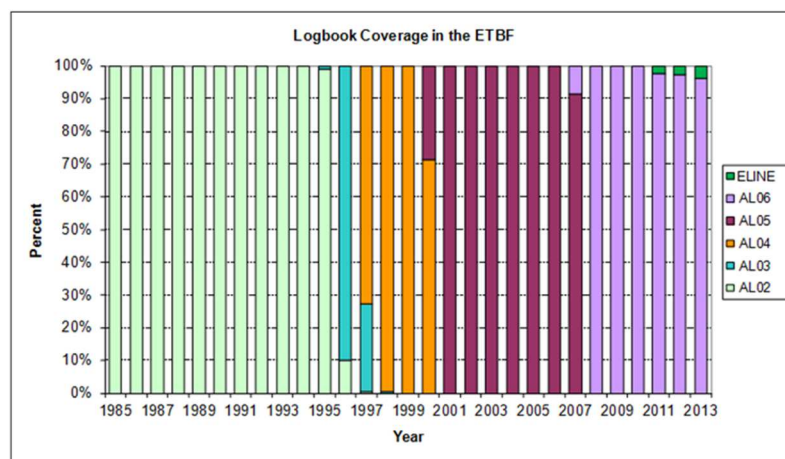
1. Introduction

The ETBF has undergone several periods of development and associated changes in targeting practices since the advent of the logbook program in the mid-1980s. For example, the fishery largely targeted only Yellowfin Tuna, and to some extent Bigeye Tuna, until the mid-1990s at which time a component of the fleet switched to targeting Broadbill Swordfish. The catches of Striped Marlin also increased considerably through the 1990s, such that by the year 2000 there were four principal target species in the fishery. The size of the fishery also changed significantly throughout the 1990s, with the effort increasing from 1.1 million hooks in 1990 to 9.6 million hooks in 2000 and the spatial extent of the fishery increasing by more than 2.5 times over this period. Effort peaked in 2003 when 12.75 million hooks were deployed and the spatial extent of the fishery reached 273 1x1-degree squares.

With the advent of lower catch rates and poor economic returns throughout the early to mid-2000s a number of vessels left the fishery and both effort and catches declined. The targeting of Albacore Tuna, and its subsequent addition as a primary target species, provided some financial assistance to the fishery with the catch of Albacore Tuna in 2006 and 2007 being the largest of the now five principal target species. With net economic returns to the fishery remaining low a government-based restructuring of the fishery in 2007 saw the number of active vessels remaining in the fishery decline to around 50 during 2008 when around 8 million hooks were deployed. Effort increased to around 8.9 million hooks in 2009 but has again declined in more recent years with around 6.8 million hooks being deployed between 2011 and 2013. The number of vessels active in the fishery (some only briefly) during 2013 was 41 (down from 49 in 2011). Total allowable catch quotas (based on individual transferrable quotas) for the five principal target species were introduced into the fishery in March 2011.

This brief history indicates that there have been a number of significant changes in the operation of the ETBF including changes in the range of species targeted. Commensurate with these changes has been the changes in the logbook used in the fishery. Since the introduction of the AL02 logbook in the mid-1980s there has been a series of four updates since this time (Figure A.1). With each new logbook there have been changes in the number of the species

Figure A.1: Annual logbook coverage (as a percentage of sets) in the ETBF. Note: ELINE refers to an electronic logbook.



reported in logbooks and this in turn has resulted in changes in the ability to report the catch by species in the fishery.

In this document the annual reported catch by species is provided for the main species caught by longline vessels operating in the ETBF over the periods 1985 to 2013. These species cover 72 of the 136 species reported caught during that period and account for 99% of all fish caught. A complete listing of the catch of all species reported in logbooks for the ETBF is provided in Appendix A.

The catches reported by species are listed by the following groups:

Tunas	(6 species),
Billfish	(6 species),
Sharks	(10 individual species, 2 species groups, 1 group of 17 species, and a single group of unidentified sharks),
Byproduct	(6 species),
Bycatch	(6 species),
Other Mackerels	(5 species in a single group), and
Skates and Rays	(1 group of stingrays, 1 group of manta rays, and a combined group).

Catch is reported by both the retained catch and the discarded catch reported in logbooks. Whilst the accuracy of the each catch component remains to be verified, it can perhaps be assumed that the retained catch is more accurately reported than the discarded catch.

It is important to note that not all species have been reported in each logbook and a listing of the number of fish caught by species in each logbook is provided in Appendix B. There have also been some changes in the identifying code used on logbooks for several species. For example, the code SPC_ID=146 was used to identify Spearfish on the AL02 logbook, while the code SPC_ID=53 has been used to identify Short-billed Spearfish on all other logbooks. Also, Mako Sharks were identified by the single code SPC_ID=138 on the AL02 logbook, while the codes SPC_ID=3 and 63 have been used to identify Short-finned Makos and Long-finned Makos respectively on all other logbooks. Similarly, the code SPC_ID=184 was used to identify Smoothed Hammerheads only on the AL05 logbook, the code SPC_ID=15 was used to identify Scalloped Hammerheads on the logbooks AL02-AL05, while the code SPC_ID=360 has been used to identified undifferentiated Hammerhead Sharks on the AL05, ALO06 and ELINE logbooks.

Finally, there may also have been changes in the manner that several species have been reported on logbooks. For example, large quantities of Mackerel (*Scomber scombrus*) have only been reported on the AL05 logbook (as SPC_ID=47) while large quantities of Snake Mackerel (*Gempylus serpens*) have only been reported on the AL06 logbook (as SPC_ID=362).

Given the above comments, some care is required in assuming that the reporting of any particular species (other than the principal catch species) has been continuous across the period reported here. Given this situation, some species may need to be grouped to more accurately reflect the nature of the species groups caught on particular longline sets.

2. SPECIES CODES

TUNAS

CODE	SPC_COMMON_NAME
YFT	Yellowfin Tuna
SKJ	Skipjack Tuna
SBT	Southern Bluefin Tuna
ALB	Albacore Tuna
BET	Bigeye Tuna
NBT	Northern Bluefin Tuna

BILLFISH

CODE	SPC_COMMON_NAME
BBL	Broad Billed Swordfish
STM	Striped Marlin
BUM	Blue Marlin
SAF	Indo-Pacific Sailfish
BAM	Black Marlin
SBS	Shortbilled Spearfish

SHARKS - Individual species

CODE	SPC_COMMON_NAME
CSH	Crocodile Shark
POR	Porbeagle
TSR	Thresher Shark
BRO	Bronze Whaler
DSK	Dusky Shark
BSH	Blue Shark
SKS	Silky Shark
TIG	Tiger Shark
OCS	Oceanic Whitetip Shark
TIP	Blacktip sharks

SHARKS - Grouped identified species

CODE	SPC_COMMON_NAME
SHK	White Shark
SHK	School Shark
SHK	Cookie-cutter Shark
SHK	Roughskin Shark
SHK	Saw Shark
SHK	Broadnose Sevengill Shark
SHK	Sandbar Shark
SHK	Bull Shark
SHK	Australian Black Shark
SHK	Australian Angel Shark
SHK	Ogilbys Ghost Shark
SHK	Australian blacktip shark
SHK	Sorrah shark
SHK	Whaler and weasel sharks
SHK	Platypus shark
SHK	whiskery shark
SHK	Grey reef shark

SHARKS - Grouped by species type

CODE	SPC_COMMON_NAME
MAK	Shortfin Mako
MAK	Longfin Mako
MAK	Mako shark species
SPN	Scalloped Hammerhead
SPN	Smoothed hammerhead
SPN	Hammerhead sharks

SHARKS - Unidentified group

CODE	SPC_COMMON_NAME
SHO	Shark "Other"
SHO	other sharks (AL02)
SHO	Sharks (other)

BYPRODUCT - Individual species

CODE	SPC_COMMON_NAME
OPA	Moonfish
DOL	Dolphinfish
POA	Ray's Bream
OIL	Oilfish
WAH	Wahoo
BOF	Black Oilfish

BYCATCH - Individual species

CODE	SPC_COMMON_NAME
LAN	Lancet fish
BAR	Barracouta
MAC	Mackerel
SUN	Ocean Sunfish
MOP	Short Sunfish
GES	Snake Mackerel

BYPRODUCT - Grouped mackerel species

CODE	SPC_COMMON_NAME
MCK	Butterfly Mackerel
MCK	Jack Mackerel
MCK	Blue Mackerel
MCK	Spanish Mackerel
MCK	Frigate mackerel

BYCATCH - Skates and rays

CODE	SPC_COMMON_NAME
STR	Stingray
RAY	Manta Ray
RAI	Skates & rays, unspecified

3. TUNAS

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
38	YFT	Yellowfin Tuna	Thunnus albacares	146151
39	SKJ	Skipjack Tuna	Katsuwonus pelamis	7725
40	SBT	Southern Bluefin Tuna	Thunnus maccoyii	8314
41	ALB	Albacore Tuna	Thunnus alalunga	102852
42	BET	Bigeye Tuna	Thunnus obesus	90059
46	NBT	Northern Bluefin Tuna	Thunnus thynnus	476

Note: N_FOPS = Number of fishing operations this species caught

Figure A.2: Logbook recorded annual catch (number of fish retained and discarded) of tuna species in the ETBF.



4. BILLFISH

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
48	BBL	Broad Billed Swordfish	<i>Xiphias gladius</i>	92432
49	STM	Striped Marlin	<i>Tetrapturus audax</i>	50996
50	BUM	Blue Marlin	<i>Makaira mazara</i>	5620
51	SAF	Indo-Pacific Sailfish	<i>Istiophorus platypterus</i>	1827
52	BAM	Black Marlin	<i>Makaira indica</i>	3613
53	SBS	Shortbilled Spearfish	<i>Tetrapturus angustirostris</i>	7666
146	SBS	Spearfish (AL02 only)	<i>Tetrapturus angustirostris</i>	587

Note: N_FOPS = Number of fishing operations this species caught

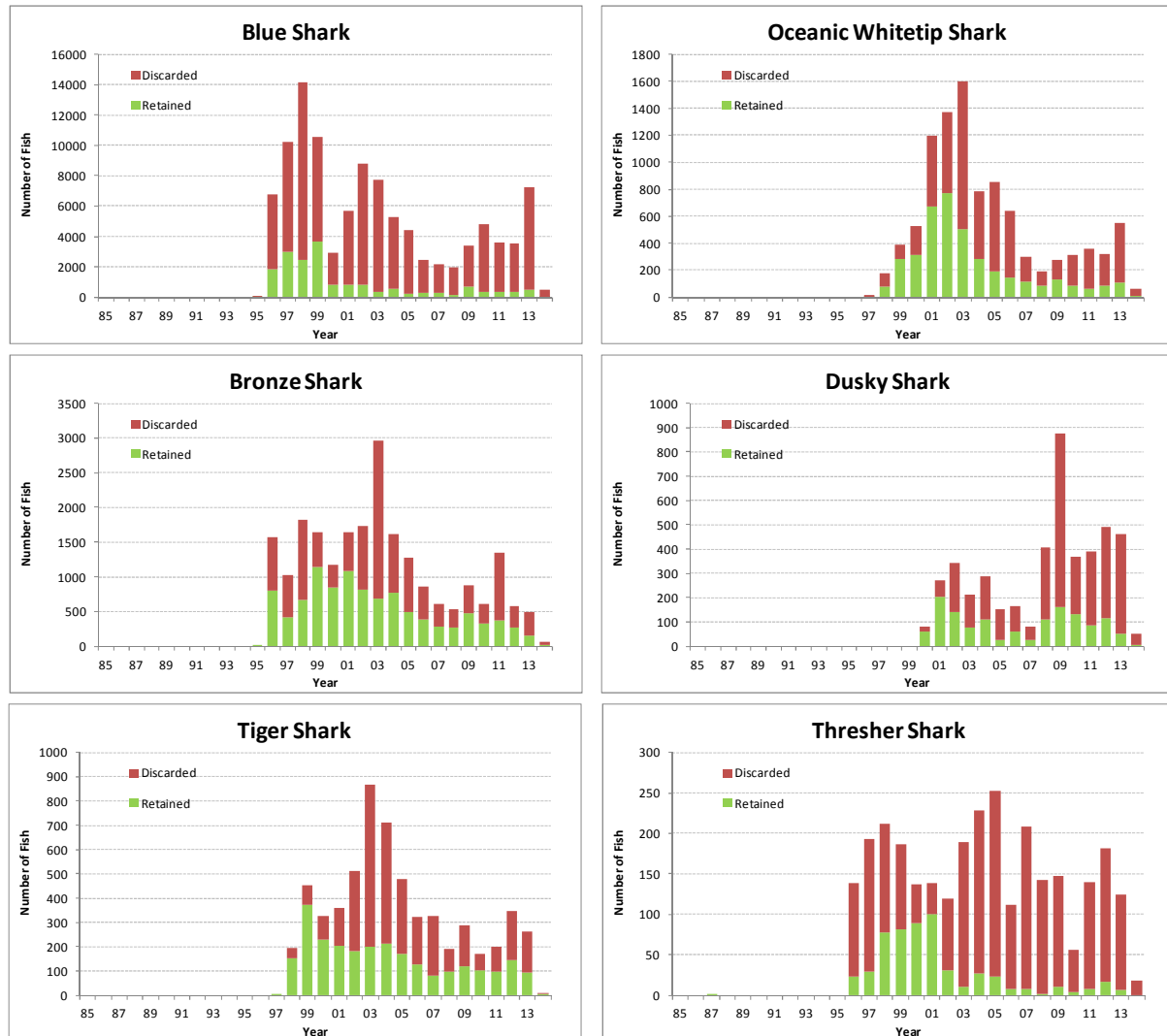
Figure A.3: Logbook recorded annual catch (number of fish retained and discarded) of billfish species in the ETBF.



5. SHARKS

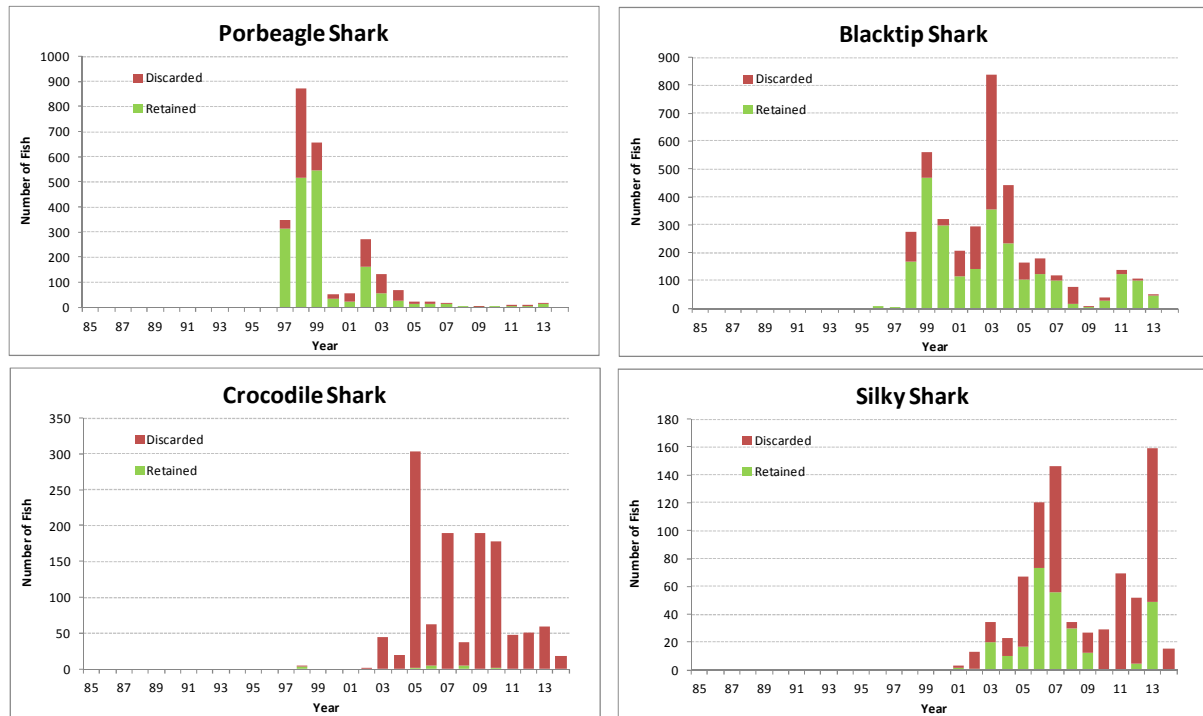
SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
2	CSH	Crocodile Shark	Pseudocarcharias kamoharai	280
5	POR	Porbeagle	Lamna nasus	653
6	TSR	Thresher Shark	Alopias vulpinus	1719
8	BRO	Bronze Whaler	Carcharhinus brachyurus	8973
9	DSK	Dusky Shark	Carcharhinus obscurus	1815
10	BSH	Blue Shark	Prionace glauca	24189
11	SKS	Silky Shark	Carcharhinus falciformis	347
12	TIG	Tiger Shark	Galeocerdo cuvier	3950
13	OCS	Oceanic Whitetip Shark	Carcharhinus longimanus	5982
14	TIP	Blacktip sharks	Carcharhinus species	1479

Figure A.4: Logbook recorded annual catch (number of fish retained and discarded) of shark species in the ETBF.



5. SHARKS (cont'd)

Figure A.4 (cont'd): Logbook recorded annual catch (number of fish retained and discarded) of shark species in the ETBF.

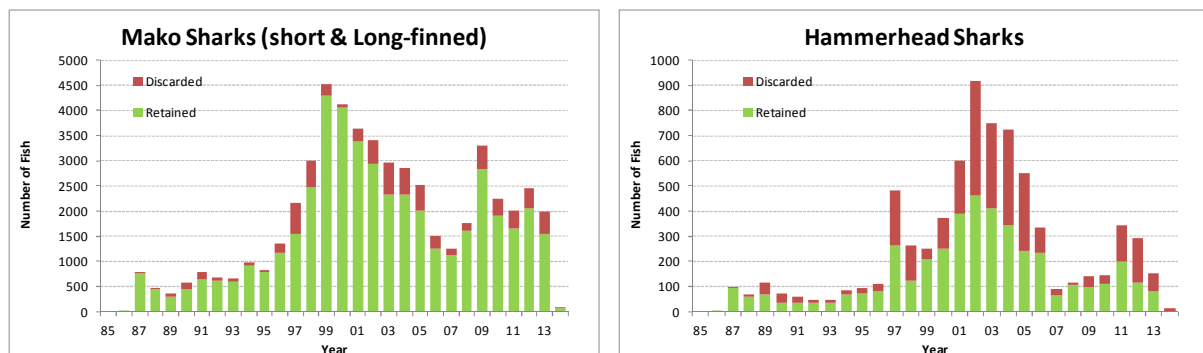


SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
3	MAK	Shortfin Mako	Isurus oxyrinchus	28858
63	MAK	Longfin Mako	Isurus paucus	51
138	MAK	Mako shark species	Lamnidae spp	3275
15	SPN	Scalloped Hammerhead	Sphyrna lewini	2477
184	SPN	Smoothed hammerhead	Sphyrna zygaena	78
360	SPN	Hammerhead sharks	Sphyrnidae - undifferentiated	730

Note: SPC_ID=3, 63, 138 grouped as Mako Sharks

Note: SPC_ID=15, 184, 360 grouped as Hammerhead Sharks

Figure A.4 (cont'd): Logbook recorded annual catch (number of fish retained and discarded) of shark species in the ETBF.



5. SHARKS (cont'd)

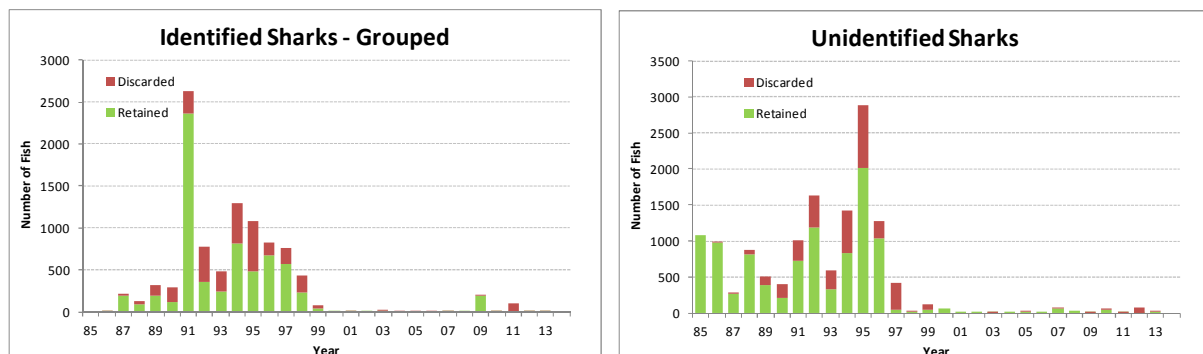
SHK: Identified Sharks (grouped) includes the following species:

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
4	SHK	White Shark	Carcharodon carcharias	4
7	SHK	School Shark	Galeorhinus galeus	480
17	SHK	Cookie-cutter Shark	Isistius brasiliensis	9
18	SHK	Roughskin Shark	Centroscyrnus and Deania.	4
19	SHK	Saw Shark	Pristiophorus	33
62	SHK	Broadnose Sevengill Shark	Notorynchus cepedianus	206
64	SHK	Sandbar Shark	Carcharhinus plumbeus	11
65	SHK	Bull Shark	Carcharhinus leucas	19
67	SHK	Australian Black Shark	Dalatis Licha	144
70	SHK	Australian Angel Shark	Squatina australis	96
71	SHK	Ogilbys Ghost Shark	Hydrolagus ogilbyi	453
93	SHK	Australian blacktip shark	Carcharhinus tilstoni	6
94	SHK	Sorrah shark	Carcharhinus sorrah	36
95	SHK	Whaler and weasel sharks	Caracharhinidae	1764
96	SHK	Platypus shark	Deania calcea	283
178	SHK	whiskery shark	Furgaleus macki	89
182	SHK	Grey reef shark	Carcharhinus amblyrhynchos	2

SHO: Unidentified Sharks includes the following:

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
58	SHO	Shark "Other"	Family "Shark" other	248
134	SHO	other sharks (AL02)	Carcharhinidae, Hemigaleidae	2989
358	SHO	Sharks (other)	sharks - other	105

Figure A.4 (cont'd): Logbook recorded annual catch (number of fish retained and discarded) of shark species in the ETBF.

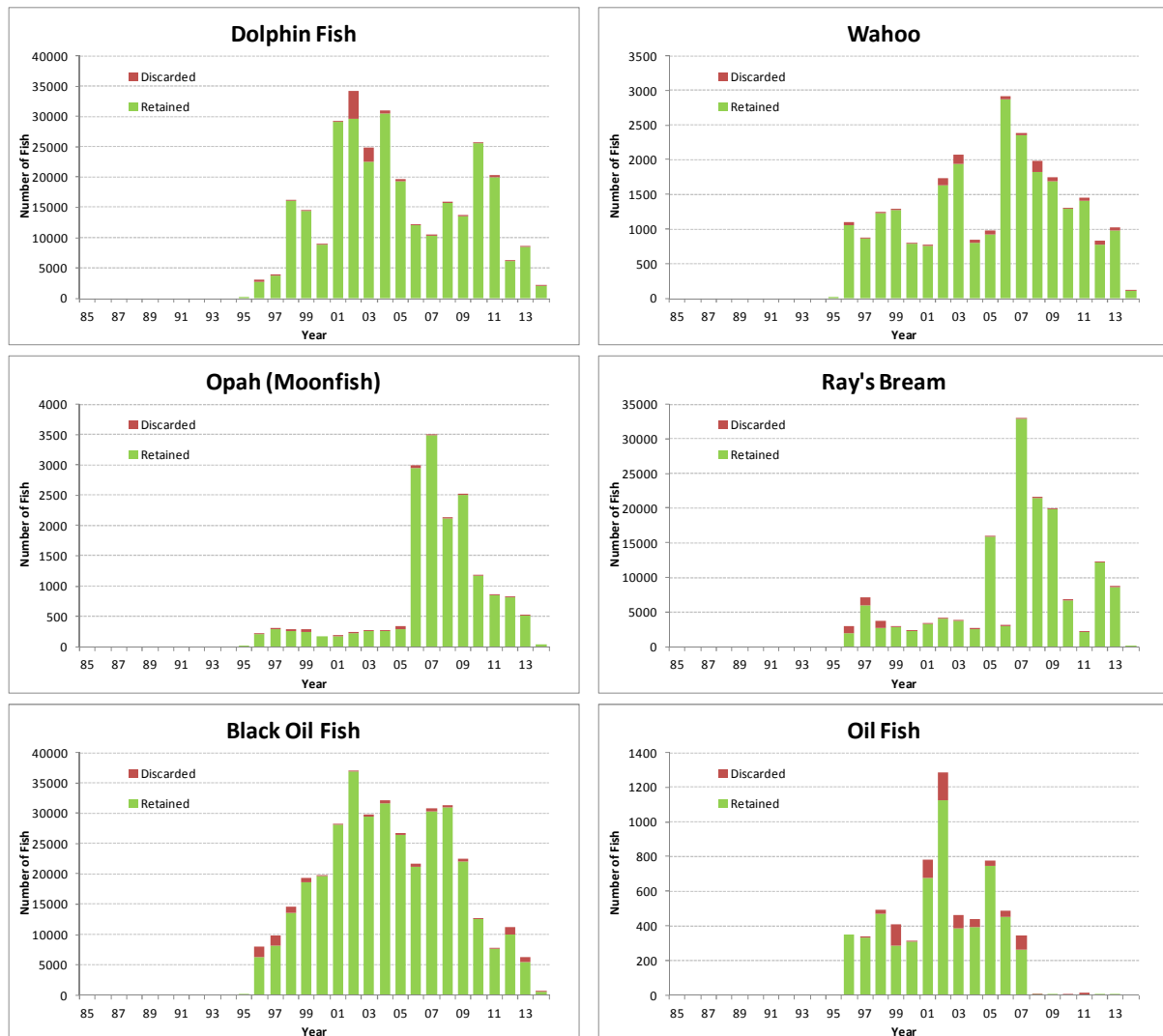


6. BYPRODUCT SPECIES

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
23	OPA	Moonfish	Lampris guttatus	6640
29	DOL	Dolphinfish	Coryphaena hippurus	60560
30	POA	Ray's Bream	Brama brama	8559
35	OIL	Oilfish	Ruvettus pretiosus	1922
45	WAH	Wahoo	Acanthocybium solandri	13789
36	BOF	Black Oilfish	Lepidocybium flavobrunneum	12053
55	BOF	Rudderfish	Centrolophus niger	50831

Note: Rudderfish are grouped with Black Oilfish (due to misidentification on logbook)

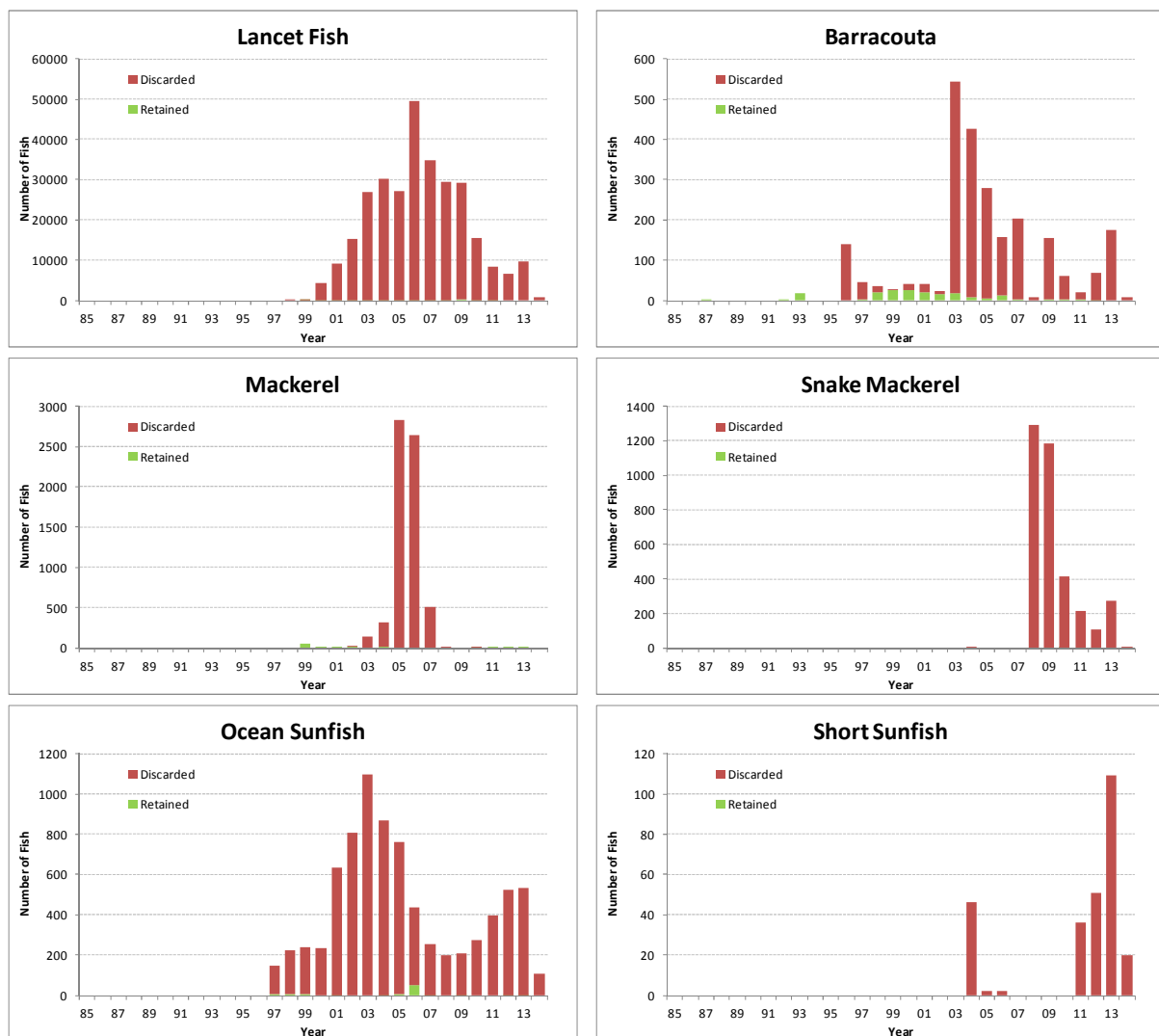
Figure A.5: Logbook recorded annual catch (number of fish retained and discarded) of byproduct species in the ETBF.



7. BYCATCH SPECIES

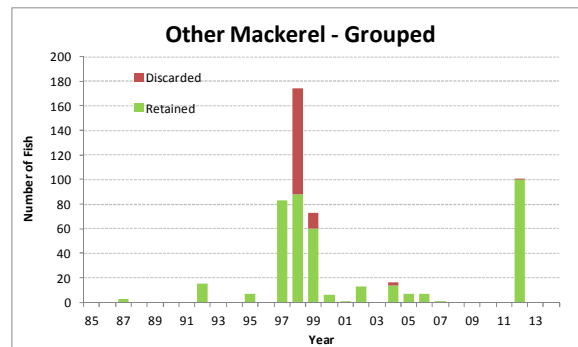
SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
22	LAN	Lancet fish	Alepisauridae	22427
34	BAR	Barracouta	Thyrsites atun	689
47	MAC	Mackerel	Scomber scombrus	726
56	SUN	Ocean Sunfish	Mola mola	3481
361	MOP	Short Sunfish	Mola ramsayi	119
362	GES	Snake Mackerel	Gempylus serpens	341

Figure A.6: Logbook recorded annual catch (number of fish retained and discarded) of bycatch species in the ETBF.



8. OTHER MACKERELS (Grouped)

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
43	MCK	Butterfly Mackerel	Gasterochisma melampus	165
75	MCK	Jack Mackerel	Trachurus declivis	6
83	MCK	Blue Mackerel	Scomber australasicus	7
84	MCK	Spanish Mackerel	Scomberomorus Commerson	11
115	MCK	Frigate mackerel	Auxis thazard	2



9. SKATES & RAYS

SPC_ID	CODE	SPC_COMMON_NAME	SPC_NAME	N_FOPS
21	STR	Stingray	Dasyatidae "family"	306
91	RAY	Manta Ray	Manta birostris	504
359	RAI	Skates & rays, unspecifie	Skates & rays, unspecified	697

Figure A.7: Logbook recorded annual catch (number of fish retained and discarded) of skate and ray species in the ETBF.

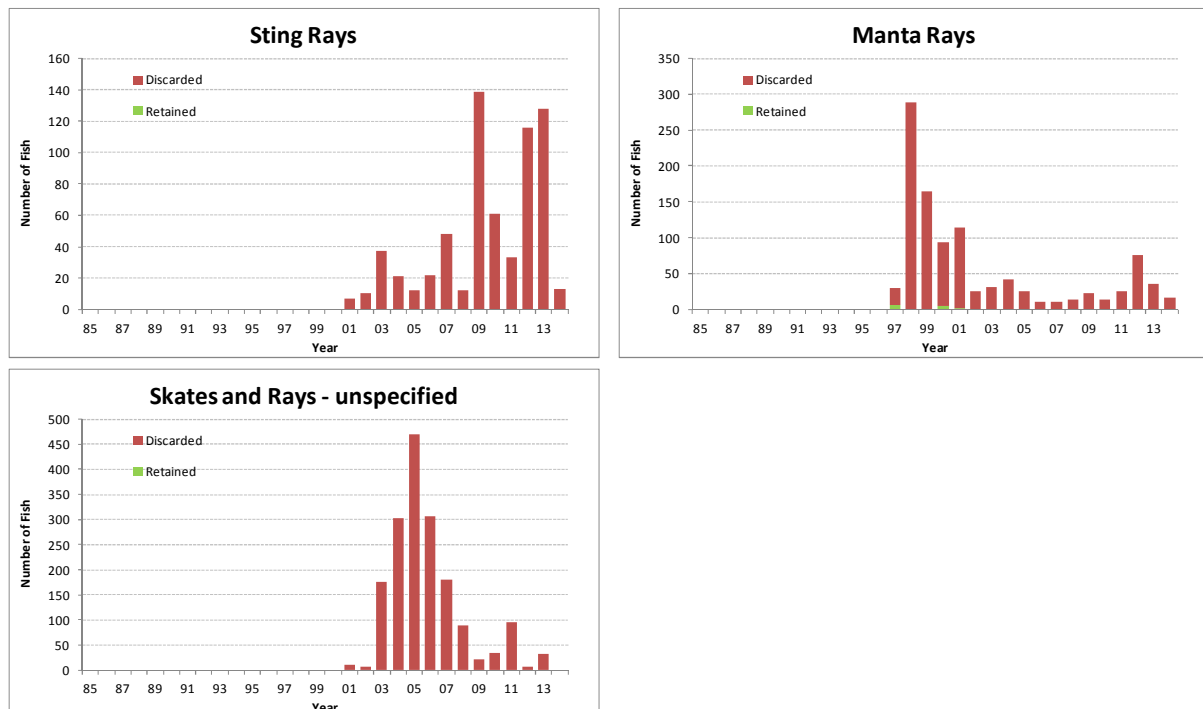


Table A.1: Complete listing of the logbook reported catch of all species reported by longline vessels operating in the ETBF. (Note, the last column indicates whether the species is reported in main document)

No.	SPC_ID	SPC	SPC_COMMON_NAME	SETS	RETAINED	DISCARDED	TOTAL	Report
1	38	YFT	Yellowfin Tuna	175569	1300099	51873	1351972	1
2	41	ALB	Albacore Tuna	114333	1315727	26802	1342529	1
3	48	SWO	Broad Billed Swordfish	95800	467613	10701	478314	1
4	42	BET	Bigeye Tuna	93493	388339	19085	407424	1
5	29	DOL	Dolphinfish	60475	289843	9549	299392	1
6	22	LAN	Lancet fish	22424	1847	296256	298103	1
7	55	CEO	Rudderfish	50698	272338	8928	281266	1
8	30	POA	Ray's Bream	8549	149978	3688	153666	1
9	10	BSH	Blue Shark	24199	18633	88694	107327	1
10	49	MLS	Striped Marlin	53687	99056	3087	102143	1
11	40	SBF	Southern Bluefin Tuna	11025	92247	8775	101022	1
12	36	BOF	Black Oilfish	12053	82121	2413	84534	1
13	3	SMA	Shortfin Mako	28822	40427	6608	47035	1
14	59	MIX	OTHER	12640	36642	2070	38712	1
15	39	SKJ	Skipjack Tuna	9612	35725	1896	37621	1
16	45	WAH	Wahoo	13706	24515	919	25434	1
17	8	BRO	Bronze Whaler	8915	10595	12201	22796	1
18	23	OPA	Moonfish	6640	16503	265	16768	1
19	134		other sharks (AL02)	5573	11823	3356	15179	1
20	53	SBS	Shortbilled Spearfish	7666	11557	1030	12587	1
21	50	BLM	Blue Marlin	6886	1362	9361	10723	1
22	13	OCS	Oceanic Whitetip Shark	5980	3907	6002	9909	1
23	95		whaler and weasel sharks	3245	6912	2421	9333	1
24	52	BKM	Black Marlin	4292	13	8087	8100	1
25	56	SUN	Ocean Sunfish	3481	130	7852	7982	1
26	47	MAC	Mackerel	723	140	6468	6608	1
27	138		Mako shark species	6352	6010	568	6578	1
28	35	OIL	Oilfish	1900	5741	707	6448	1
29	12	TSH	Tiger Shark	3937	2414	3431	5845	1
30	15	SPL	Scalloped Hammerhead	2751	3534	2226	5760	1
31	9	DSK	Dusky Shark	1812	1370	3270	4640	1
32	14	TIP	Blacktip sharks	1479	2266	1386	3652	1
33	51	SAF	Indo-Pacific Sailfish	2037	1838	1693	3531	1
34	362	GES	Snake Mackerel	341	1	3495	3496	1
35	6	TSR	Thresher Shark	1717	559	2365	2924	1
36	5	POR	Porbeagle	645	1722	806	2528	1
37	34	BAR	Barracouta	689	179	2300	2479	1
38	359	RAI	Skates & rays	697	0	1732	1732	1
39	360	SPN	hammerhead sharks	730	1059	643	1702	1
40	146		Spearfish	1143	1097	125	1222	1
41	2	CSH	Crocodile Shark	280	15	1191	1206	1
42	91	RAY	Manta Ray	504	42	1026	1068	1
43	11	SKS	Silky Shark	347	274	516	790	1
44	21	STR	Stingray	306	0	659	659	1
45	71	SHT	Ogilbys Ghost Shark	448	439	190	629	1
46	96		platypus shark	276	564	64	628	1
47	46	NBF	Northern Bluefin Tuna	478	496	22	518	1
48	73	GRE	Blue Grenadier	98	228	151	379	0
49	58	SHO	Shark "Other"	215	161	213	374	1
50	61	MAR	Marlin/Sailfish	115	2	367	369	0
51	43	MAB	Butterfly Mackerel	165	235	99	334	1

Table A.1: (cont'd)

No.	SPC_ID	SPC	SPC_COMMON_NAME	SETS	RETAINED	DISCARDED	TOTAL	n Report
52	67	SLH	Australian Black Shark	144	170	158	328	1
53	26	BKF	Black Kingfish	149	290	6	296	0
54	361	MOP	Short Sunfish	119	0	266	266	1
55	184		Smoothed hammerhead shark	78	86	179	265	1
56	102		Moonfish	181	236	3	239	1
57	7	SHS	School Shark	83	209	10	219	1
58	353		pufferfish	93	0	214	214	0
59	182		grey reef shark	2	193	0	193	1
60	358	SHO	Sharks (other)	105	68	114	182	1
61	44	ABO	Australian Bonito	35	163	3	166	0
62	70	SHA	Australian Angel Shark	65	68	97	165	1
63	25	OAR	Oarfish	80	75	48	123	0
64	37	SFF	Southern Frostfish	98	37	80	117	0
65	83	MAS	Blue Mackerel	7	105	2	107	1
66	178		whiskery shark	8	1	98	99	1
67	31	SNA	Snapper	21	93	2	95	0
68	27	YEK	Yellowtail Kingfish	35	79	8	87	0
69	80	BRE	Black Bream	19	83	0	83	0
70	63	LFM	Longfin Mako	51	36	16	52	1
71	28	RRR	Rainbow Runner	29	48	3	51	0
72	24	RIB	Dealfish	42	2	47	49	0
73	94		Sorrah shark	29	34	14	48	1
74	54	TBE	Blue Eye Trevalla	16	47	1	48	0
75	65	BUL	Bull Shark	19	16	27	43	1
76	78	BSP	Big-scale Pomfret	7	32	0	32	0
77	85	MKT	Eastern Little Tuna	5	27	0	27	0
78	64	SBH	Sandbar Shark	11	18	4	22	1
79	157	CAR	Cardinal Fish	20	12	8	20	0
80	114		Gemfish	57	2	17	19	0
81	87	DTT	Dogtooth Tuna	9	17	0	17	0
82	1	SHY	Grey Nurse	13	0	15	15	0
83	317		yellow-spotted boarfish	15	13	0	13	0
84	75	MAJ	Jack Mackerel	6	12	1	13	1
85	367		tunas	5	9	4	13	0
86	79	JOB	Jobfish	3	12	1	13	0
87	84	SNM	Spanish Mackerel	11	11	0	11	1
88	159	JOR	Rosy Jobfish / King Snapper	2	11	0	11	0
89	17	CCS	Cookie-cutter Shark	9	4	6	10	1
90	227		"true" dories	1	10	0	10	1
91	18	SRH	Roughskin Shark	4	9	0	9	1
92	62	SHL	Broadnose Sevengill Shark	6	6	2	8	1
93	93		Australian blacktip shark	6	8	0	8	1
94	115		Frigate mackerel	2	8	0	8	1
95	16	SHD	Dogfish	21	2	5	7	0
96	103		Reef ocean perch	75	6	0	6	0
97	111		jackass morwong	53	6	0	6	0
98	354		porcupine fish	5	0	6	6	0
99	32	MOO	Moonlighter	3	6	0	6	0
100	60	TUN	Tuna/Mackerel	1	6	0	6	0
101	155		trevallies and jacks	3	4	1	5	0
102	4	JAW	White Shark	4	0	4	4	1

Table A.1: (cont'd)

No.	SPC_ID	SPC	SPC_COMMON_NAME	SETS	RETAINED	DISCARDED	TOTAL	n Report
103	82	TRU	Bastard Trumpeter	4	4	0	4	0
104	88	LUV	Luvaru	4	4	0	4	0
105	250		eastern orange perch	3	4	0	4	0
106	156	MUL	Jewfish/Mulloway	1	0	4	4	0
107	264		coral trout	1	4	0	4	0
108	86	LTT	Long Tailed Tuna	3	3	0	3	0
109	363	BRA	pomfrets	3	3	0	3	0
110	290		green jobfish	2	3	0	3	0
111	209		cod	1	3	0	3	0
112	90	SQO	Squid	103	2	0	2	0
113	170		cow shark	2	1	1	2	0
114	216		garfishes	1	0	2	2	0
115	343		large-headed hairtail	1	2	0	2	0
116	368		tonquefish & soleidae	1	0	2	2	0
117	89	TRS	Spotted Warehou	55	0	1	1	0
118	351		triggerfishes and leatherjackets	25	1	0	1	0
119	33	BOA	Boarfish	9	1	0	1	0
120	265		bar rock cod	4	0	1	1	0
121	81	REM	Red Mullet/Blue-lined Goatfish	3	1	0	1	0
122	312		silver bream	2	0	1	1	0
123	68	DFS	White-Spotted Dogfish	1	1	0	1	0
124	72	PIK	Common Pike Eel	1	1	0	1	0
125	74	GRB	Hapuku and Bass Groper-NSW	1	1	0	1	0
126	76	SAM	Samsonfish	1	1	0	1	0
127	137		perches, basses, rock cods	1	1	0	1	0
128	158	RAE	Southern Eagle Ray	1	0	1	1	0
129	277		bigeye trevally	1	1	0	1	0
130	280		black pomfret	1	1	0	1	0
131	310		sea breams	1	1	0	1	0
132	345		medusa fish	1	0	1	1	0
133	356		squid	1	1	0	1	0
134	380		wedgefishes	1	1	0	1	0
135	399		Saddleback Snapper	1	1	0	1	0
136	425	WHO	whales	1	0	1	1	0
				Total	4710775	629117	5339892	136
				In Report	4709416	628115	5337531	72
					99.97%	99.84%	99.96%	52.94%

Table A.2: Listing by logbook of the reported catch of all species by longline vessels operating in the ETBF (Note, the last column indicates whether the species is reported in main document)

No.	SPC ID	SPC	SPC_COMMON_NAME	N_FISH	ALO2	ALO3	ALO4	ALO5	ALO6	ELINE	Report
1	38	YFT	Yellowfin Tuna	1351972	276810	84041	184616	527284	273399	5822	1
2	41	ALB	Albacore Tuna	1342529	145909	38435	138753	587110	429508	2814	1
3	48	SWO	Broad Billed Swordfish	478314	6213	14083	124253	205137	126643	1985	1
4	42	BET	Bigeye Tuna	407424	12200	12366	89959	194340	97589	970	1
5	29	DOL	Dolphinfish	299392	0	4206	37732	163036	92180	2238	1
6	22	LAN	Lancet fish	298103	0	0	952	195997	100145	1009	1
7	55	CEO	Rudderfish	281266	0	9283	56007	149769	65754	453	1
8	30	POA	Ray's Bream	153666	0	2645	14379	65604	70590	448	1
9	10	BSH	Blue Shark	107327	0	9461	35222	37436	24738	470	1
10	49	MLS	Striped Marlin	102143	5819	2943	24459	48591	20063	268	1
11	40	SBF	Southern Bluefin Tuna	101022	36596	8225	30473	5732	19996	0	1
12	36	BOF	Black Oilfish	84534	0	0	0	56331	28203	0	1
13	3	SMA	Shortfin Mako	47035	0	1714	12515	18802	13871	133	1
14	59	MIX	OTHER	38712	36841	887	454	380	150	0	1
15	39	SKJ	Skipjack Tuna	37621	10198	1784	3906	14161	7548	24	1
16	45	WAH	Wahoo	25434	0	1420	3676	11589	8648	101	1
17	8	BRO	Bronze Whaler	22796	0	2596	4429	11244	4527	0	1
18	23	OPA	Moonfish	16768	0	0	934	7151	8671	12	1
19	134		other sharks (ALO2)	15179	14472	550	0	8	149	0	1
20	53	SBS	Shortbilled Spearfish	12587	0	122	734	7618	4082	31	1
21	50	BLM	Blue Marlin	10723	1610	220	3	6908	1908	74	1
22	13	OCS	Oceanic Whitetip Shark	9909	0	0	720	7117	1974	98	1
23	95		whaler and weasel sharks	9333	9333	0	0	0	0	0	1
24	52	BKM	Black Marlin	8100	0	0	26	5538	2448	88	1
25	56	SUN	Ocean Sunfish	7982	0	0	817	4922	2243	0	1
26	47	MAC	Mackerel	6608	0	0	67	6481	60	0	1
27	138		Mako shark species	6578	6578	0	0	0	0	0	1
28	35	OIL	Oilfish	6448	0	464	1278	4681	15	10	1
29	12	TSH	Tiger Shark	5845	0	0	679	3528	1617	21	1
30	15	SPL	Scalloped Hammerhead	5760	837	214	1401	3308	0	0	1
31	9	DSK	Dusky Shark	4640	0	0	5	1568	2890	177	1
32	14	TIP	Blacktip sharks	3652	0	0	908	2328	412	4	1
33	51	SAF	Indo-Pacific Sailfish	3531	695	107	375	2090	263	1	1
34	362	GES	Snake Mackerel	3496	0	0	0	6	3330	160	1
35	6	TSR	Thresher Shark	2924	0	163	619	1329	732	81	1
36	5	POR	Porbeagle	2528	0	0	1866	625	37	0	1
37	34	BAR	Barracouta	2479	0	194	109	1663	462	51	1
38	359	RAI	Skates and rays	1732	0	0	0	1450	282	0	1
39	360	SPN	hammerhead sharks	1702	0	0	0	500	1081	121	1
40	146		Spearfish	1222	1222	0	0	0	0	0	1
41	2	CSH	Crocodile Shark	1206	0	0	4	622	425	155	1
42	91	RAY	Manta Ray	1068	0	0	560	304	177	27	1
43	11	SKS	Silky Shark	790	0	0	0	403	265	122	1
44	21	STR	Stingray	659	0	0	0	157	474	28	1
45	71	SHT	Ogilbys Ghost Shark	629	0	398	231	0	0	0	1
46	96		platypus shark	628	0	628	0	0	0	0	1
47	46	NBF	Northern Bluefin Tuna	518	3	0	101	368	46	0	1
48	73	GRE	Blue Grenadier	379	0	16	363	0	0	0	0
49	58	SHO	Shark "Other"	374	0	0	374	0	0	0	1
50	61	MAR	Marlin/Sailfish	369	2	0	0	0	367	0	0
51	43	MAB	Butterfly Mackerel	334	0	0	326	8	0	0	1

Table A.2: (cont'd)

No.	SPC_ID	SPC	SPC_COMMON_NAME	N_FISH	ALO2	ALO3	ALO4	ALO5	ALO6	ELINE	n Report
52	67	SLH	Australian Black Shark	328	0	0	328	0	0	0	1
53	26	BKF	Black Kingfish	296	0	30	61	188	16	1	0
54	361	MOP	Short Sunfish	266	0	0	0	50	0	216	1
55	184		Smoothed hammerhead	265	0	0	0	265	0	0	1
56	102		Moonfish	239	0	239	0	0	0	0	1
57	7	SHS	School Shark	219	0	88	125	3	3	0	1
58	353		pufferfish	214	0	0	0	17	186	11	0
59	182		grey reef shark	193	0	0	0	0	193	0	1
60	358	SHO	Sharks (other)	182	0	0	0	65	117	0	1
61	44	ABO	Australian Bonito	166	0	0	4	104	58	0	0
62	70	SHA	Australian Angel Shark	165	0	0	161	4	0	0	1
63	25	OAR	Oarfish	123	0	0	50	38	27	8	0
64	37	SFF	Southern Frostfish	117	0	1	31	71	14	0	0
65	83	MAS	Blue Mackerel	107	0	0	1	5	101	0	1
66	178		whiskery shark	99	0	0	0	0	99	0	1
67	31	SNA	Snapper	95	0	0	7	71	17	0	0
68	27	YEK	Yellowtail Kingfish	87	0	0	15	33	39	0	0
69	80	BRE	Black Bream	83	0	0	83	0	0	0	0
70	63	LFM	Longfin Mako	52	0	0	2	16	34	0	1
71	28	RRR	Rainbow Runner	51	0	0	21	20	6	4	0
72	24	RIB	Dealfish	49	0	0	4	32	6	7	0
73	94		Sorrah shark	48	0	48	0	0	0	0	1
74	54	TBE	Blue Eye Trevalla	48	0	3	0	1	44	0	0
75	65	BUL	Bull Shark	43	0	0	28	12	3	0	1
76	78	BSP	Big-scale Pomfret	32	0	0	32	0	0	0	0
77	85	MKT	Eastern Little Tuna	27	0	0	27	0	0	0	0
78	64	SBH	Sandbar Shark	22	0	0	0	4	18	0	1
79	157	CAR	Cardinal Fish	20	0	0	0	15	5	0	0
80	114		Gemfish	19	0	0	0	12	7	0	0
81	87	DTT	Dogtooth Tuna	17	0	0	8	7	2	0	0
82	1	SHY	Grey Nurse	15	0	0	0	15	0	0	0
83	317		yellow-spotted boarfish	13	0	0	0	13	0	0	0
84	75	MAJ	Jack Mackerel	13	0	0	0	13	0	0	1
85	367		tunas	13	0	0	0	0	9	4	0
86	79	JOB	Jobfish	13	0	0	2	11	0	0	0
87	84	SNM	Spanish Mackerel	11	0	0	0	11	0	0	1
88	159	JOR	Rosy Jobfish / King Snapper	11	0	0	0	3	8	0	0
89	17	CCS	Cookie-cutter Shark	10	0	0	1	8	1	0	1
90	227		"true" dories	10	0	0	0	0	10	0	1
91	18	SRH	Roughskin Shark	9	0	0	0	0	9	0	1
92	62	SHL	Broadnose Sevengill Shark	8	0	0	8	0	0	0	1
93	93		Australian blacktip shark	8	0	8	0	0	0	0	1
94	115		Frigate mackerel	8	0	0	0	8	0	0	1
95	16	SHD	Dogfish	7	0	0	7	0	0	0	0
96	103		Reef ocean perch	6	0	0	0	6	0	0	0
97	111		jackass morwong	6	0	6	0	0	0	0	0
98	354		porcupine fish	6	0	0	0	6	0	0	0
99	32	MOO	Moonlighter	6	0	0	2	4	0	0	0
100	60	TUN	Tuna/Mackerel	6	0	6	0	0	0	0	0
101	155		trevallies and jacks	5	0	0	0	1	4	0	0
102	4	JAW	White Shark	4	0	0	0	4	0	0	1

Table A.2: (cont'd)

No.	SPC_ID	SPC	SPC_COMMON_NAME	N_FISH	ALO2	ALO3	ALO4	ALO5	ALO6	ELINE	n Report
103	82	TRU	Bastard Trumpeter	4	0	2	2	0	0	0	0
104	88	LUV	Luvaru	4	0	0	2	1	1	0	0
105	250		eastern orange perch	4	0	0	0	4	0	0	0
106	156	MUL	Jewfish/Mulloway	4	0	0	0	4	0	0	0
107	264		coral trout	4	0	0	0	4	0	0	0
108	86	LTT	Long Tailed Tuna	3	2	0	0	0	1	0	0
109	363	BRA	pomfrets	3	0	0	0	2	1	0	0
110	290		green jobfish	3	0	0	0	3	0	0	0
111	209		cod	3	0	0	0	3	0	0	0
112	90	SQO	Squid	2	0	0	2	0	0	0	0
113	170		cow shark	2	0	0	0	2	0	0	0
114	216		garfishes	2	0	0	0	2	0	0	0
115	343		large-headed hairtail	2	0	0	0	2	0	0	0
116	368		tonguefish & soleidae	2	0	0	0	2	0	0	0
117	89	TRS	Spotted Warehou	1	0	0	1	0	0	0	0
118	351		triggerfishes and leatherjackets	1	0	0	0	0	1	0	0
119	33	BOA	Boarfish	1	0	0	0	0	1	0	0
120	265		bar rock cod	1	0	0	0	1	0	0	0
121	81	REM	Red Mullet/Blue-lined Goatfish	1	0	0	1	0	0	0	0
122	312		silver bream	1	0	0	0	1	0	0	0
123	68	DFS	White-Spotted Dogfish	1	0	0	1	0	0	0	0
124	72	PIK	Common Pike Eel	1	0	0	1	0	0	0	0
125	74	GRB	Hapuku and Bass Groper-NSW	1	0	0	1	0	0	0	0
126	76	SAM	Samsonfish	1	0	0	1	0	0	0	0
127	137		perches, basses, rock cods	1	0	0	0	0	0	1	0
128	158	RAE	Southern Eagle Ray	1	0	0	0	0	1	0	0
129	277		bigeye trevally	1	0	0	0	1	0	0	0
130	280		black pomfret	1	0	0	0	1	0	0	0
131	310		sea breams	1	0	0	0	1	0	0	0
132	345		medusa fish	1	0	0	0	0	0	1	0
133	356		squid	1	0	0	0	1	0	0	0
134	380		wedgefishes	1	0	0	0	0	1	0	0
135	399		Saddleback Snapper	1	0	0	0	0	1	0	0
136	425	WHO	whales	1	0	0	0	0	1	0	0

Appendix B. Spatial and temporal and distributions of CPUE in the ETBF

The following are shown in this attachment:

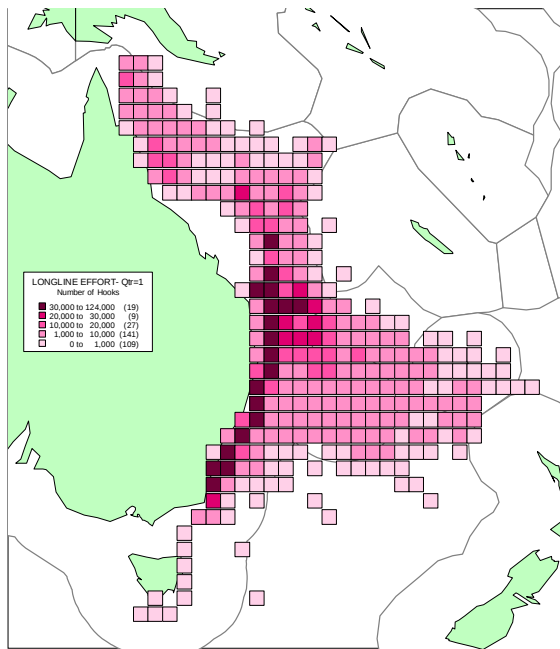
1. Spatial plots (1x1-degree) of aggregate effort and nominal CPUE over the years 2000-2013 for each quarter of the year where:
 - Quarter 1 = Jan - Mar
 - Quarter 2 = Apr – Jun
 - Quarter 3 = Jul – Sep
 - Quarter 4 = Oct - Dec
2. Pearson correlation between quarters of the spatial distributions of CPUE for each species. (Note, in order to avoid non-representative CPUE values due to a small amount of effort the spatial distribution over which the correlation was calculated was limited to the 76 one-degree squares where the deployed longline effort was greater than 3000 hooks in each quarter.)
3. Pearson correlation, by quarter, between the spatial distributions of CPUE for a given specie and each other species. (Note: uses the same 76 one-degree squares as for 2 above.)

Results are shown for the following species:

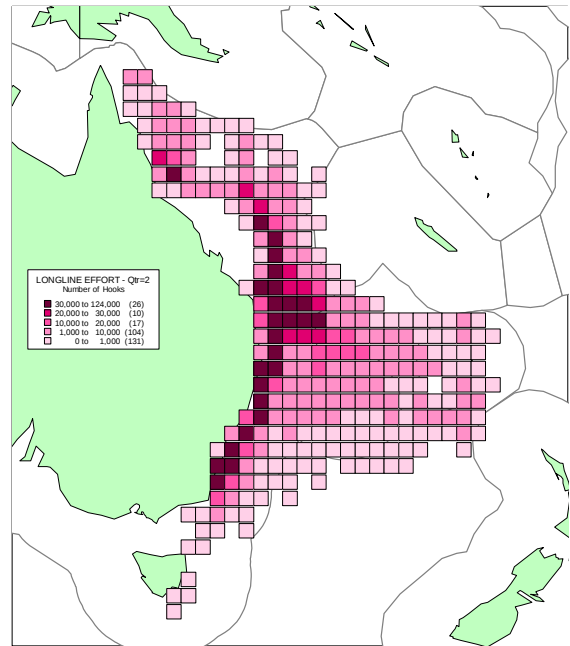
1	YFT	Yellowfin Tuna
2	BET	Bigeye Tuna
3	ALB	Albacore Tuna
4	SBT	Skipjack Tuna
5	NBT	Northern Bluefin Tuna
6	SKJ	Southern Bluefin Tuna
7	SWO	Broadbill Swordfish
8	STM	Striped Marlin
9	BAM	Black Marlin
10	BUM	Blue Marlin
11	SBS	Short-billed Spearfish
12	SAF	Sailfish
13	DOL	Dolphin Fish
14	WAH	Wahoo
15	OPA	Opah
16	POA	Pomfrets
17	OIL	Oilfish

Figure B.1: Longline Effort

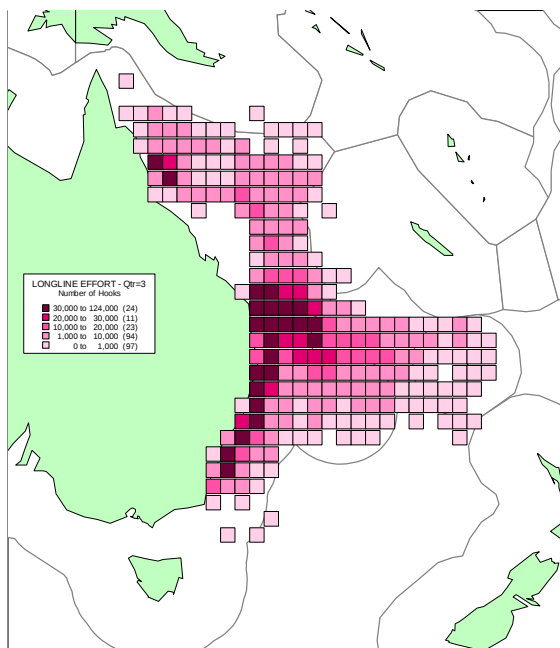
Quarter 1



Quarter 2



Quarter 3



Quarter 4

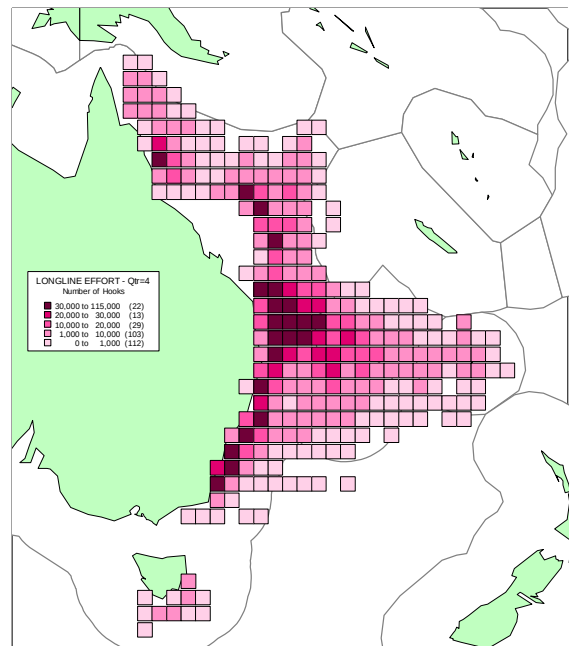
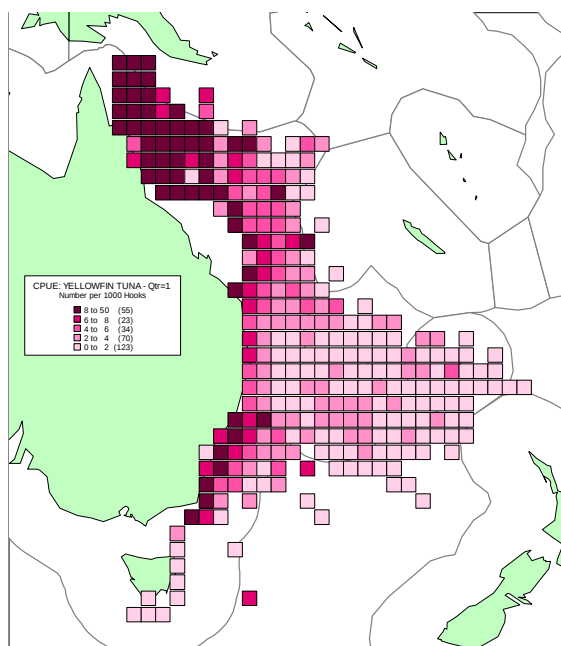
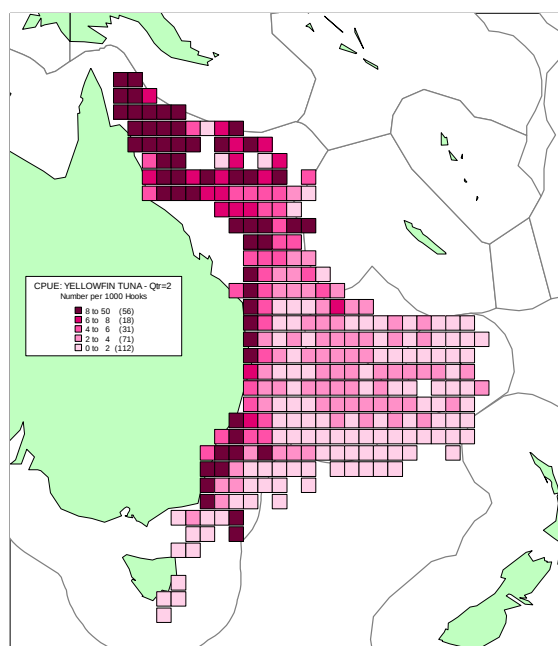


Figure B.2a: Yellowfin Tuna

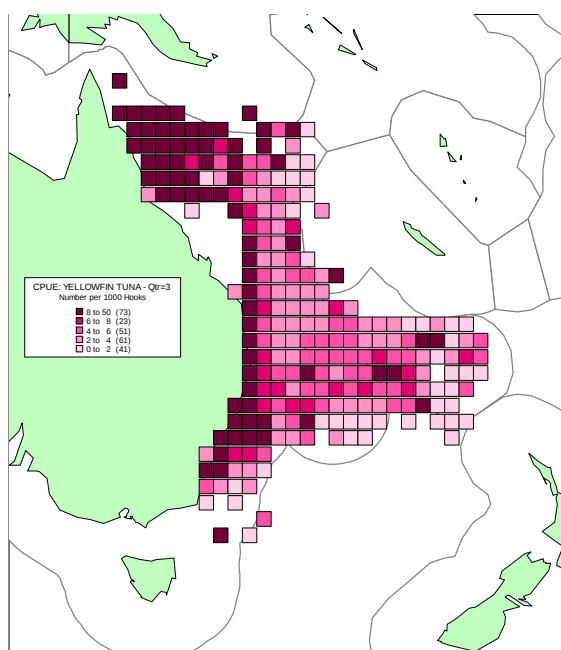
Quarter 1



Quarter 2



Quarter 3



Quarter 4

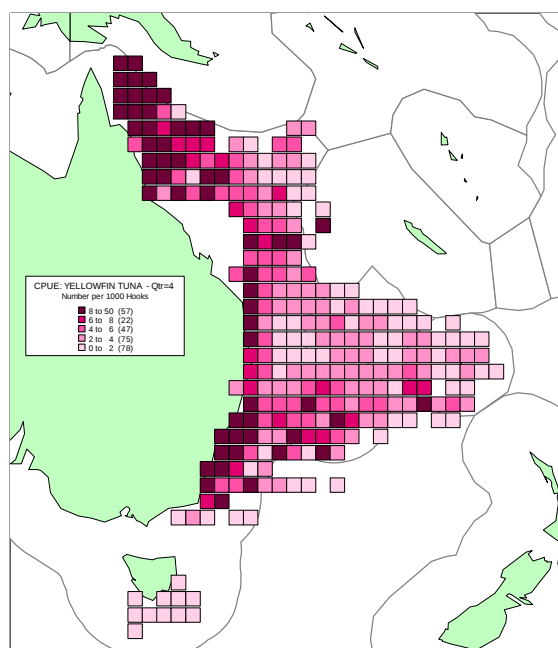
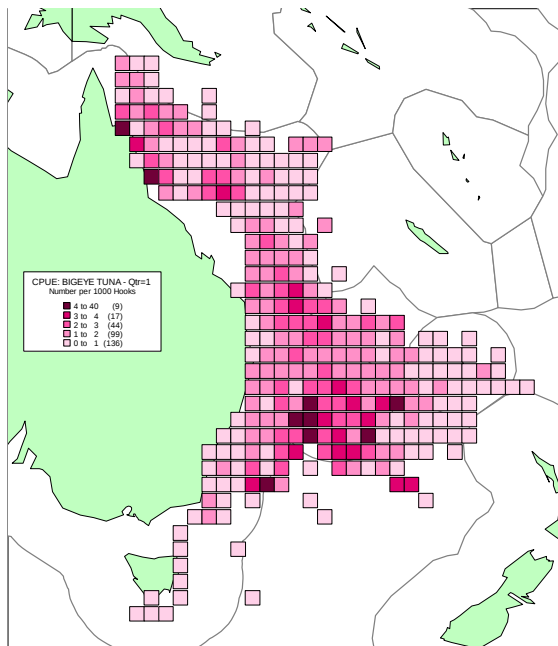
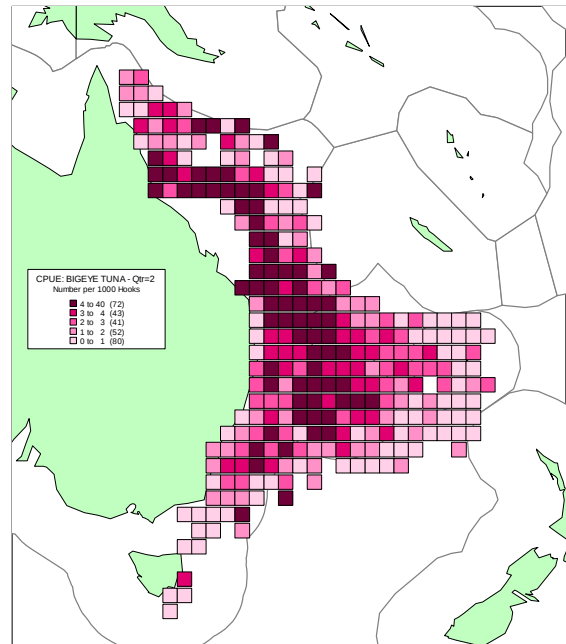


Figure B.2b: Bigeye Tuna

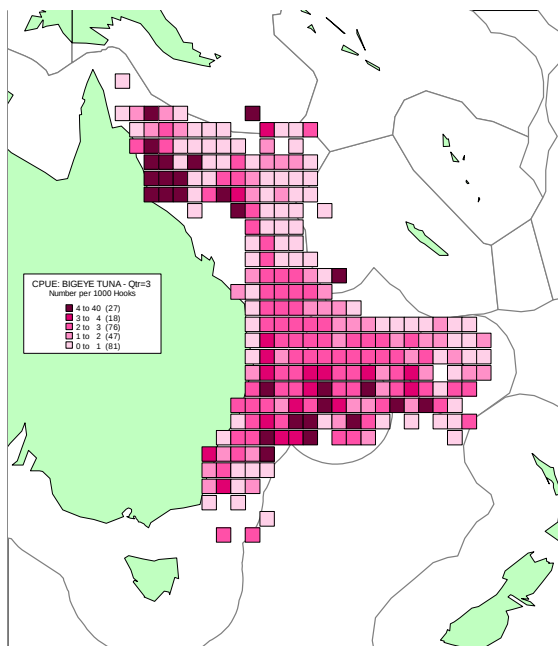
Quarter 1



Quarter 2



Quarter 3



Quarter 4

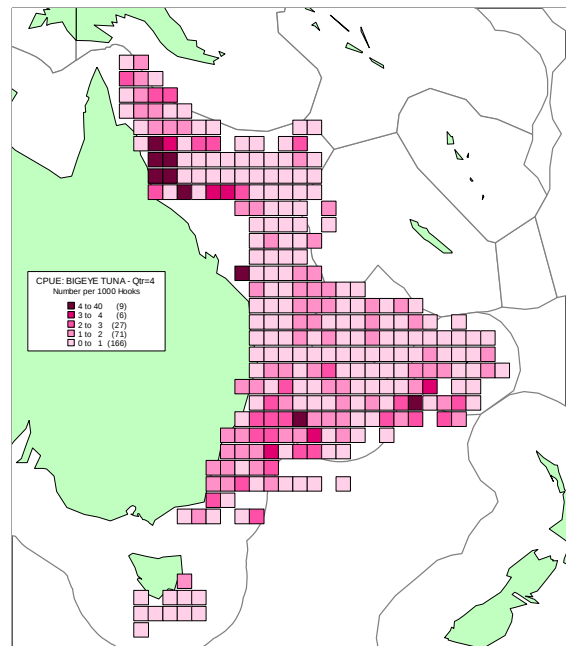
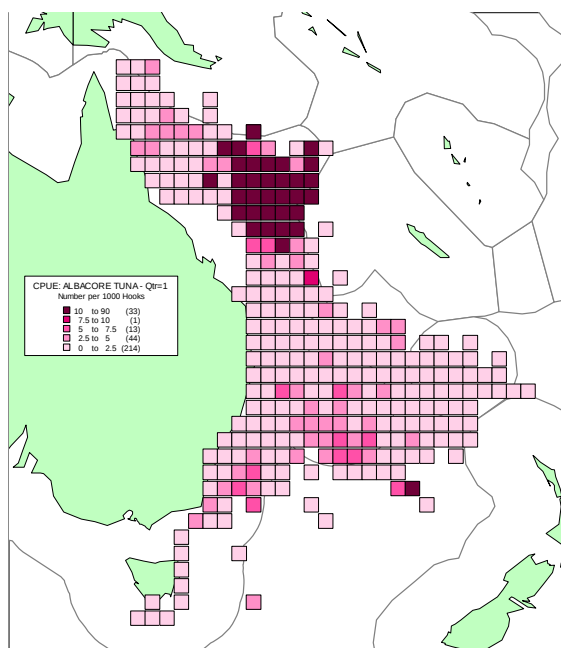
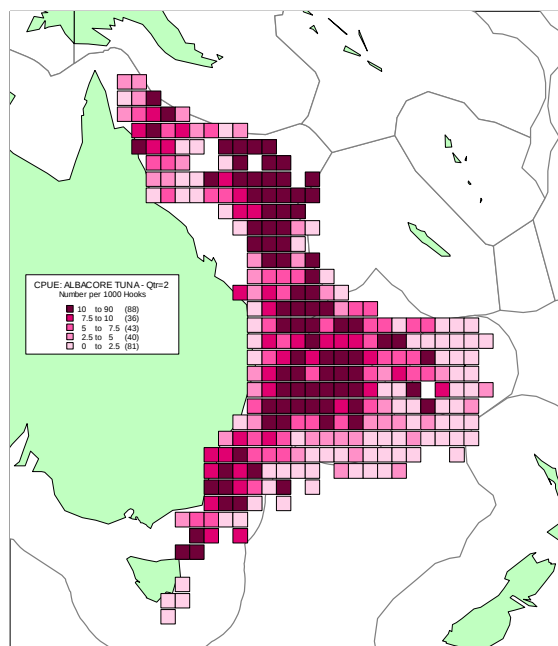


Figure B.2c: Albacore Tuna

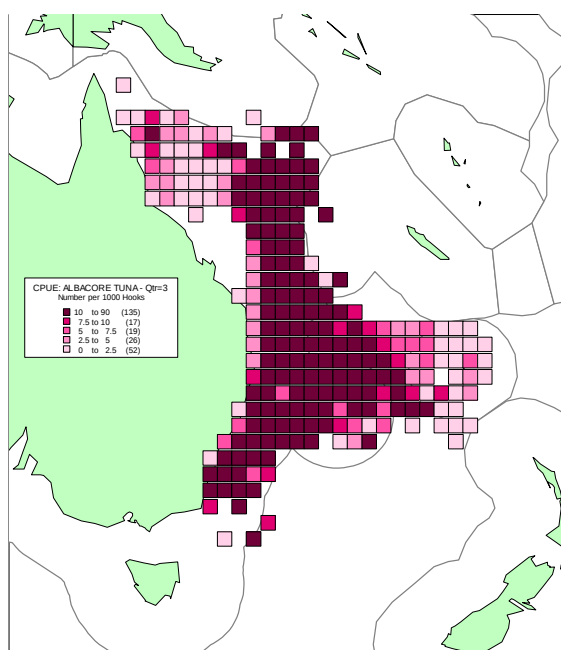
Quarter 1



Quarter 2



Quarter 3



Quarter 4

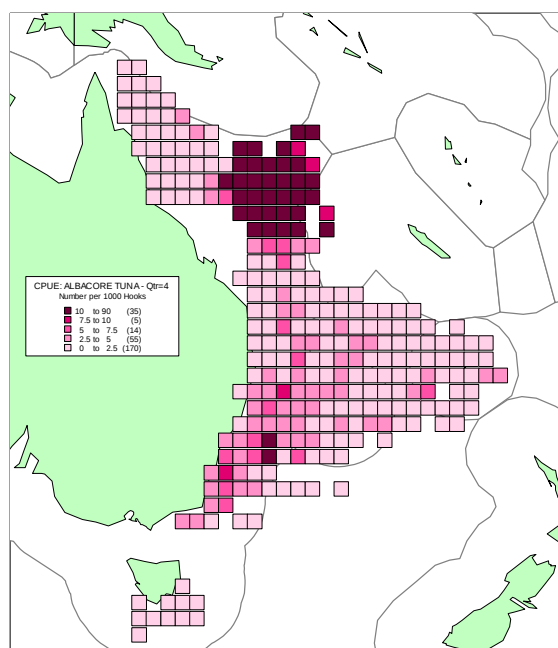
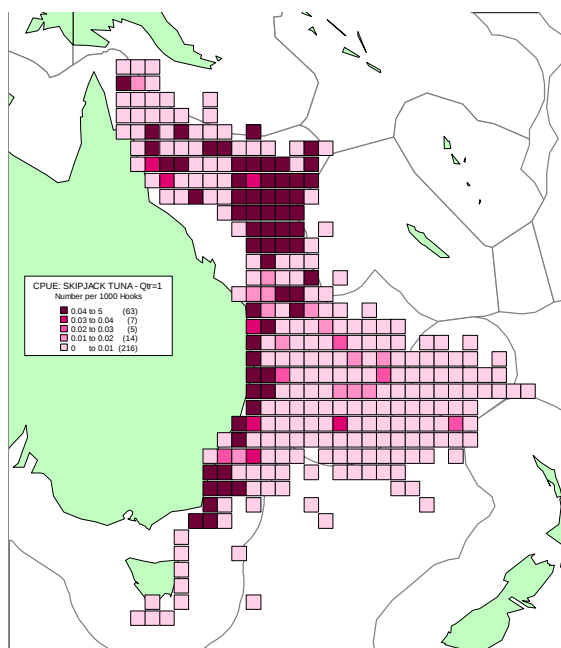
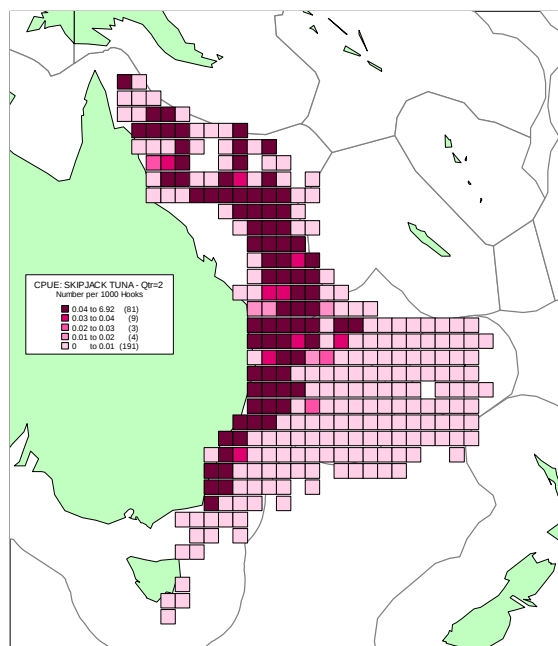


Figure B.2d: Skipjack Tuna

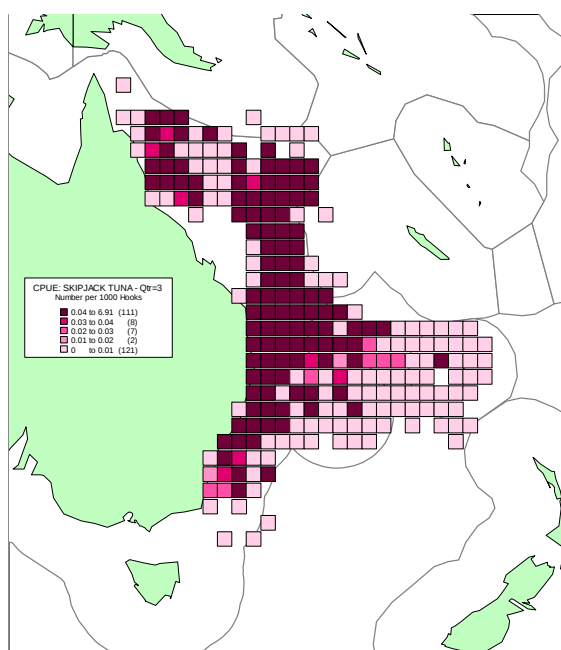
Quarter 1



Quarter 2



Quarter 3



Quarter 4

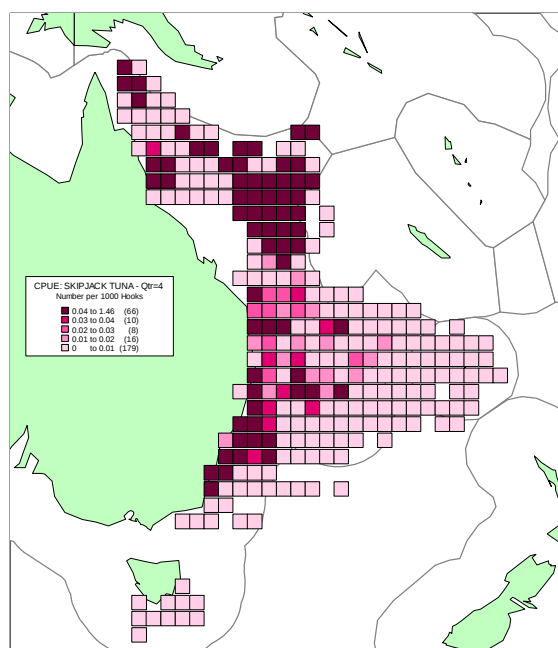
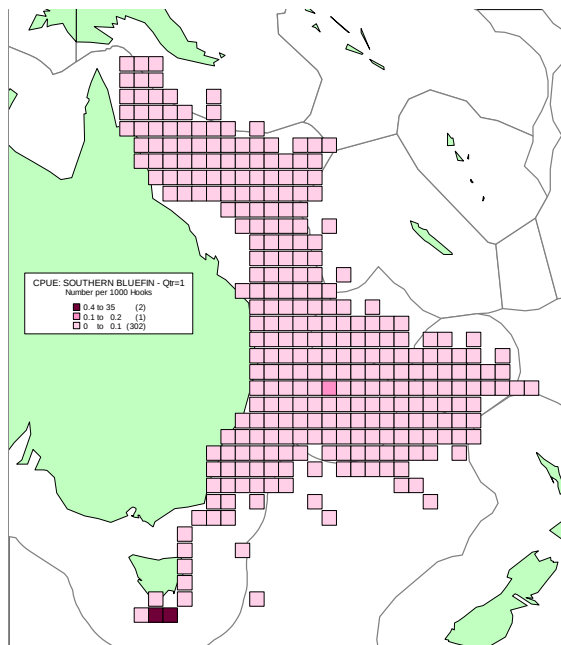
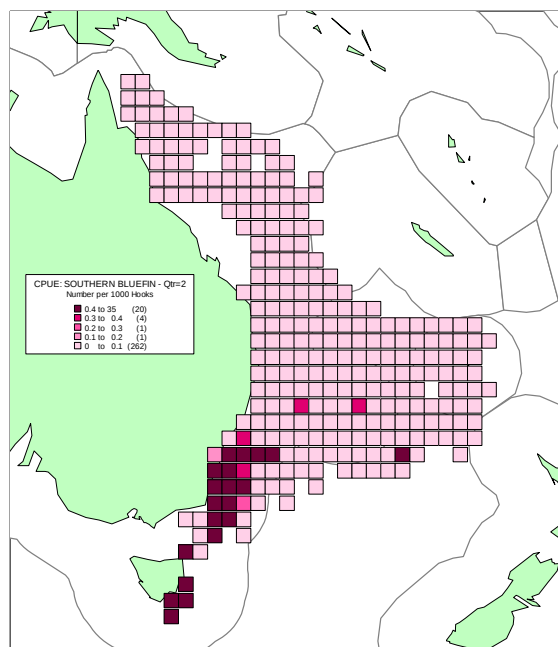


Figure B.2e: Southern Bluefin Tuna

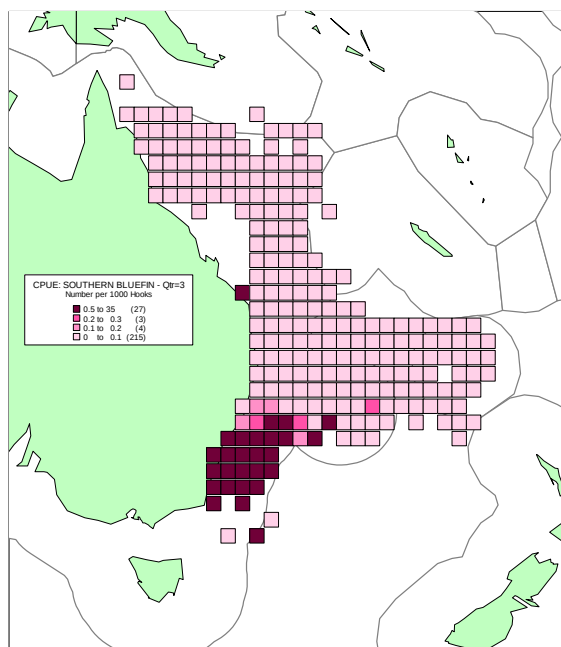
Quarter 1



Quarter 2



Quarter 3



Quarter 4

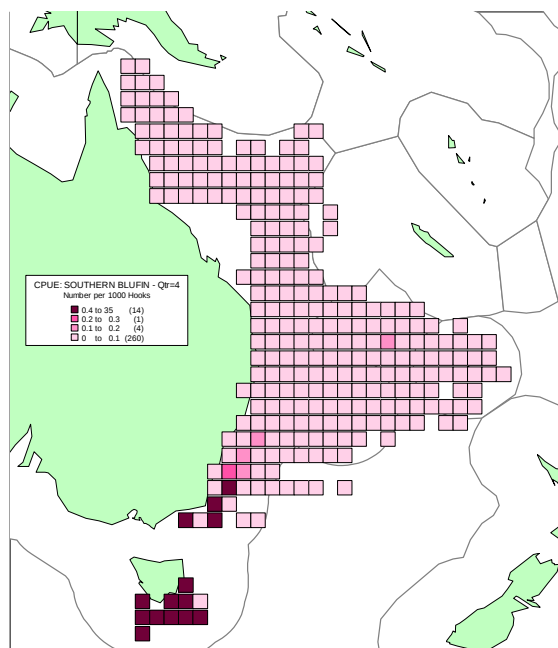
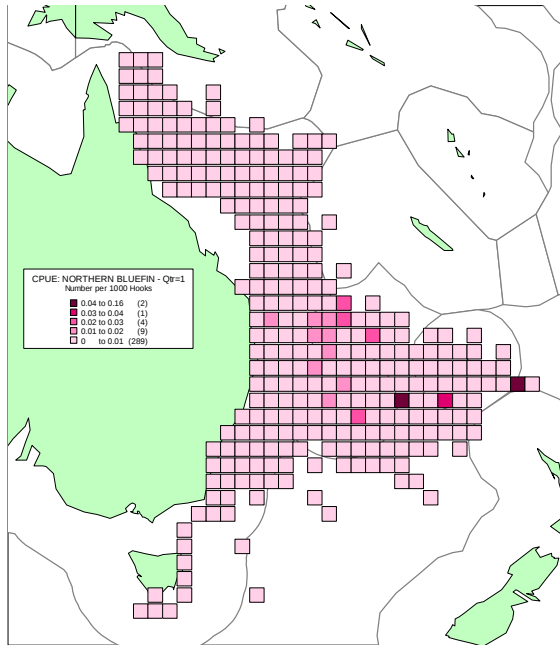
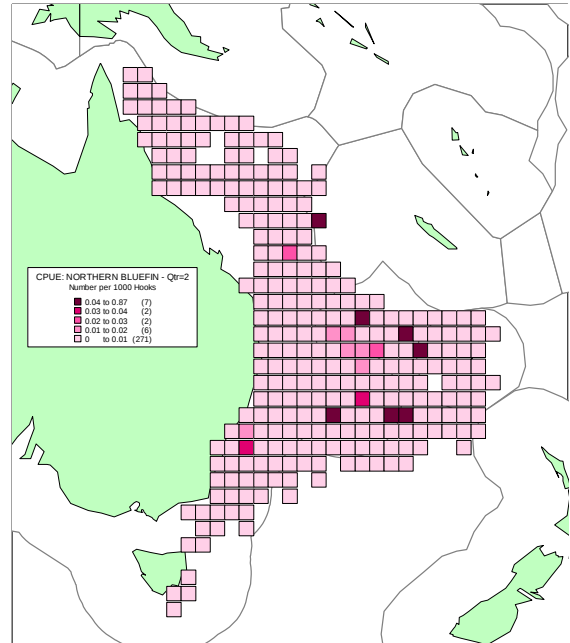


Figure B.2f: Northern Bluefin Tuna

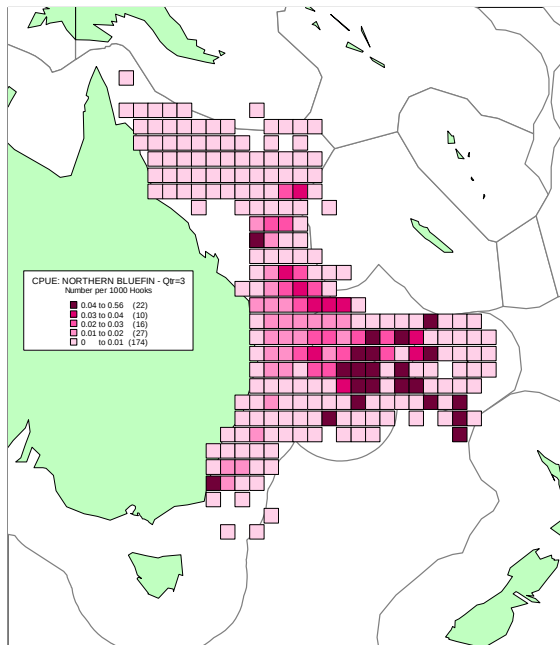
Quarter 1



Quarter 2



Quarter 3



Quarter 4

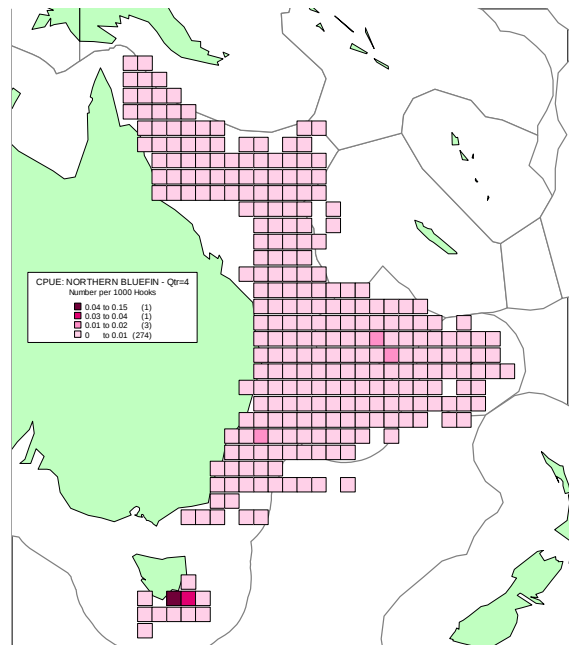
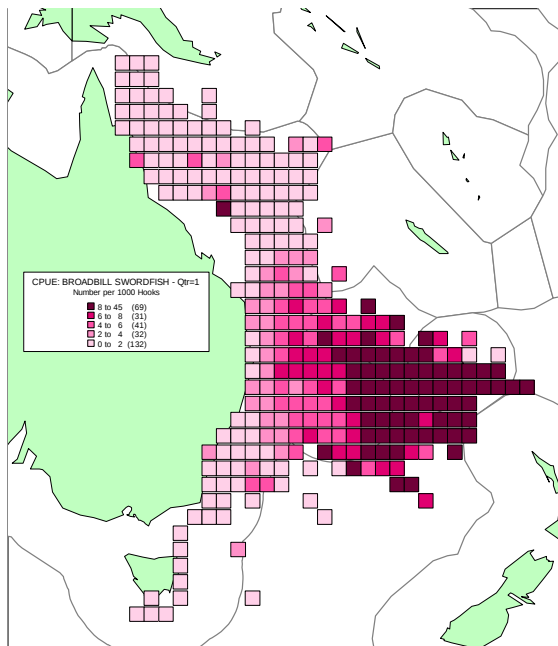
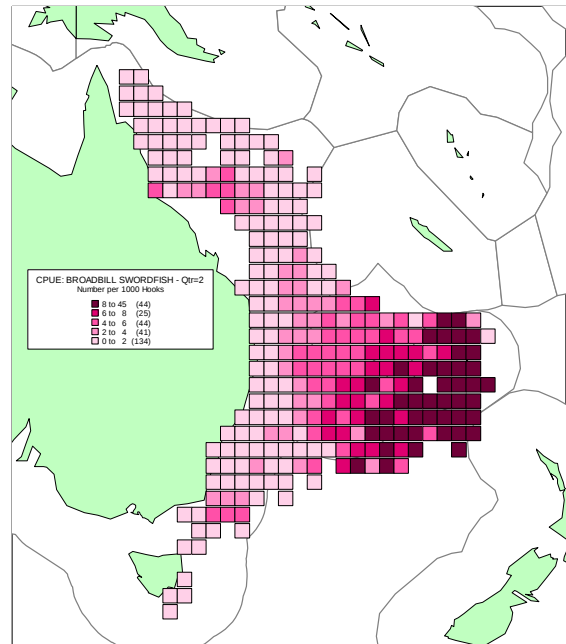


Figure B.2g: Broadbill Swordfish

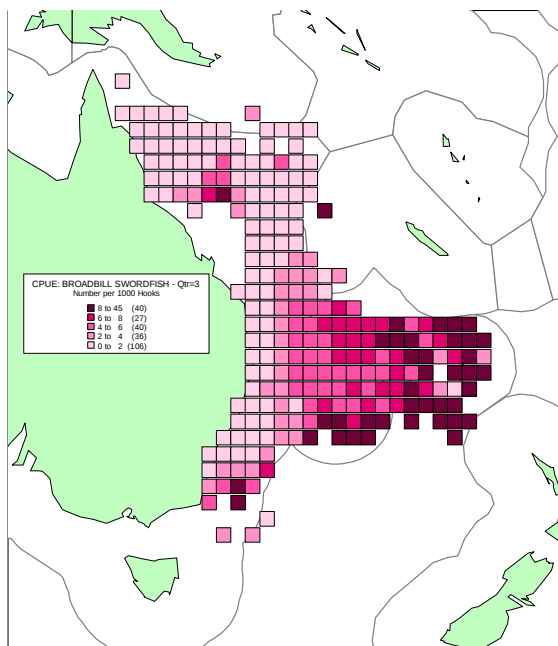
Quarter 1



Quarter 2



Quarter 3



Quarter 4

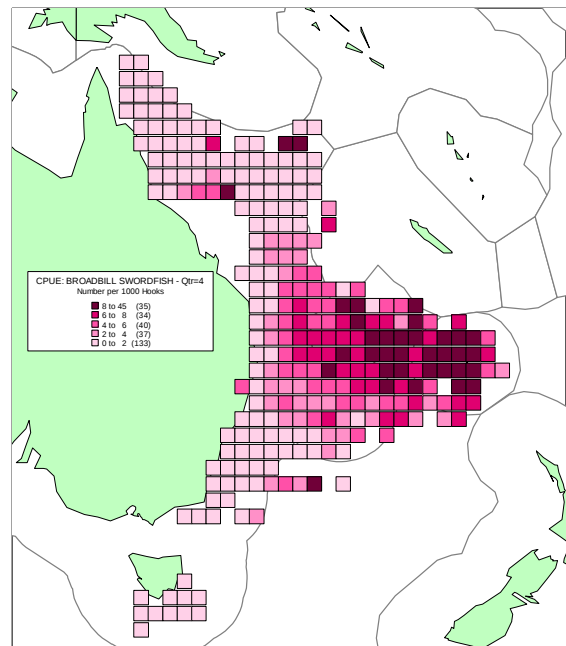
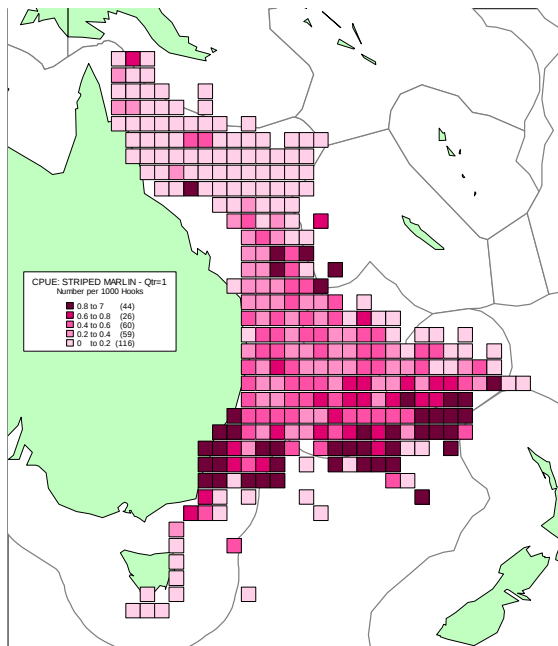
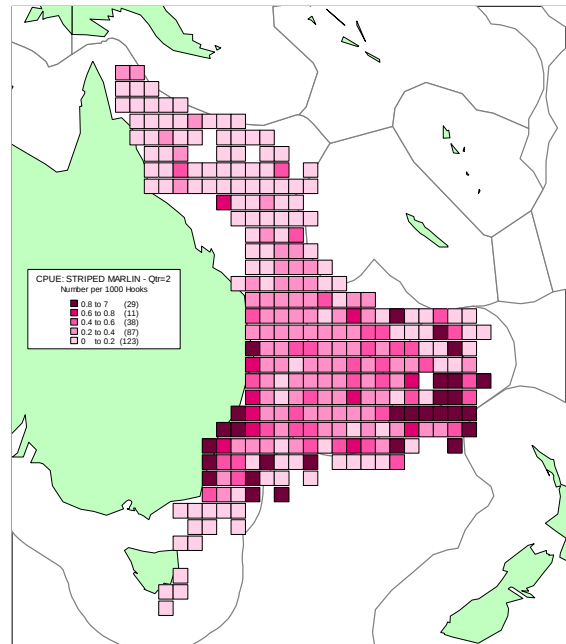


Figure B.2h: Striped Marlin

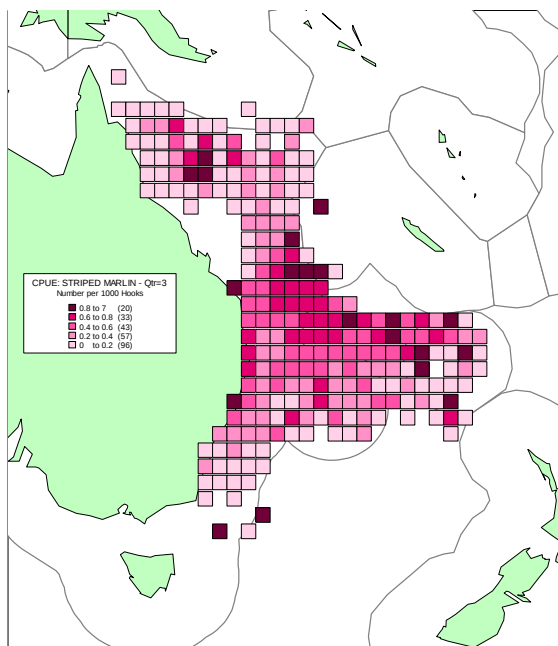
Quarter 1



Quarter 2



Quarter 3



Quarter 4

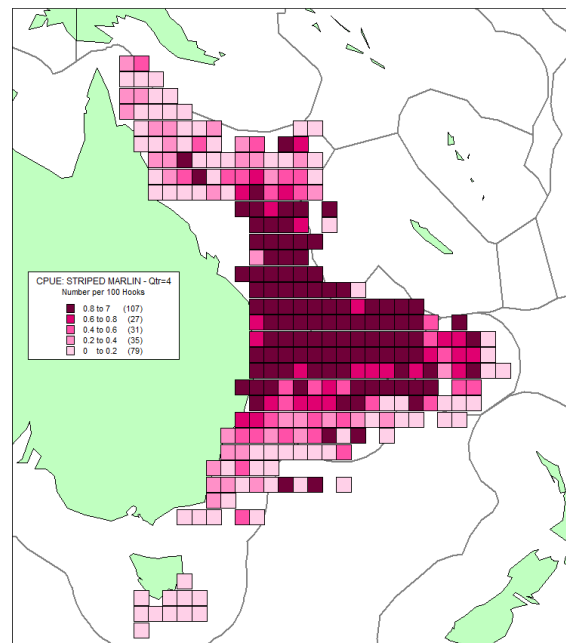
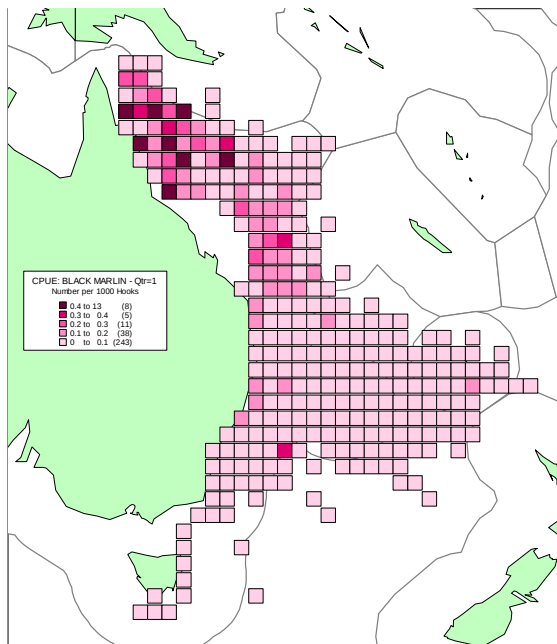
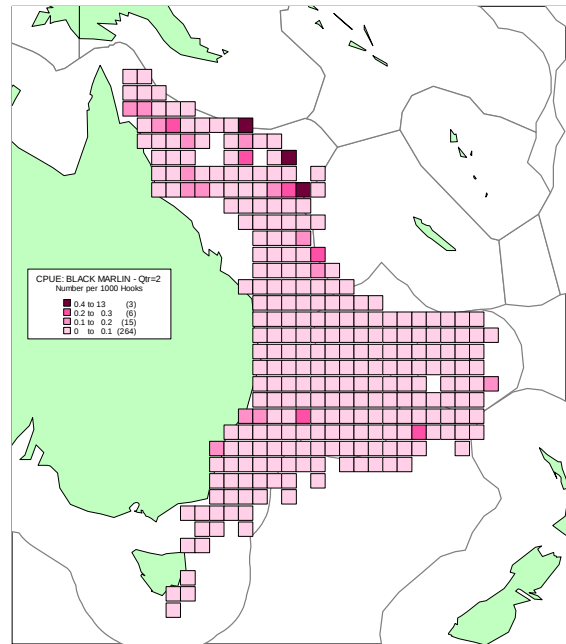


Figure B.2i: Black Marlin

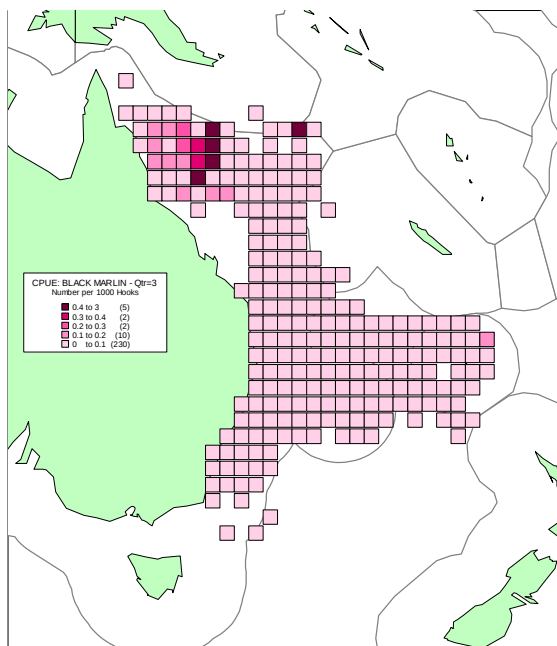
Quarter 1



Quarter 2



Quarter 3



Quarter 4

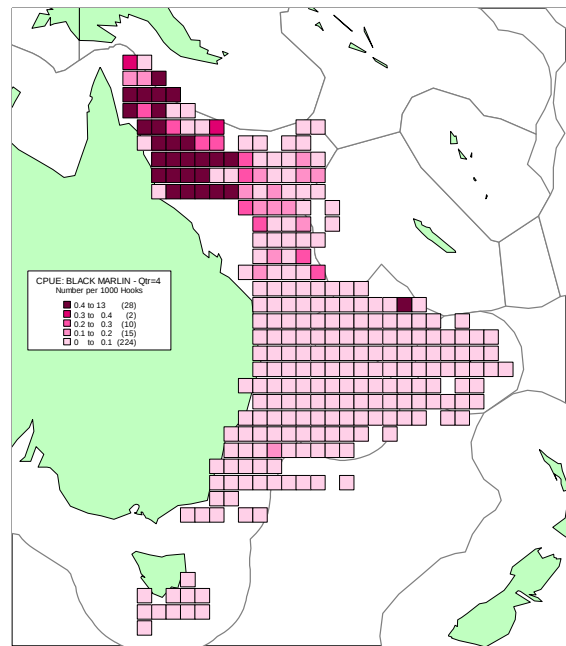
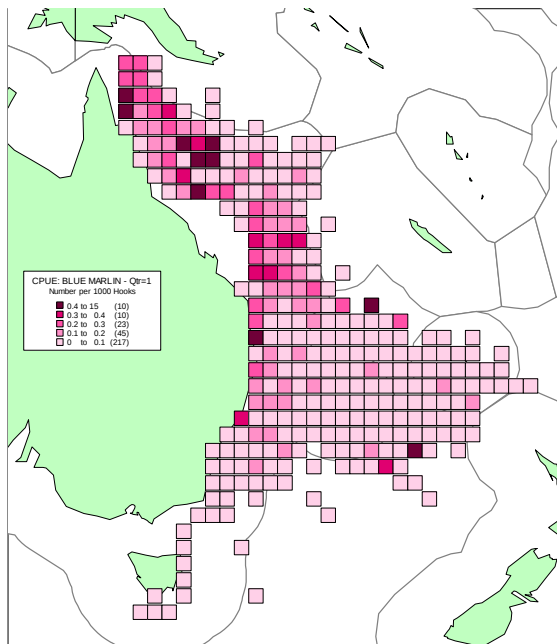
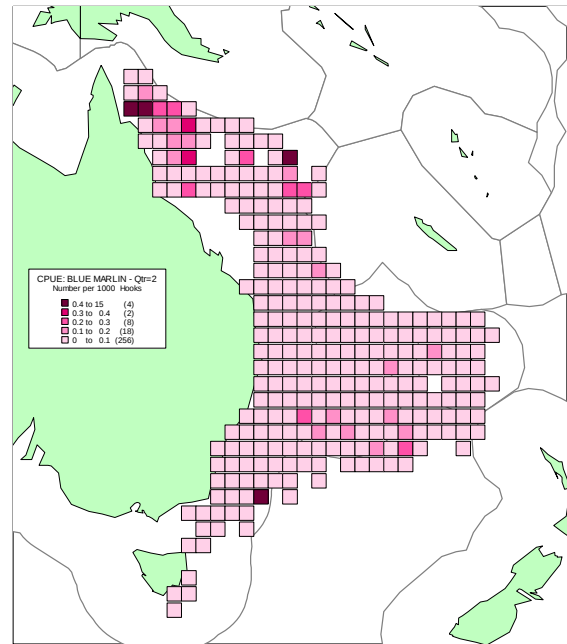


Figure B.2j: Blue Marlin

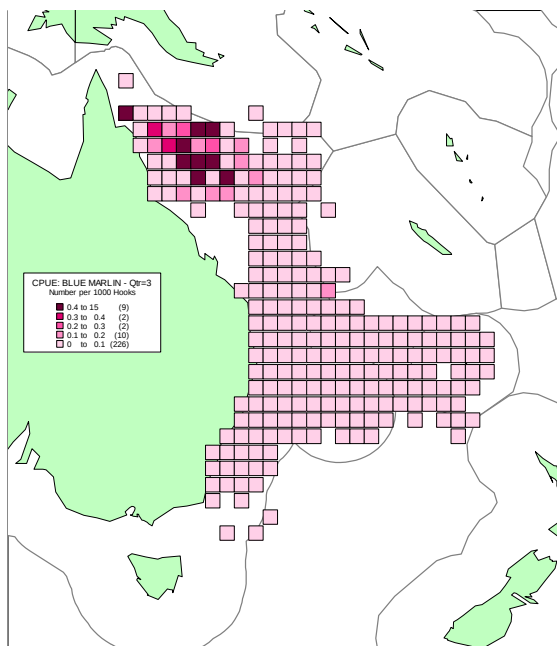
Quarter 1



Quarter 2



Quarter 3



Quarter 4

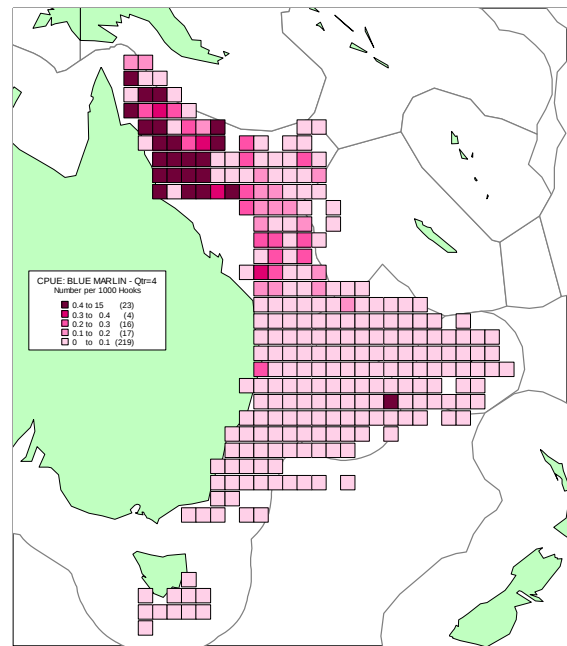
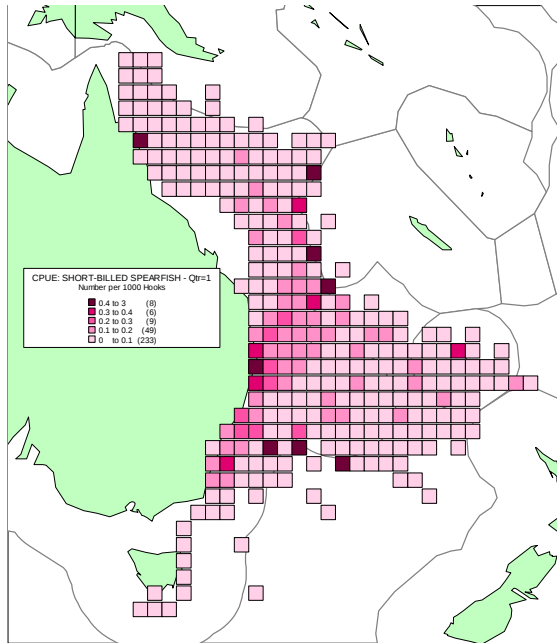
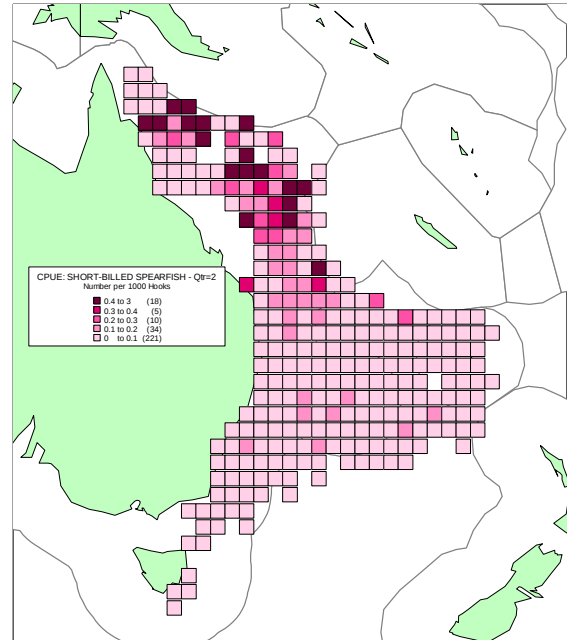


Figure B.2k: Short-billed Spearfish

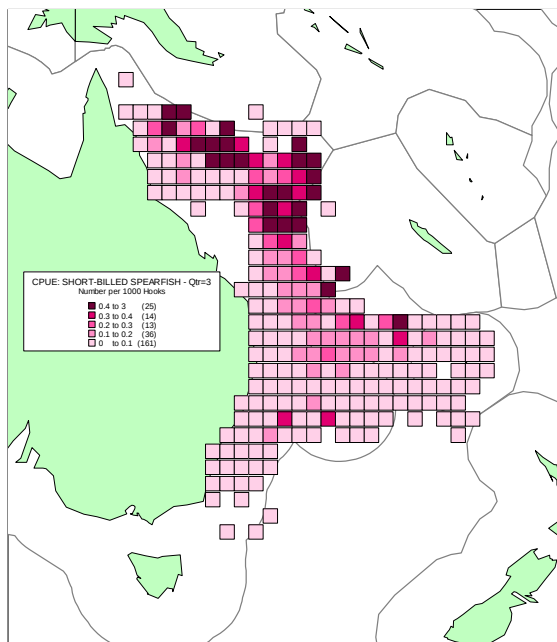
Quarter 1



Quarter 2



Quarter 3



Quarter 4

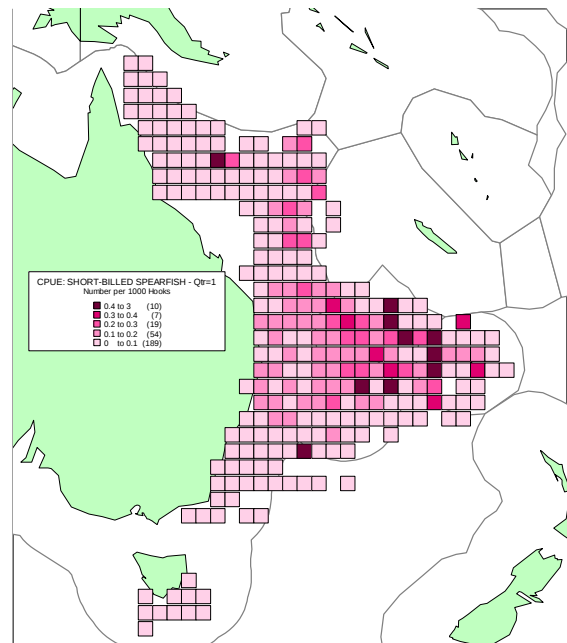
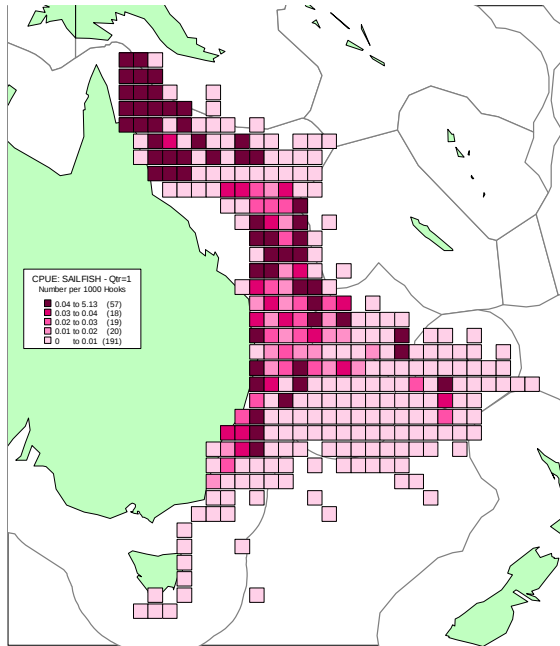
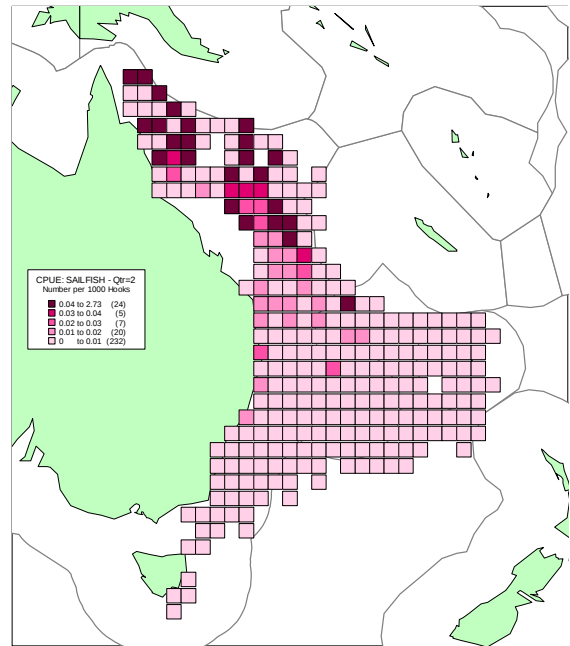


Figure B.21: Sailfish

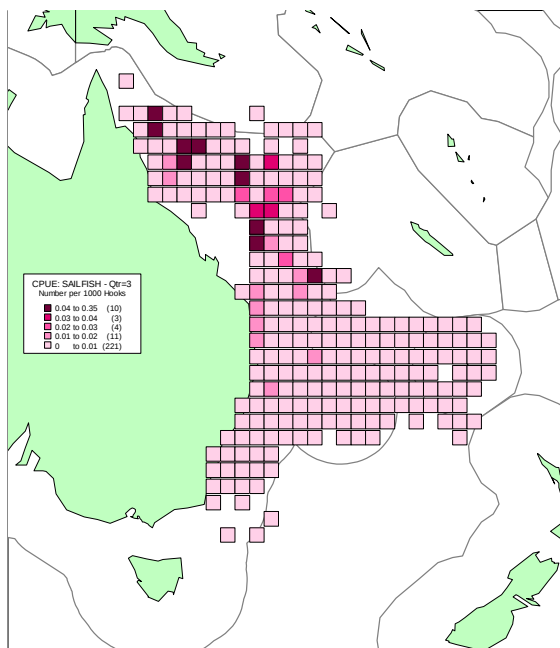
Quarter 1



Quarter 2



Quarter 3



Quarter 4

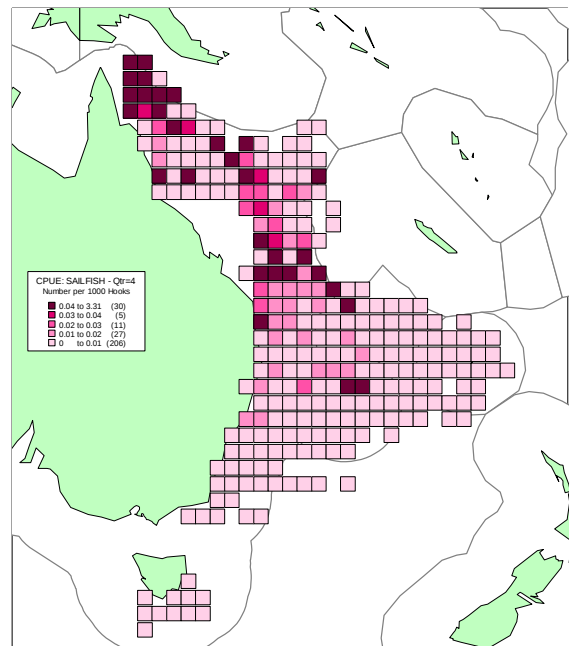
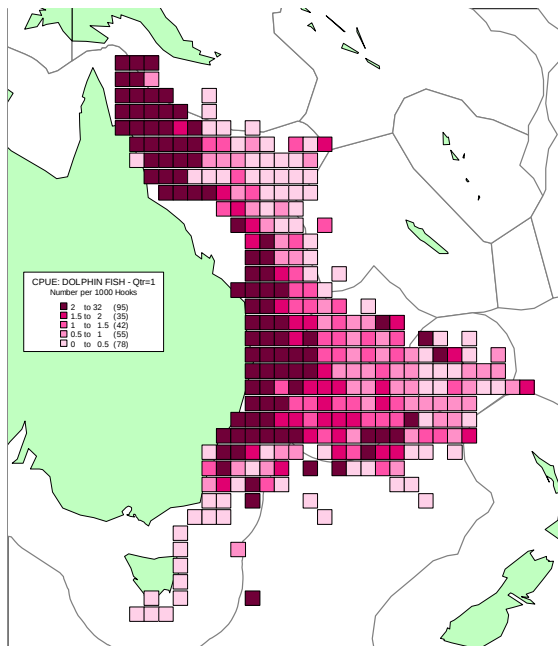
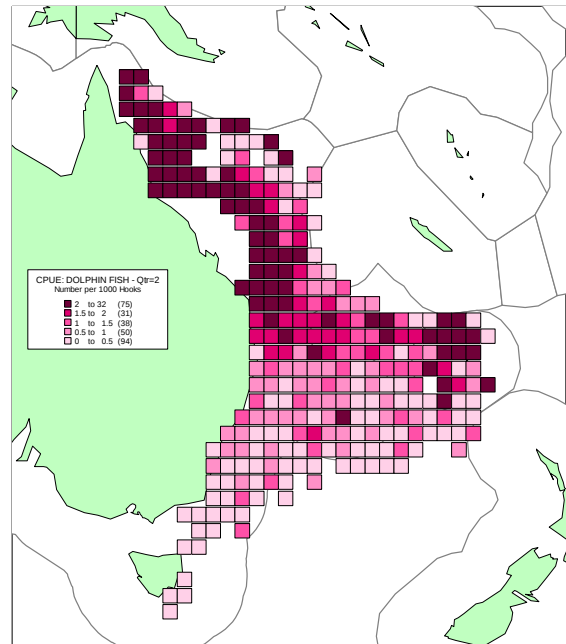


Figure B.2m: Dolphin Fish

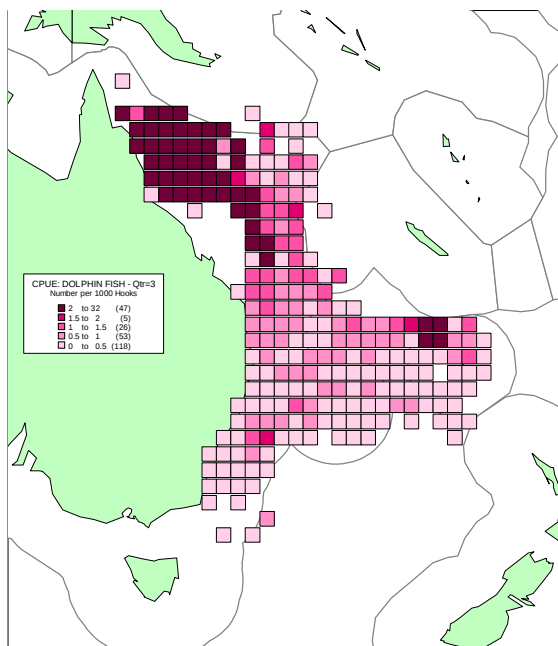
Quarter 1



Quarter 2



Quarter 3



Quarter 4

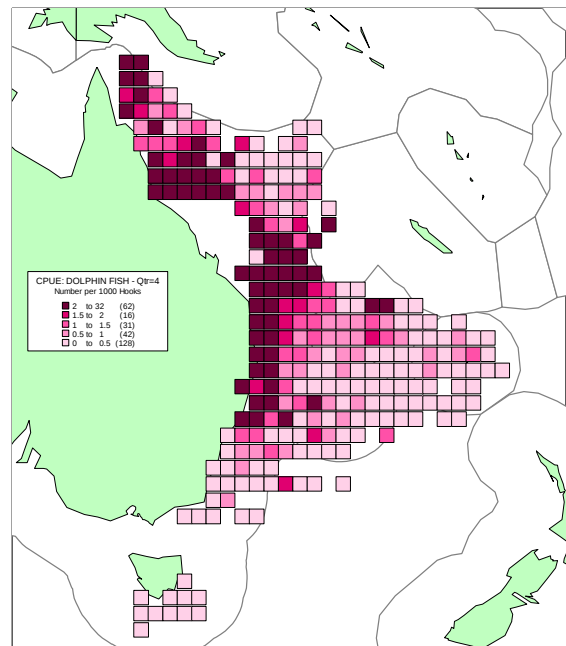
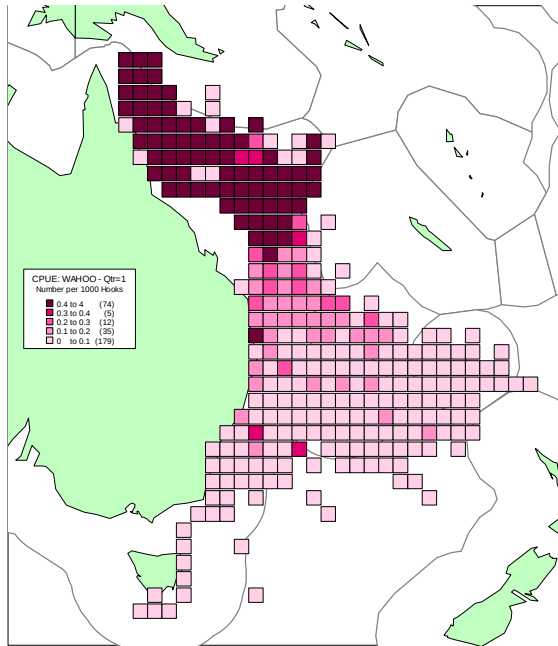
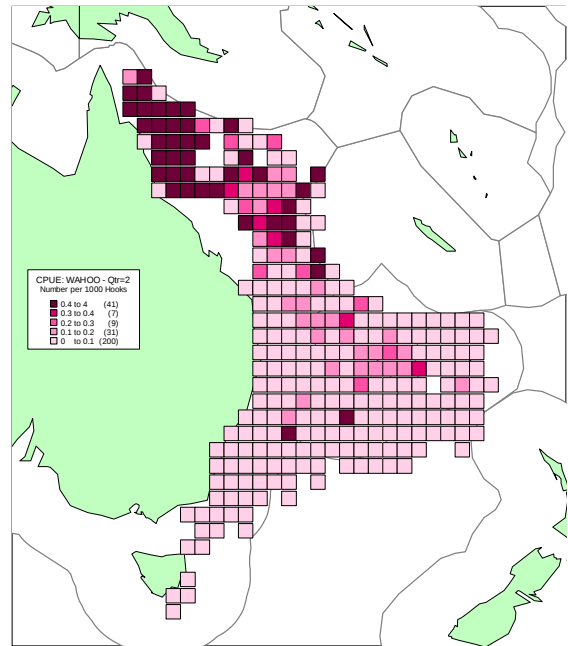


Figure B.2n: Wahoo

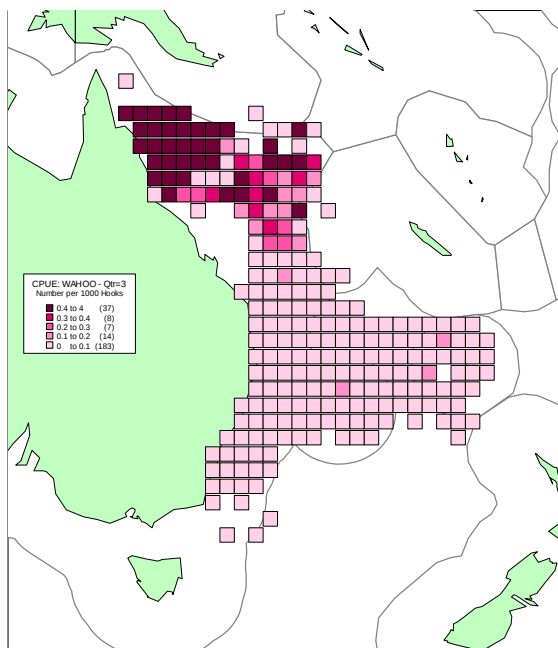
Quarter 1



Quarter 2



Quarter 3



Quarter 4

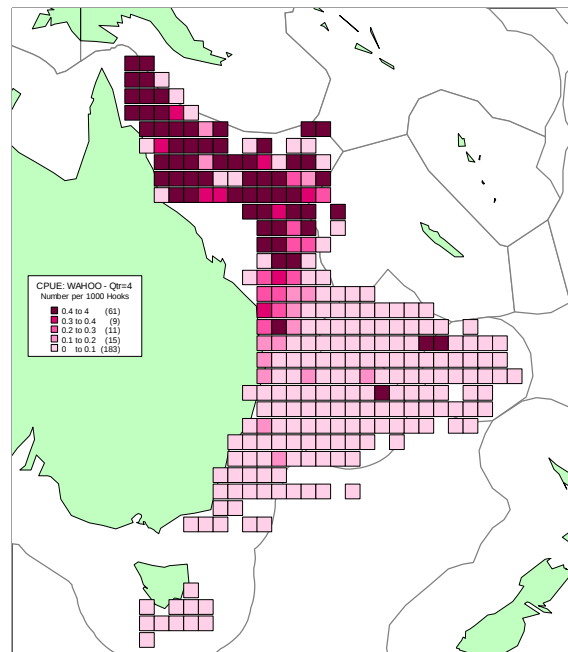
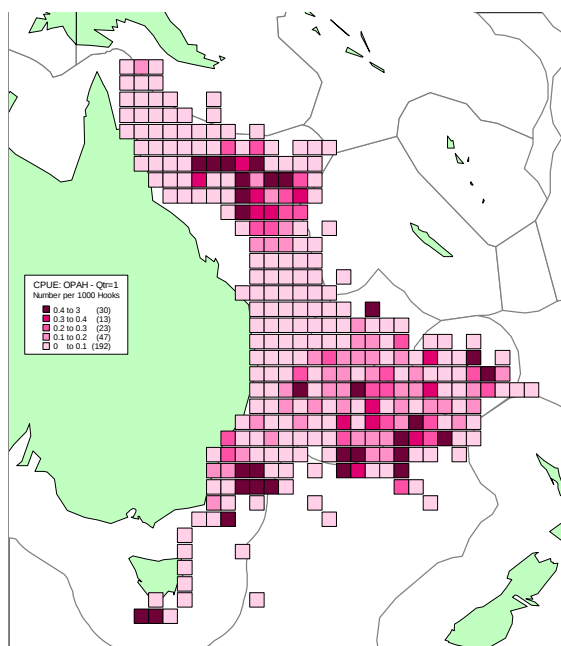
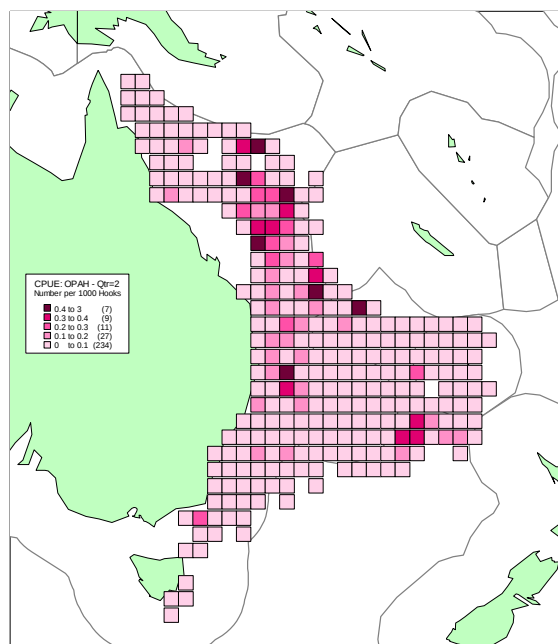


Figure B.2o: Opah

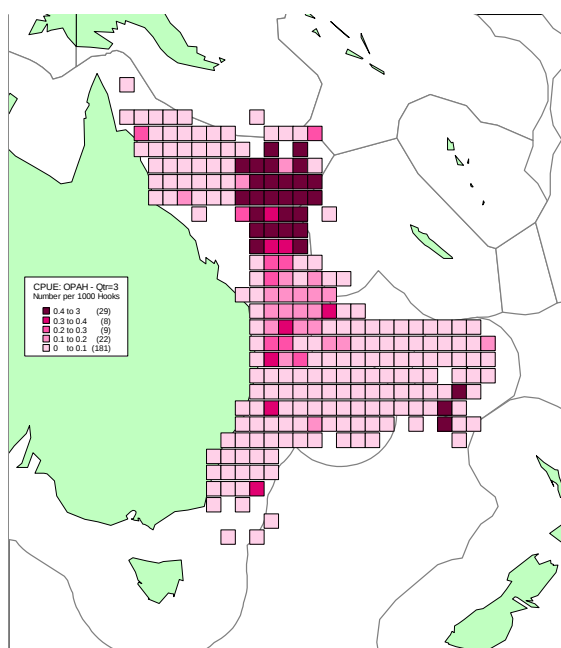
Quarter 1



Quarter 2



Quarter 3



Quarter 4

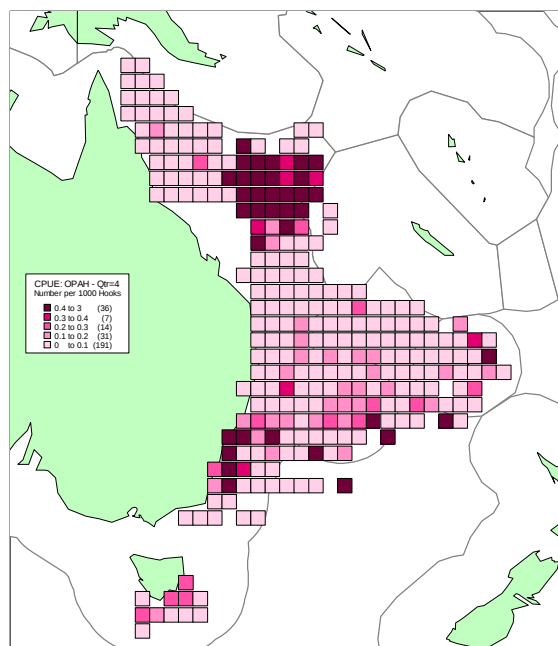
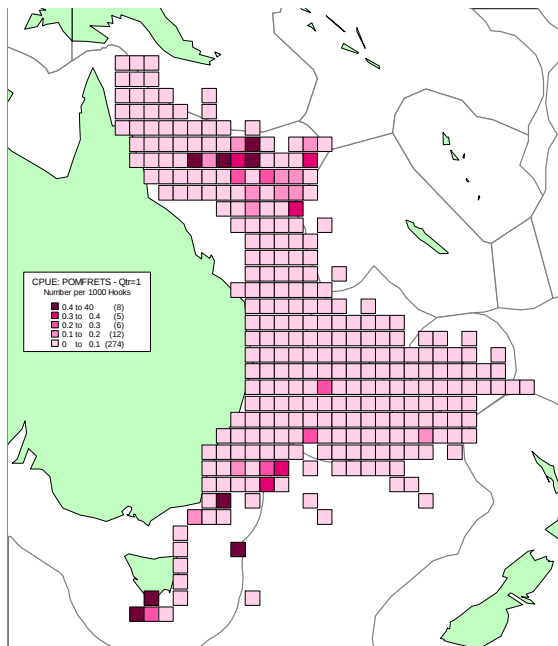
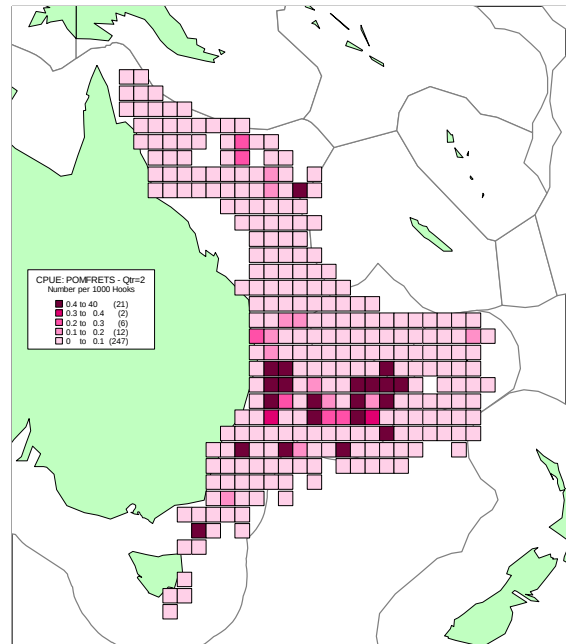


Figure B.2p: Pomfrets

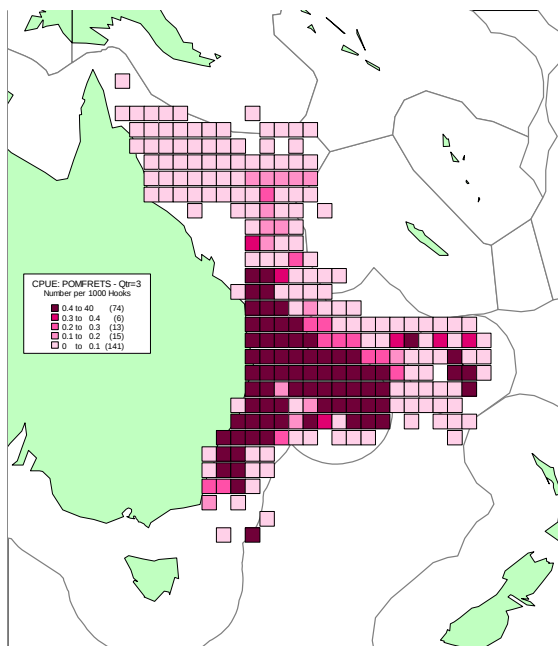
Quarter 1



Quarter 2



Quarter 3



Quarter 4

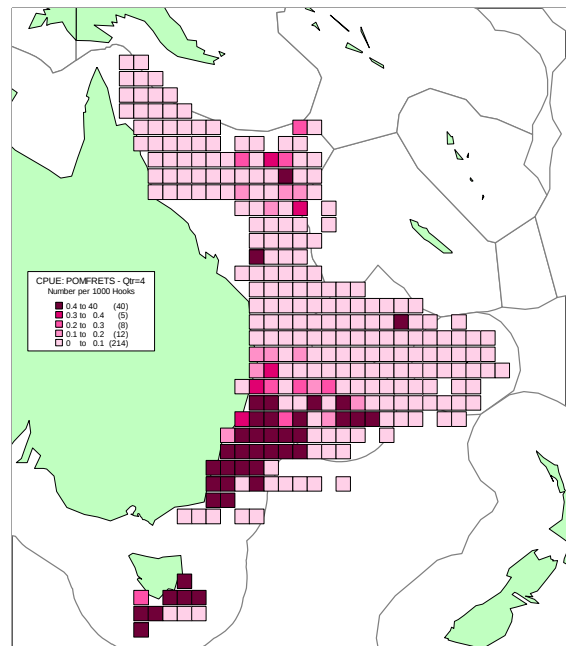
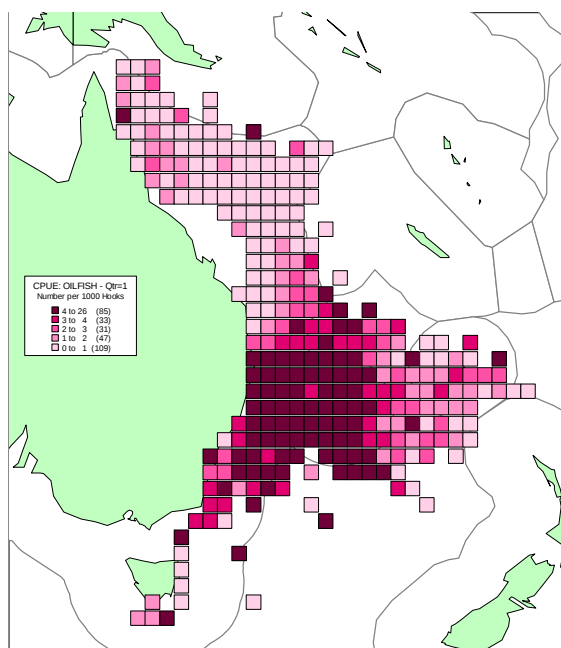
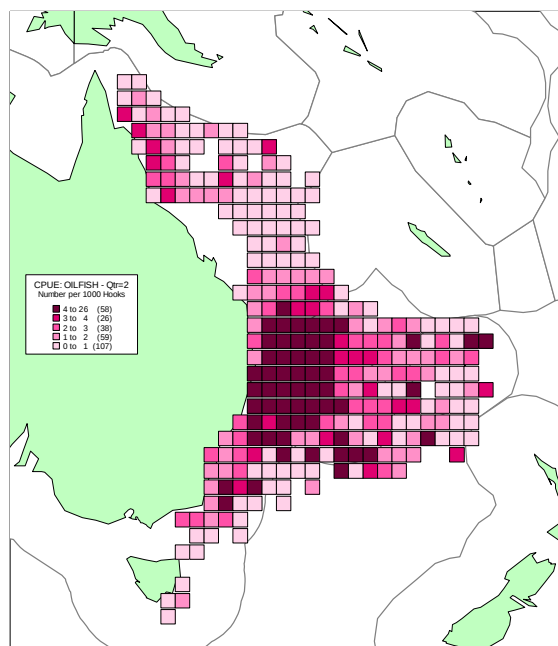


Figure B.2q: Oilfish

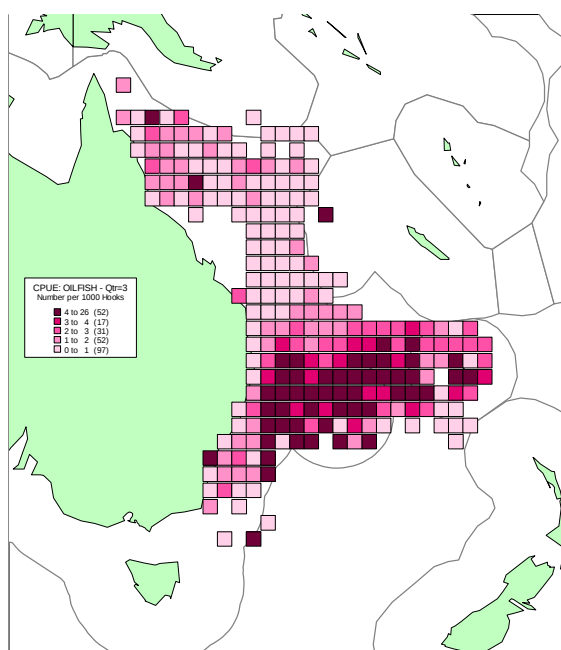
Quarter 1



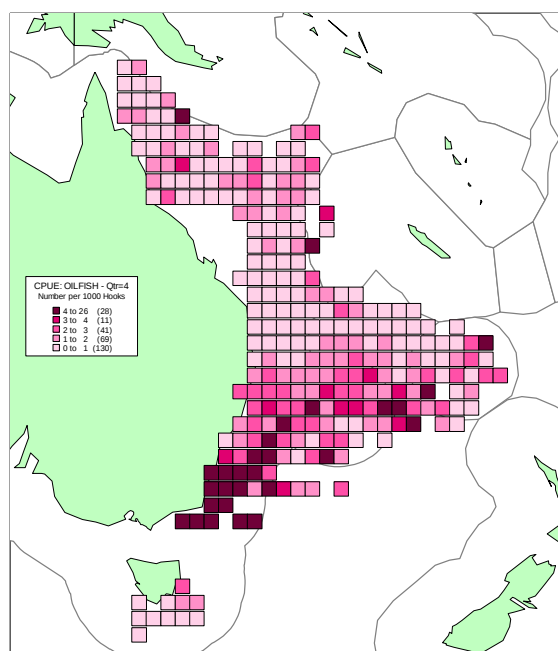
Quarter 2



Quarter 3



Quarter 4



Appendix C. Clustering by deployed fishing gears

Numerical data (such as catch number) have meaning as a measurement and as such a 'distance' (or dissimilarity measure for use in a cluster analysis) between any two such data points is easy to define. On the other hand, for categorical data which represent characteristics such as a fish's gender, reproductive stage, or the types of gears used in a fishing operation, it is not obvious how one define a 'distance' or 'dissimilarity' measure. One can associate categorical data with numerical values (such as "1" indicating male and "2" indicating female) but these numbers don't have mathematical meaning. In order to define a 'distance' measure for clustering by gear-type the following approach was used.

First, the following gears-settings were identified:

- | | |
|----------------------|--|
| A) Hooks-per-float | 7 levels (<8, 8-9, 10-11,12-14,15-19,20-29,30-40) |
| B) Start-time | 6 levels (0-4am, 4-8am, 8-12am, 0-4pm, 4-8pm, 8-12pm) |
| C) Bait | 8 levels |
| D) Light-stick usage | 7 levels (0-19%, 20-39%, 40-59%, 60-79%, 80-99%, 100%) |

Second, for each set the use of each gear and level was coded by a series of 0's and 1's. This provides a string of 28 characters long consisting of 4 '1's and 24 '0's for each data record. Third, the Jaccard measure of similarity was then calculated between each set where the *Jaccard coefficient* J is defined as the number of variables that are coded as 1 for both states divided by the number of variables that are coded as 1 for either or both states. Since dissimilarity measures are required by the cluster procedure, the *Jaccard measure of dissimilarity* ($1-J$) is selected. Finally, using this distance measure a hierarchical cluster analysis using the *Ward* method is undertaken on each of the data records.

As with the cluster analysis based on species composition, the above analysis was first carried out separately for each month and then as a single analysis across all records combined. For the monthly analysis seven cluster-types were identified for each month and the aggregate species composition for each cluster-type was calculated. These 84 species-compositions (7 clusters x 12 months) were then subject to a second cluster analysis in order to identify a set of dominant cluster-types. Based on an inspection of the Pseudo-F clustering criterion for each month (c.f. Figure C.1) this final number of clusters was set to be nine. A pictorial representation of the nine clusters for each month based on plots of the first two canonical variables are shown in Figure C.2.

The distribution of the 84 monthly cluster-types identified in stage 1 of the analysis across the nine cluster types identified in second cluster analysis based on species-composition are shown in Figure C.3. For each cluster-type the catch composition based on 11 species is shown. Note, where more than one of the 7 cluster-types identified for a month related to the same stage-2 cluster-type, these clusters were combined into a single cluster. This resulted in a total of 52-monthly clusters being distributed across the 12 months.

The distribution of the nine cluster types (as a percentage of sets) across each year and month are shown in Figures C.4 while the distribution of settings for four gear-types (bait-type, hooks-per-float, light-stick usage and set start-time) across the seven cluster-types is shown in Figure C.5. While this latter figure shows that a large range of gear-settings are associated with each cluster-type, again certain combinations of gear-settings are more likely to be associated with certain cluster-types (and resulting catches) than others.

Figure C.1: Clustering criteria used to identify the appropriate final number of clusters based on clustering the 84 monthly cluster types identified in the first stage of the gear-based analysis.

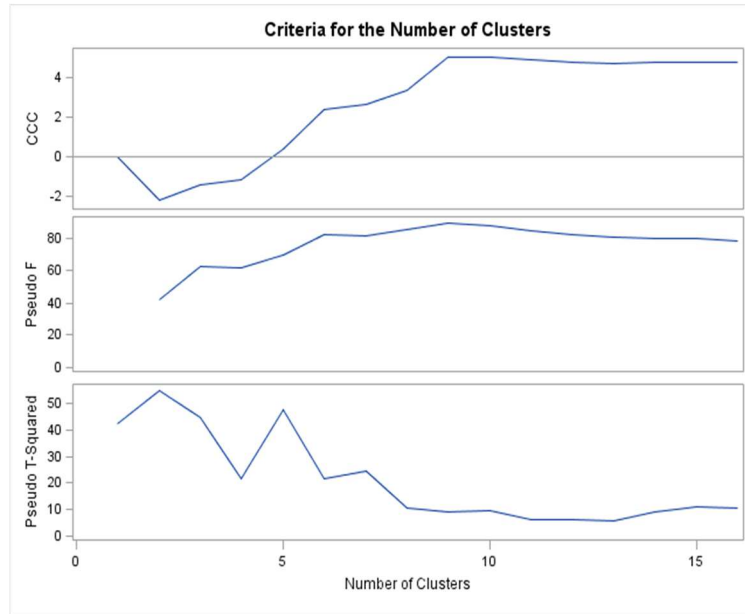
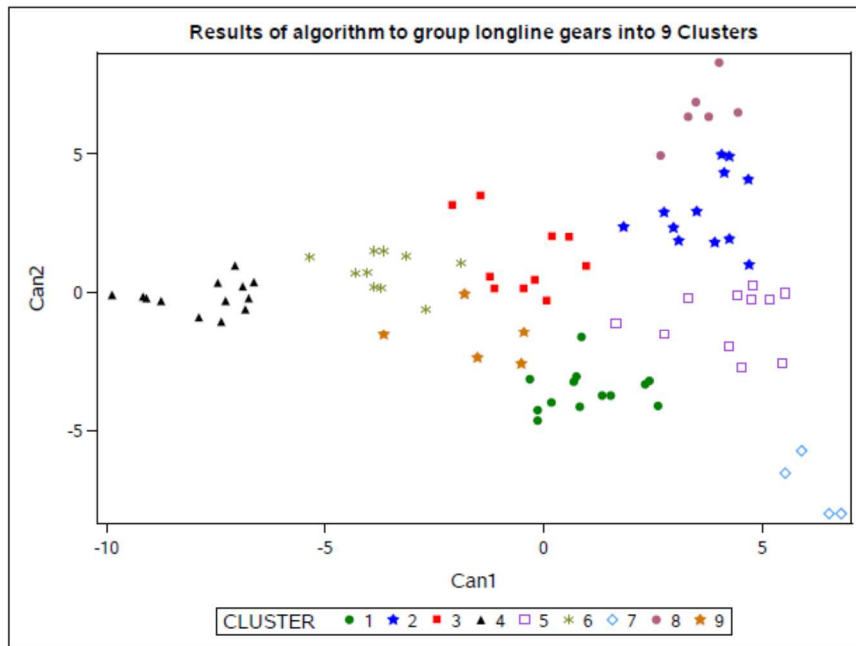


Figure C.2: Pictorial representation of the nine cluster types based on plots of the first two canonical variables for the 84 monthly gear-based clusters.



Finally, the catch composition of the nine cluster-types identified in the clustering the 84 monthly clusters (Figure C.6a) can be compared with the catch composition of the nine cluster-types identified directly by a single analysis of all 108,650 records (Figure C.6b).

Figure C.3: Distribution of the 84 monthly cluster types identified in stage 1 of the gear-based analysis across the nine cluster types identified in the second cluster analysis.

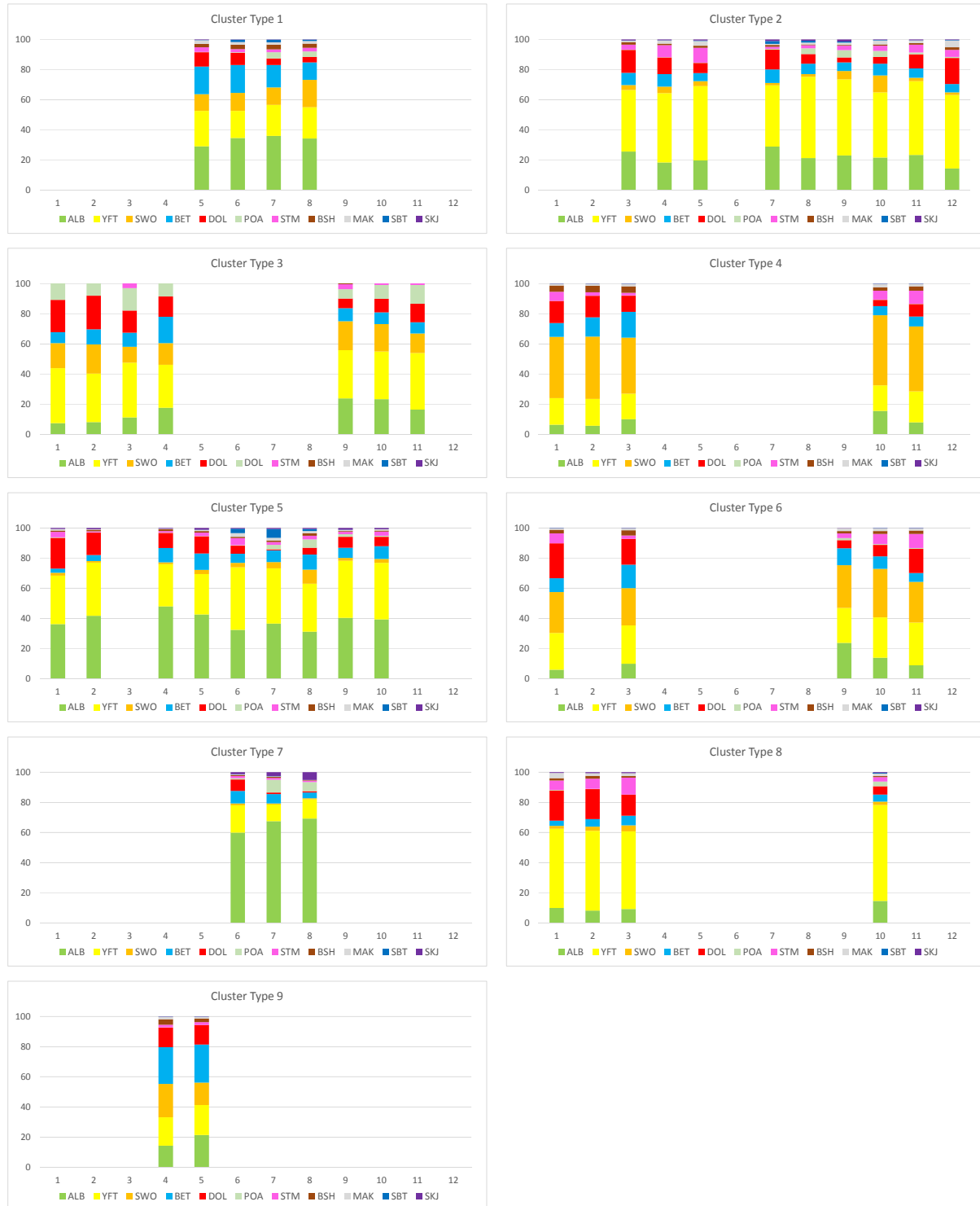


Figure C.4: Distribution of nine cluster types (percentage of sets) based on the gear-based cluster analysis across (a) month, and (b) year.

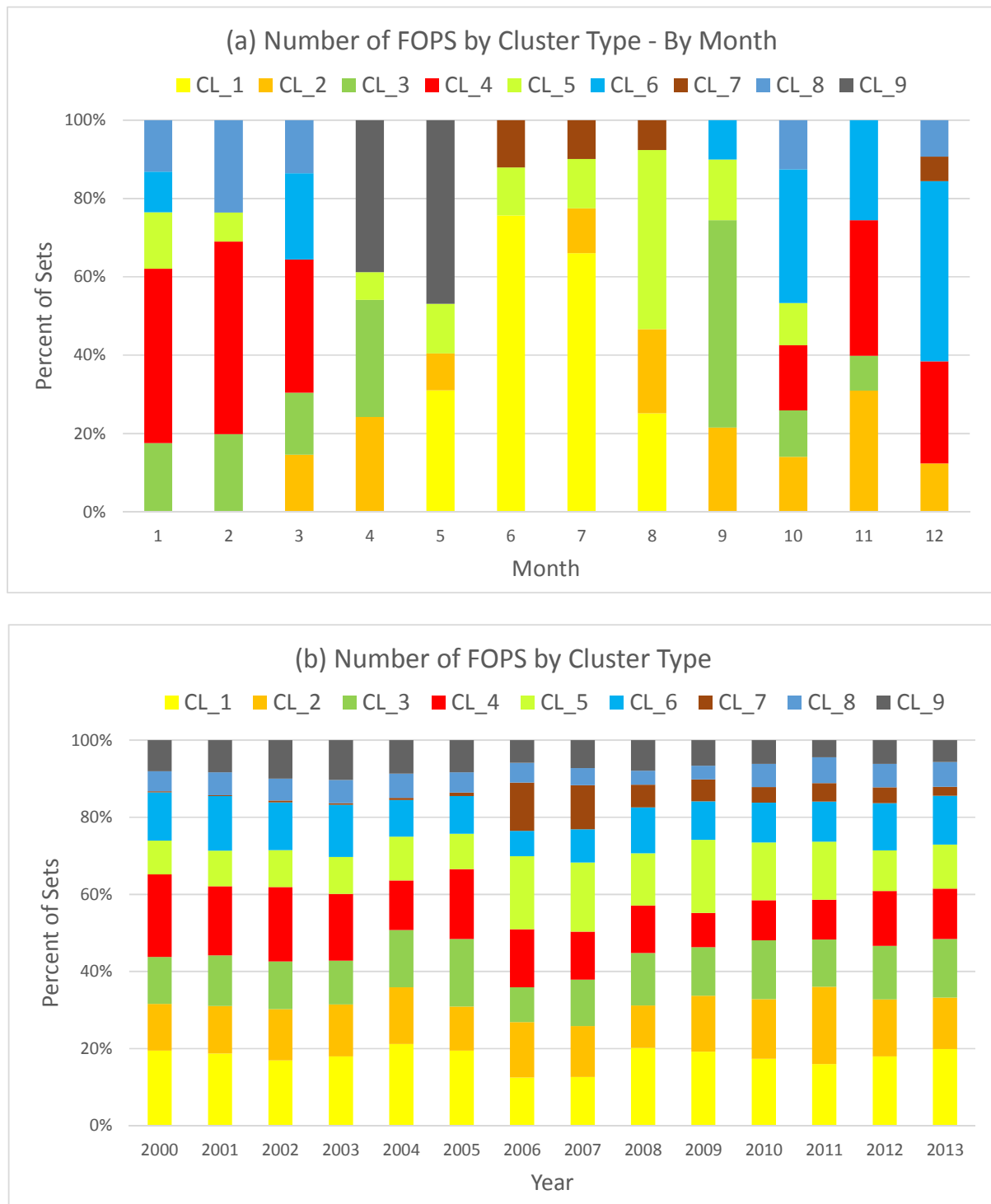


Figure C.5: Distribution of gear-settings across the nine gear-based cluster types identified in the 2-stage monthly analysis.

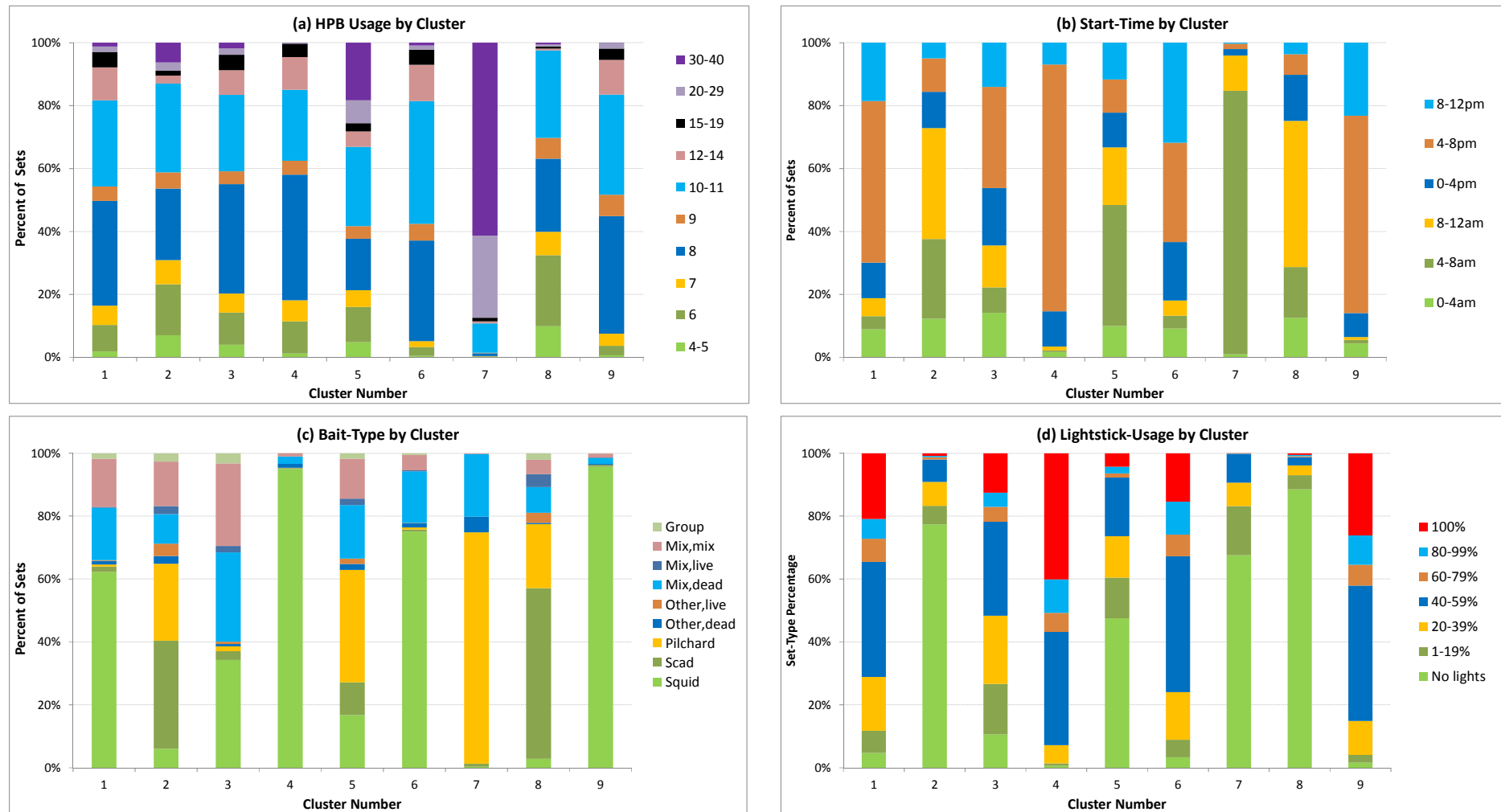
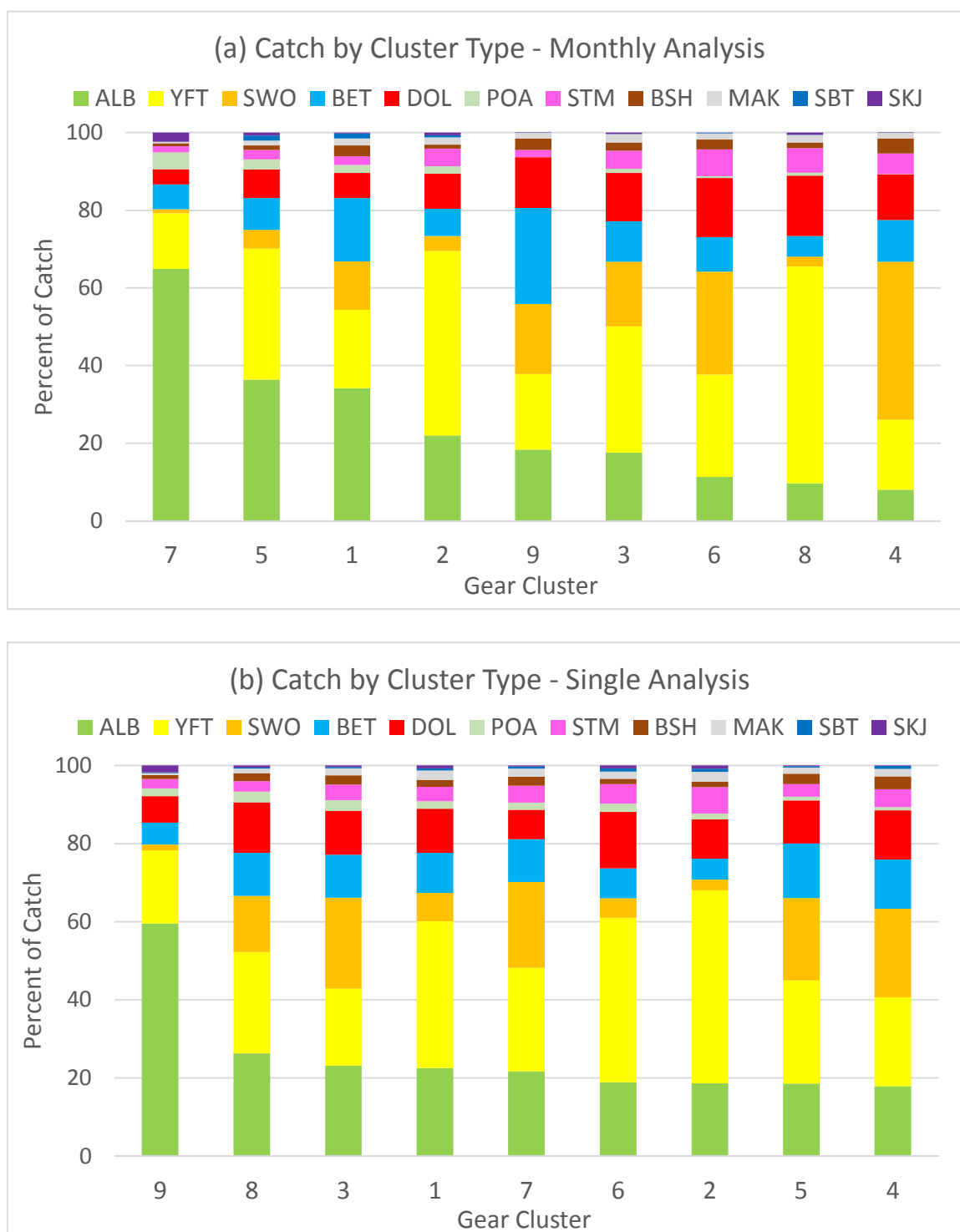


Figure C.6: Catch composition of the nine clusters based on (a) the analysis by month, and (b) the single analysis of all sets.



Appendix D. Spatial Dynamics Factor Analysis Model

SDFA method

We explored the potential of a method known as Spatial Dynamic Factor Analysis (SDFA) (Thorson *et al.* 2016) to provide improved indices of abundance. The approach simultaneously models several effects on catch rates in a statistically coherent way, and has the potential to provide robust, precise, and reliable indices of abundance. Issues that affect catch rates and which are included in the model include:

- i) spatial variation in the density of each species;
- ii) auto-correlated spatiotemporal variation for each species, since species distributions change through time;
- iii) correlations in spatiotemporal variation among species, such as those caused by species having similar habitat requirements; and
- iv) correlated catch rates among species arising from the behaviour of the fisheries.

The method explains species density as a log-linear combination of factors, in which each factor represents unobserved spatial variation in density for a species assemblage.

The SDFA method is very different from the two-stage clustering-based models, otherwise used in this project. The SDFA analysis is done in a single model, so variation in one part of the model (e.g. time trends in abundance) is allowed for elsewhere in the model (e.g. in the allocation of effort among fishing strategies).

Thorson *et al.* (2016) report the following advantages for the SDFA method:

1. Uncertainty in the identification of fishing tactics is propagated throughout the analysis, rather than being ignored in the regression analysis.
2. The design of the overall model and the interactions among components are clearly expressed.
3. The number of subjective decisions is reduced, such as the number of clusters to use, the choice of data transformation when estimating fishing tactics.
4. Changes in relative abundance among species are accounted for in the estimation of fishing tactics.

However, there are also disadvantages. The model is computationally demanding, which limits the ability to take covariates into account in the large datasets used for CPUE standardization. In addition, the approach is new (published in April 2016) and still under development, and software changes occur regularly. For example, the SpatialDFA package is no longer being actively maintained and a new package VAST is recommended instead (Jim Thorson, personal communication). VAST can do almost everything that the SpatialDFA package does, but it may run slower, and was not tested here. Code for the SDFA model itself was obtained from the SpatialDFA package in the github repository https://github.com/James-Thorson/spatial_DFA, and other code was based on the example analyses therein.

The aim of the analyses undertaken and outlined below was to explore the viability of the approach given the large size of the ETBF datasets and the number of species involved; to determine RAM and time requirements for analysis, and to see whether the results were reasonable. We did not aim to optimise the approach by including covariates known to be important, such as gear factors.

Methods

We conducted analyses using the SDFA model. For a full description of the approach see Thorson *et al.* (2016).

Paraphrasing Thorson *et al.* (2016), SDFA models the catch for species s based on the catch equation $c_s = q_s f_s d_s$, where q is catchability, f is nominal fishing effort after accounting for small-scale targeting, and d is density. Density, fishing mortality and catchability are modelled as follows:

$$\begin{aligned}\log(d_s(a, t)) &= \sum_{j=1}^J A_{s,j} \psi_j(a, t) + \sum_{l=1}^L \gamma_{s,l} x_l(a, t) \\ \log(f_s(a, t)) &= \sum_{k=1}^K B_{s,k} \epsilon_k(i) \\ \log(q_s(a, t)) &= \sum_{l=1}^L \nu_{s,l} y_l(i)\end{aligned}$$

$A_{s,j}$ is an S by J matrix of factors on log-density, $\psi_j(a, t)$ is the value of factor j in area a at time t , x_l is the l -th measured variable at site s and time t , and $\gamma_{s,l}$ is the effect of covariate l on density for species s . $B_{s,k}$ is an S by K matrix representing the impact on species s of small-scale variables $\epsilon_k(i)$, such that $\epsilon_k(i)$ is one or more latent variables representing variation in fishing tactics. $\nu_{s,l}$ is the effect of covariate l on catch rates from species s , and $y_l(i)$ is the l -th measured variable affecting catch rates for fishing occasion i . Residual variation follows the zero-inflated Gamma distribution.

Analyses were carried out for the following datasets:

- i) the reduced simulated ETBF dataset (N3) with 31,959 rows of data, using the deterministic catches;
- ii) the full simulated ETBF dataset with 113,711 rows of data, using the deterministic D-catches;
- iii) the full simulated ETBF dataset, using 10 realisations of the randomized R-catches (called hear-after the R10 results).

Set locations, specified in the dataset by latitude and longitude, were modelled with a 139 location mesh using INLA. Data were entered into the model as catch rates (number per hook per set). The number of spatial factors J , which represent unobserved spatial variation in density for a species assemblage, was set to 4. The number of targeting factors K , which represent small-scale spatiotemporal variation in targeting of species assemblages, was set to 3.

To increase computation speed, covariates that may affect catchability, such as HBF and set time, were not included in the standardization. Time variation was annual, and quarterly variation in catchability or spatial effects was not included. Models were run using Microsoft R Open with R version 3.3.2 and a CRAN mirror snapshot from 1 November 2016. We used the Intel Math Kernel Library for parallel mathematical computing. The processor was Intel core i7-6700HQ with 64 GB RAM, running 64 bit Windows 10 Pro.

Results

All models ran successfully. The model for the reduced N3 deterministic dataset ran for 1.4 hours, while the full deterministic ETBF dataset run took 6.5 hours. The full random ETBF datasets took similar times, averaging 6.7 hours and with a range from 5.4 to 7.3 hours. Memory use was reasonable with up to 6.8 GB per model run for the full dataset, varying through each run. In the following we only report those results fitted to the full simulated EBF dataset.

A comparison of the annual abundance indices based on the SDFA method with the assumed and nominal model indices is shown in Figure D.1. Performance of the SDFA method fitted to the 10 realisations of the randomized R-catches in comparison with the Group-A models (fitted to the same set of 10 realisations) varied by species (Figure D.2). The SDFA method had the second-best performance for YFT, but worst for DOL and (particularly) SWO. Averaging errors across all species gave the lowest rank to the SDFA method for each error type (Figure D.3). Similar results (not shown) were obtained for the performance of the SDFA method fitted to the single realisations of the deterministic D-catches in comparison with the Group-A models, though the relative performance of the SDFA model was found to be best (4th) for ALB.

Figure D.1: Comparisons by species of annual indices between true, nominal, and SDFA model results. The SDFA-R10 model indices are the means of the 10 model runs using random data, while the SDFA-D model indicates the index for the single run using the deterministic catches.

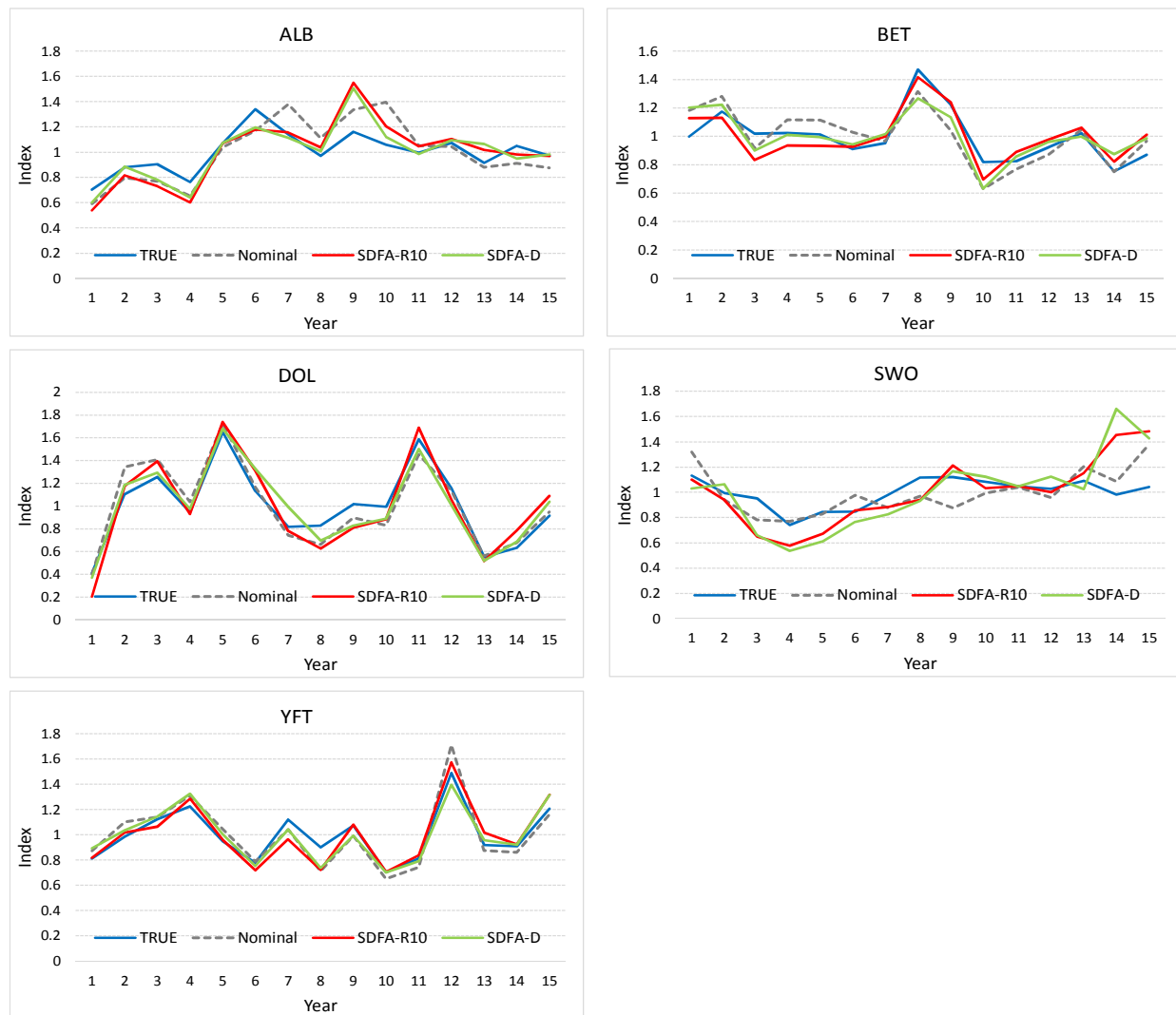


Figure D.2: Comparisons among the Group-A and SDFA models by species, for the four error types. All results based on fitting each model to the 10 realisations of the ETBF dataset using the random catches.

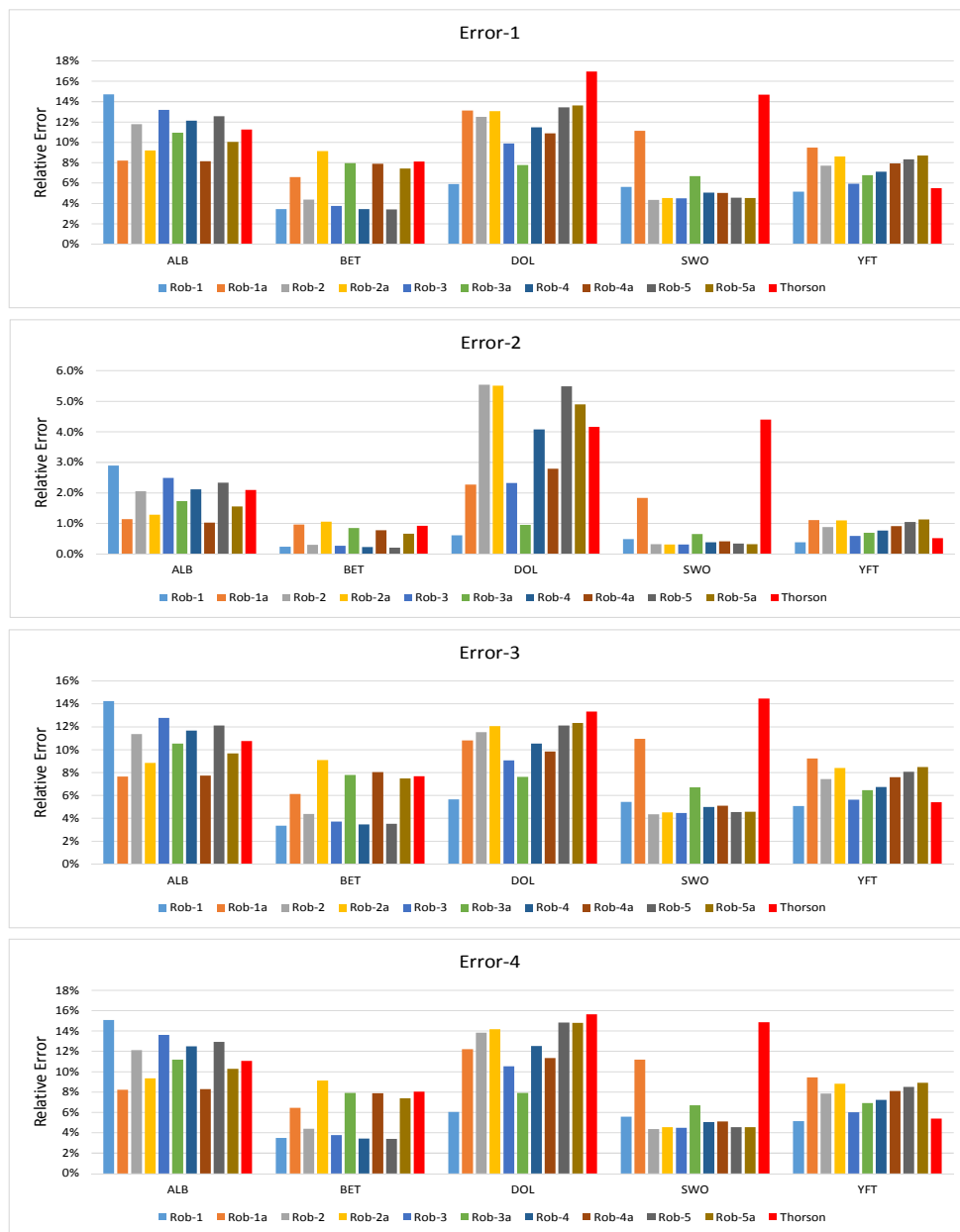
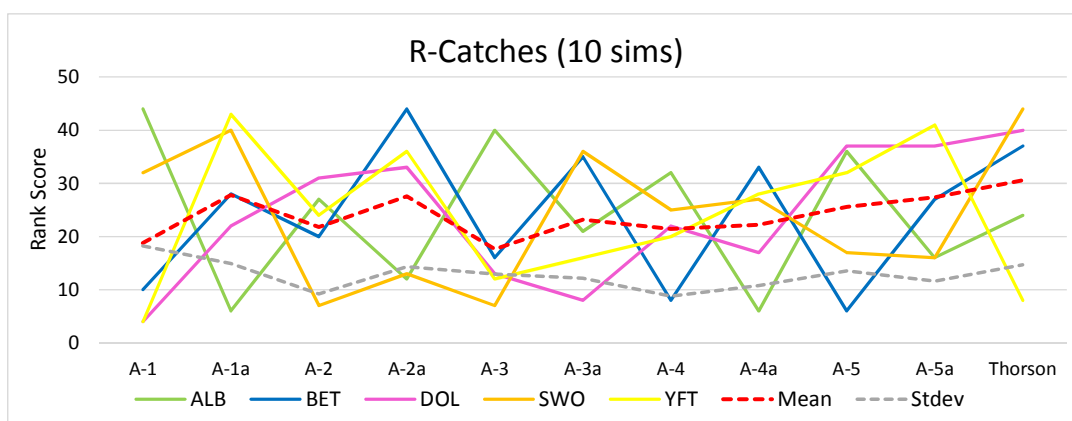


Figure D.3: Comparison of the Species Ranked Scores (SRS) for the Group-A and SDFA models by species, when fitted to the 10 realisations of the ETBF dataset using the random catches.



Discussion

The SDFA method appears promising. Results were generally worse than those for the Group-A models, indicating that the approach, as implemented here, requires more development before it can be used for management advice. Resource use was less than we expected with analyses completing in about 6 hours for a large dataset of over 100,000 records.

However, the worse model performance may be caused more by aspects of the model implementation than by the modelling approach itself. The SDFA model structure differed from the Group A models in several important ways, quite separate from the major difference in statistical approach. Seasonal variation (*Qtr* and its interactions with *Year* and *Area*) was not accounted for, with *Year* the only modelled time effect. Seasonal effects are important for these species. The gear effects *HPF* and *STIME* were also omitted, so that the structure of the model was effectively $CPUE \sim Year * Area$. However, gear effects were omitted in the Group A models with subscript *a*, and these models also outperformed the SDFA model. The SDFA model also has a more complex spatial representation than the standard models, and incorporates time-area effects.

Our limited exploration of this new modelling approach suggest that it has potential, and we recommend further work to explore and develop its capabilities.

Code

```
# File structure
TmbFile = system.file("executables", package="SpatialDFA")
basedir <- "~/SDFA_memtests/"
setwd(basedir)
DateFile = paste(getwd(), '/', Sys.Date(), '/', sep='')
dir.create(DateFile)
setwd(DateFile)

# Settings
Version = "spatial_dfa_v18"

# Settings
Nfactors = 4 # spatial factors
Nobsfactors = 3 # targeting factors

# Libraries
library( INLA )
library( TMB )

# Specific libraries
library( SpatialDFA )
setMKLthreads(1)

# Compile TMB model
setwd( TmbFile )
compile( paste0(Version, ".cpp") )

# Data
dssets <- list.files(path=paste0(basedir, "../Robs100datasets/"), pattern="datagen_", full.names = TRUE)
for(ds in 1:3) {
  nm <- read.table( dssets[ds], skip=5, nrows=1, stringsAsFactors = FALSE)
```

```

Data = read.table( dsets[ds],header=FALSE,col.names=nm,skip=7 )
#Data = read.table( paste0(basedir,"../Robs100datasets/datagen_1.txt"),header=TRUE )
Data = Data[,1:23]
Data = cbind(Data, 'Lat'=Data[, 'LAT'], 'Lon'=Data[, 'LON'])
# lat_set
Data$Lat = Data$Lat - 12
Data$Lon = Data$Lon - 144
latlong_set = unique(paste( Data[, 'Lat'], Data[, 'Lon'], sep="_"))
lat_set = as.numeric(sapply( latlong_set, FUN=function(Char){strsplit(Char,"_")[[1]][1]}))
long_set = as.numeric(sapply( latlong_set, FUN=function(Char){strsplit(Char,"_")[[1]][2]}))
loc_xy=data.frame(x=lat_set,y=long_set)
# Rename stuff
Data$FOP_ID <- factor(Data$FOP_ID)
YRs = sort(unique(Data[, 'YEAR']))
DF_blank = data.frame( 'sitenum'=match(paste( Data[, 'Lat'], Data[, 'Lon'], sep="_"),latlong_set), 'spp'=NA,
'year'=match(Data[, 'YEAR'],YRs), 'catch'=NA, 'TowID'=Data[, 'FOP_ID'])
# DF_blank = data.frame( 'sitenum'=match(paste( Data[, 'Lat'], Data[, 'Lon'], sep="_"),latlong_set),
'spp'=NA, 'year'=Data[, 'yr'], 'catch'=NA, 'TowID'=Data[, 'X']) # try this later
DF = NULL
for( i in 1:5 ){
  Temp = DF_blank
  Temp[, 'spp'] = c("ALB_R", "BET_R", "DOL_R", "SWO_R", "YFT_R")[i]
  Temp[, 'catch'] = Data[Temp[, 'spp'], 'catch'] / Data[, 'HOOKS']
  DF = rbind(DF, Temp)
}
DF[, 'spp'] = as.factor(DF[, 'spp'])
EncounterFunction = 1
estimation_method = c("mesh", "grid")[1]

## Build SPDE object using INLA
mesh = inla.mesh.create( cbind(long_set, lat_set), plot.delay=NULL, refine=FALSE ) # loc_samp ;
,max.edge.data=0.08,max.edge.extra=0.2
plot.inla.mesh(mesh)

# Bundle inputs
if(estimation_method=="grid") {
  InputList = MakeInput_Fn( Version=Version, DF=DF, Nfactors=Nfactors, Nobsfactors=Nobsfactors,
loc_xy=loc_xy, EncounterFunction=EncounterFunction,method=estimation_method )
} else {
  InputList = MakeInput_Fn( Version=Version, DF=DF, Nfactors=Nfactors, Nobsfactors=Nobsfactors,
loc_xy=loc_xy, inla_mesh=mesh, a_n=rep(1,mesh$n), EncounterFunction=EncounterFunction )
}
InputList$TmbData$a_n[InputList$TmbData$a_n==0] <- 1
InputList$Map$gamma_ptl = NULL

# Link TMB
dyn.load( dynlib(Version) )

# Initialization
obj <- MakeADFun(data=InputList[["TmbData"]], parameters=InputList[["TmbParams"]],
random=InputList[["Random"]], map=InputList[["Map"]], hessian=FALSE, inner.control=list(maxit=1000) )
obj$control <- c( obj$control, list(trace=1, parscale=1, REPORT=1, reltol=1e-12, maxit=100) )
obj$env$inner.control <- c(obj$env$inner.control, list("step.tol"=1e-8, "tol10"=1e-6, "grad.tol"=1e-8) )

```

```
# Bounds
Upper = rep(Inf, length(obj$par) )
Upper[grep("rho_",names(obj$par))] = 0.99
Lower = rep(-Inf, length(obj$par) )
Lower[grep("rho_",names(obj$par))] = -0.99

# Run model
# Rprof(tf <- "rprof.log", memory.profiling=TRUE)
Start_time = Sys.time()
for(i in 1:2) opt = nlminb(start=obj$env$last.par.best[-c(obj$env$random)], objective=obj$fn,
gradient=obj$gr, upper=Upper, lower=Lower, control=list(eval.max=1e4, iter.max=1e4, trace=1, rel.tol=1e-
14) )
opt[["final_gradient"]] = obj$gr( opt$par )
opt[["total_time_to_run"]] = Sys.time() - Start_time
print(opt[["total_time_to_run"]])
Report = obj$report()
SD = sdreport( obj, bias.correct=FALSE )

# Loadings matrix
L_pj = Report$L_pj
dimnames(L_pj) = list(levels(DF[, 'spp']), paste("Factor", 1:Nfactors))

# Extract factors
Psi = Report$psi_njt

# Rotate
if(Nfactors>1){
  RotateList = Rotate_Fn( L_pj=L_pj, Psi=Psi, RotationMethod="PCA", testcutoff=1e-5 )
  L_pj_rot = RotateList[["L_pj_rot"]]
  Psi_rot = RotateList[["Psi_rot"]]
}else{
  L_pj_rot = L_pj
  Psi_rot = Psi
}
opt[["total_time_to_run_plus"]] = Sys.time() - Start_time

# Save stuff
Save = list("opt"=opt, "Report"=Report, "Sdreport"=SD, "ParHat"=obj$env$parList(opt$par),
"TmbData"=InputList[["TmbData"]])
save(Save, file=paste0(DateFile,"Save",ds,"_",estimation_method,".RData"))
capture.output( opt, file=paste0(DateFile,"Opt",ds,"_",estimation_method,".txt"))
capture.output( SD, file=paste0(DateFile,"Sdreport",ds,"_",estimation_method,".txt"))
}
```

Appendix E. List of Project Staff

Dr Robert Campbell

CSIRO Oceans and Atmosphere, Private Bag No. 1, Aspendale, VIC 3195, Australia

Principal Investigator, Data Manager and Analyst

Dr Shijie Zhou

CSIRO Oceans and Atmosphere, GPO Box 2583, Brisbane, QLD 4001, Australia

Analyst

Dr Simon Hoyle

Hoyle Consulting, 20 Bisley Ave, Nelson, 7011, New Zealand

Analyst

Dr Rich Hillary

CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, TAS 7001, Australia

Analyst

Dr Malcolm Haddon

CSIRO Oceans and Atmosphere, GPO Box 1538, Hobart, TAS 7001, Australia

Contributed to review and development of methodologies

Mr Steve Auld

Department of Agriculture and Water Resources, GPO Box 858, Canberra ACT, Australia

Contributed to review and development of methodologies and provided expertise on ETBF

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help
improve today – for our customers, all
Australians and the world.

Our innovations contribute billions of
dollars to the Australian economy
every year. As the largest patent holder
in the nation, our vast wealth of
intellectual property has led to more
than 150 spin-off companies.

With more than 5,000 experts and a
burning desire to get things done, we are
Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE.
WE INNOVATE.

FOR FURTHER INFORMATION

Oceans and Atmosphere

Dr Robert Campbell

t +61 3 9239 4681
e Robert.Campbell@csiro.au
w www.csiro.au/en/Research/OandA

Oceans and Atmosphere

Dr Shijie Zhou

t +61 7 3833 5968
e Shijie.Zhou@csiro.au
w www.csiro.au/en/Research/OandA