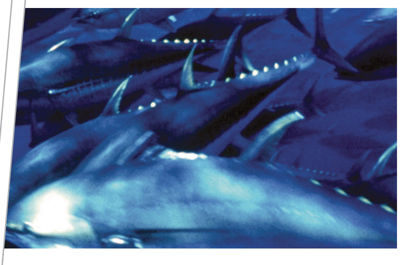
National Research
FLAGSHIPS
Wealth from Oceans

CSIRO



www.csiro.au

Adaptive frameworks for Australian fishery observer programs

Wealth from Oceans National Research Flagship

FRDC Project 2005/002 • May 2010

• Christian H. Moeseneder • Jeffrey M. Dambacher
• Robert A. Campbell • Fiona M. Giannini
• Emma K. Lawrence • Alistair J. Hobday • Jock W. Young

Australian Government
Bureau of Rural Sciences

Australian Government
Fisheries Research and Development Corporation

Adaptive frameworks for Australian fishery observer programs

Christian H. Moeseneder
Jeffrey M. Dambacher
Robert A. Campbell
Fiona M. Giannini
Emma K. Lawrence
Alistair J. Hobday
Jock W. Young

Project No. 2005/002

CSIRO

Australian Government
Bureau of Rural Sciences

Australian Government
Fisheries Research and Development Corporation

Copyright and disclaimer

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.

Authors: Christian H. Moeseneder, Jeffrey M. Dambacher, Robert A. Campbell, Fiona M. Giannini, Emma K. Lawrence, Alistair J. Hobday, Jock W. Young

Title: Adaptive frameworks for Australian fishery observer programs

Includes: contents, bibliography, appendices.

Subjects: Bycatch, Eastern Tuna and Billfish Fishery, Fishery logbook, Fishery observer program

[Cite as: Moeseneder, C. H., J. M. Dambacher, R. A. Campbell, F. M. Giannini, E. K. Lawrence, A. J. Hobday and J. W. Young. 2010. Adaptive frameworks for Australian fishery observer programs. FRDC Final Report 2005/002. 110 pp.]

2005/002 - Adaptive frameworks for Australian fishery observer programs

Principal Investigator: Jeffrey M. Dambacher

Address: CSIRO Mathematics, Informatics and Statistics
GPO Box 1538
Hobart, Tas. 7001, Australia
Telephone: 03 6232 5096 Fax: 03 6232 5000

Objectives

1. Develop software tools and analytical framework to support AFMA management decisions and research projects with streamlined and timely analyses of fishery observer data.

2. Develop statistical and organizational frameworks to assist in the allocation of observer effort within Australian fisheries.

3. Develop guidelines for sufficient and appropriate methods of analysis and reporting of AFMA observer data.

4. Develop an understanding of the relationship between catch rates of selected bycatch species, fishing gear and practices, and surrounding oceanographic environments.

5. Develop performance indicators and predictive models to support management of bycatch and discard mortality for sensitive or endangered species.

6. Assess the quality of AFMA logbook data for bycatch species.

Non technical summary

The Australian Fisheries Management Authority (AFMA) maintains fishery logbook and observer databases for many of Australia's commercial fisheries. And while these databases are regularly analysed for specific management purposes, recent expansion in observer programs and an increased need for management of catch and bycatch species has created a demand for a standardized analysis approach that meets the needs of a wide range of stakeholders and management applications at a national level.

This research project has addressed this need for the Eastern Tuna and Billfish Fishery (ETBF) by providing an analytical framework to facilitate quality control of the databases and, through a standardized reporting system, providing feedback to fishery managers and stakeholders (Figure 1).

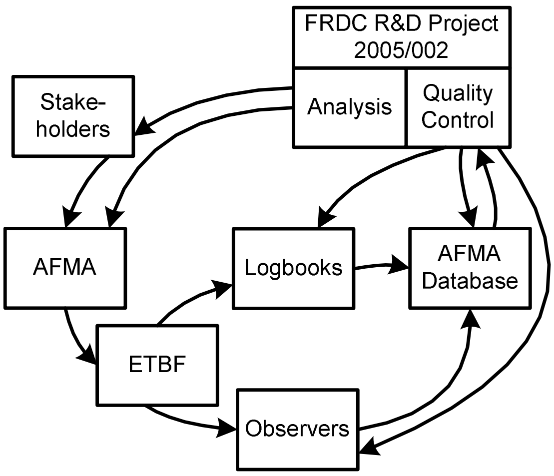


Figure 1. Schematic of feedbacks between research and development (R&D) project and monitoring and management of the Eastern Tuna and Billfish Fishery (ETBF).

A basic issue that this project has dealt with is the provisioning of standardized reports of catch and bycatch species in the ETBF that meet the common needs of management, industry and stakeholder groups. Additionally, this project has attempted to better understand the relationships between catch rates of species, fishing gear and practices, and the surrounding ocean environment, with the goal to developing management options to minimize bycatch. This project also provides the basis to better distribute observers throughout the ETBF.

OUTCOMES ACHIEVED

This project has created standardized summaries and analysis of oberserver and logbook data for the ETBF. An analytical framework was developed that allowed for the rapid provision of estimates of catch and bycatch for the ETBF, which has been used on an annual basis for RAG meetings and National Fishery Report for the Scientific Committtee of the WCPFC.

We created a centralised database and flexible software interface for integrated queries and analyses of observer data, thus permitting regular and ad hoc summaries of observer data to AFMA, DEWHA, Research Advisory Groups and the Ecological Risk Assessment Team. Queries are designed to permit summaries based on selected areas and times, thus allowing for specific cross referencing to existing dependent data sets.

Our reporting system detects data errors in a semi-automated process which has enabled a more streamlined and rapid method of preparing raw AFMA data for analysis. Moreover, this project identified ways to improve the usefulness of observer and logbook databases methods. While this information was conveyed to AFMA, it remains to be seen if these suggestions will be implemented in their planned Data Warehouse and Reporting System.

Our reporting system produces statistical estimates of the catch and for all species caught in the ETBF. Additionally, it gives the fate and life status of all species encountered in the fishery. This information is particularly relevant and fundamental to reoccurring discussions at national and international fishery management meetings.

The provision of statistical estimates of bycatch provides the basis for management discussion on the conservation of threatened and endangered species. Our analysis of associations between targeted and bycatch species, as well as associations with environmental variables can be used to inform and thus improve mitigation efforts to minimize bycatch of selected species.

Our reporting system provides an analysis of observer coverage in the ETBF, and therefore creates the basis for improved allocation of observer effort in the fishery. The statistical analyses of catch and bycatch can be used to guide an adaptive approach to allocation of observer effort across the ETBF, this final step however, requires completion of, and integration with, the AFMA Data Warehouse and Reporting system, which is not yet completed.

Keywords

Bycatch, Eastern Tuna and Billfish Fishery, Logbook, Observer program

Acknowledgements

We thank members of the ETBF RAG for helpful comments and criticisms.

Background

Multiple-use management of marine resources requires an understanding of the effects of fishing on ecosystem integrity, including its effects on ecological processes, such as trophic interactions and biodiversity (Sainsbury et al. 1997). Fishing affects target species, but it can also affect non-target species, both directly and indirectly. Management regulations often require fishers to report their daily activities and catches in logbooks. However, logbooks need to be independently verified. Furthermore, commercial fishing operations focus on maximizing profits, and crewmembers cannot be expected to collect reliable data on non-target species, discard levels or fish size. The deployment of independent observers is therefore a key component of multiple-use or “ecosystem-based” approaches to fisheries management.

Observer programs play an integral role in the provision of data to fisheries research and improved scientific advice to fisheries stakeholders and managers. For example, observer-verified catch and effort data plays an important role in undertaking stock assessments, and also provides information on fishing strategies and targeting practices required for standardising catch per unit effort. Observer data is often the only source of information available for management of bycatch issues. Bycatch issues include estimation of the number of bycatch species caught, interactions with protected species, the survival of bycatch that are released from the gear, and the level of discarding that occurs.

Although the Resource Assessment Group (RAG) for the ETBF has long identified the need for an effective observer program for the ETBF, in response to Australia’s Tuna Fisheries Bycatch Action Plan (BAP, AFMA 2001) they commissioned a report on what level of observer coverage would be required to meet the BAP objectives (Bravington et al. 2002). After consideration of this report by the RAG and the Australian Fisheries Management Authority (AFMA), an observer program was commenced in July 2003. This program is to run for an initial period of 5 years and has an intended coverage rate of 5.1% of annual effort.

Concomitant with the above actions, the Ecological Risk Assessment for Effects of Fishing Project (ERA) has also highlighted a large number of gaps in the data and information available for assessing the sustainable nature of fishing practices (Hobday et al. 2004, 2007). These data gaps include a lack of information on the biology of many of the species caught and a lack of data on the catch, life-status and fate of a large number of by-product and by-catch species not recorded in logbooks, especially those that are sensitive or endangered (e.g., sea birds and turtles). A primary output of the ERA project to date has been a classification of the risk profile for all species, and it is important that risk profiles be reassessed as new information is obtained, this reassessment is now a RAG responsibility.

Australia also has international obligations to develop management plans regarding a number of species (e.g., seabirds, sharks, southern bluefin tuna). For example, in recent years AFMA has undertaken a number of trials off NSW aimed at reducing the incidental catch of sea-birds and has restricted access to the southern parts of the ETBF to limit bycatch of southern bluefin tuna. The continued collection and analysis of data pertinent to these and other areas issues is vital if Australia is to continue to meet these obligations.

Finally, the Fourth International Fisheries Observer Conference, held in Sydney in 2004, raised the profile of Australia's observer programs, and there is an evolving and growing debate regarding analytical techniques and adequate levels of observer coverage to monitor the catch of rare species.

Need

Although each year millions of dollars are spent observing bycatch and assessing bycatch mitigation across Australian Commonwealth fisheries, there currently is no comprehensive analysis of observer data. For example, AFMA's observer section compiles reports on seabird Threat Abatement Programs (TAP) containing summaries for numbers of birds caught, effort, and catch rates. While these statistics satisfy TAP reporting requirements, they nonetheless lack analysis of variation in catch rates over time, space, and among vessels. Hence understanding of why mitigation methods may or may not be working is hampered, ultimately to the detriment of industry. Also, past analyses of observer data (Dambacher et al. 2003) have demonstrated a potential for predictive models based on fishing methods and oceanographic conditions to support management strategies minimising discard mortality and bycatch. Unfortunately such analyses are beyond the current scope of AFMA's observer section, and a need exists for managers to receive and respond to information coming from observer programs in a timely manner. The need for streamlined analyses of observer data is not limited to bycatch issues alone, but extends to various management and research areas including ecological risk assessments for judging ecological sustainability, and stock assessments requiring observer data to develop calibrated analyses of standardized catch per unit effort.

A second driver for this project is that while most fishery stakeholders agree on the need for observer programs, disagreement exists on appropriate levels of observer effort. NGO's cite literature from within the NGO community (Babcock et al. 2003) that coverage levels exceeding 50% are required to rigorously estimate bycatch, while the fishing industry argues that the capacity to pay limits what is possible. Clearly this isn't a case of "one-size-fits-all", and decisions on levels of observer coverage need to be framed against an array of what will sometimes be competing management objectives. We contend that this debate will benefit from a more rigorous statistical approach. Ultimately, managers require an adaptive approach that is both practical and transparent to the trade-offs involved. This project set out to address these needs by developing a reporting system to provide analyses of AFMA observer data to support Australian fisheries management.

Objectives

1. *Develop software tools and analytical framework to support AFMA management decisions and research projects with streamlined and timely analyses of fishery observer data.*

   Achieved fully as stated.

2. *Develop statistical and organizational frameworks to assist in the allocation of observer effort within Australian fisheries.*

   This objective was completed in part, but due to unforeseen circumstances with AFMA's database, we are unable to complete it fully. A reporting system was developed to describe the distribution of observer coverage within the ETBF relative to fishing effort and catch of species of interest. Embedding more refined analyses within AFMA's organizational framework required our analytical tools to be integrated with AFMA's planned transaction database and data warehouse. Unfortunately the database and warehouse are not operational. Moreover, AFMA has made substantial changes to the structure of their databases, and presently, the new database structure is incompatible with our reporting system, and presently the reporting system we developed is nonoperational.

3. *Develop guidelines for sufficient and appropriate methods of analysis and reporting of AFMA observer data.*

   Achieved fully as stated. The original project proposal anticipated that there would be difficulty in establishing concensus among the various stakeholders regarding the design of the analysis and reporting framework. This objective was designed to formulate an agreed reporting framework based on formal guidelines. As it turned out, the first generation of reporting developed by the Principal Investigator was relatively close to the mark, and required only minor revision, as critiqued and guided by stakeholders at a number of annual meetings of the ETBF RAG.

4. *Develop an understanding of the relationship between catch rates of selected bycatch species, fishing gear and practices, and surrounding oceanographic environments.*

   Mostly achieved, but data gaps for fishing gear and practices precluded these variables from analysis.

5. *Develop performance indicators and predictive models to support management of bycatch and discard mortality for sensitive or endangered species.*

   Partially achieved. Performance indicators developed but no reliable predictive models were found.

6. *Assess the quality of AFMA logbook data for bycatch species.*

   Achieved fully as stated. Also extended assessment to include observer data.

Methods

Reporting System

The core objective of this project was to develop a standardised report which would meet stakeholder needs. Thus we consulted with the ETBF RAG as well as other interested parties (i.e., individuals from the BRS, DEWHA and CSIRO). Feedback from workshops and individual peer review was used to structure and improve the reporting system.

The methods and tools employed were standard, off-the-shelf, software products as well as statistical theory and design. Microsoft (MS) Access was used for data storage, manipulation and report generation. Reports were driven by MS Access queries and code written in Visual Basic for Applications (VBA). ArcGIS 8.3 was used 1) for spatial display of logbook and observer records, 2) detection of outlying data records, and 3) developing and verifying average shot locations. A MS-SQL 7.0 database was developed for final data storage and subsequent web-enabled report generation, but not implemented due to incompletion of AFMA's Database Warehouse and Reporting System.

Intitially, the reporting system was created as an ad-hoc solution in an MS Excel file by the Principal Investigator. To provide the capability of automated report creation, an MS Access database was created which contains all objects that are necessary for execution of the reports. There are links to the original AFMA observer and logbook data which is uploaded for each new season into CSIRO's Oracle database. The Word document "ETBF data work Version 2.doc" (Appendix 3) contains detailed instructions in step-by-step form which allows a non-techncial user to complete the entire reporting process. Generally, either queries or VBA procedures are used to create intermediate data tables that contain the result of table joins or summaries of the source tables. All final tables are based on a multitude of such underlying queries and tables.

The reporting system's main MS Access database (ETBF.mdb, Supplement 1-compact disk) has 83 tables, including temporary tables, 83 queries and 15 VBA functions. Its size is approximately 570 MB. The system's GIS database (ETBF_GIS.mdb, Supplement 1-compact disk) has 80 tables and is approximately 180 MB. The number of tables in the database increases as data is added for each new season.

All queries and tables can be viewed in SQL-mode which produces a standard SQL statement that can be copied to SQL-Server 7.0 or an Oracle database system in order to recreate the reporting system. This enables AFMA to generate reports from their planned Data Warehouse and Reporting Sytem. There are a number of web-enabled front-end solutions that provide data entry capabilities to allow a user to enter time periods and spatial boundaries that are then written to the reporting database and executed.

The statistical framework for analysis of observer and logbook data was based on standard statistical analyses (Cochran 1977, Box 1). The point estimate of catch was determined from a ratio of the observed catch and the observed number of hooks in a sampling stratum multiplied by the total number of hooks, as reported from fisherman logbooks. Calculation of 95% confidence intervals for the point estimates was performed by two separate methods, one based on simple random sampling estimator of variance, and the other on a cluster estimator of variance, which treated all sets within a fishing trip of a particular boat as a sampling cluster.

Ratio estimate (adapted from Eq. [6.1] Cochran 1977)

$$\widehat{Y_R} = \frac{y}{h} H$$

$\widehat{Y_R}$ = ratio (R) estimate of Y, the total catch in a strata
y = observed total catch in a strata
h = observed total hooks in a strata
H = total no. of hooks in a strata (from log book data)

Simple random sampling (SRS) estimate of variance (adapted from Eq. [6.9] Cochran 1977)

$$\widehat{\text{var}_{\text{SRS}}}\left(\widehat{Y_R}\right) = N \underbrace{\frac{(N-n)}{n}}_{\text{finite population correction factor}} \underbrace{\frac{\sum_{i=1}^{n}\left(y_i - \frac{y}{h}h_i\right)^2}{n-1}}_{\text{sample variance } (s^2_{\text{SRS}})}$$

$$= \frac{N(N-n)}{n\,(n-1)} \sum_{i=1}^{n}\left(y_i - \frac{y}{h}h_i\right)^2$$

$\widehat{\text{var}_{\text{SRS}}}$ = estimate of variance for simple random sampling
n = observed no. of sets in a strata
N = total no. sets in strata (from logbook data)
y_i = observed catch of a species in a set
h_i = observed no. of hooks in a set
i = individual observed set

Cluster sample (CS) estimate of variance

$$\widehat{\text{var}_{\text{CS}}}\left(\widehat{Y_R}\right) = M \underbrace{\frac{(M-m)}{m}}_{\text{finite population correction factor}} \underbrace{\frac{\sum_{g=1}^{m}\left(y_g - \frac{y}{h}h_g\right)^2}{m-1}}_{\text{sample variance (CS)}}$$

$$= \frac{M(M-m)}{m\,(m-1)} \sum_{g=1}^{m}\left(y_g - \frac{y}{h}h_g\right)^2$$

$\widehat{\text{var}_{\text{CS}}}$ = estimate of variance for cluster sampling
m = observed no. of trips in a strata
M = total no. of trips in strata (from logbook data)
y_g = observed catch of a species in a trip
h_g = observed no. of hooks in a trip
g = individual observed trip (i.e., cluster)

Coefficient of variation (CV) of sample

$$\text{CV}_{\text{SRS}} = \frac{\sqrt{s^2_{\text{SRS}}}}{\left(\frac{y}{n}\right)} \qquad \text{CV}_{\text{CS}} = \frac{\sqrt{s^2_{\text{CS}}}}{\left(\frac{y}{m}\right)}$$

Box 1. Formula used for estimation of catch and bycatch in the ETBF.

Bycatch modelling

A number of approaches were taken by Giannini and Lawrence in their analysis of bycatch in the ETBF (Appendix 4), with the goal of better understanding the relationship between target species, bycatch species, and oceanographic variables. In exploratory work they employed cluster analysis and principle components analysis (PCA) in an attempt to reveal underlying structure in catch data in terms of significant or commonly encountered species associations. In separate modelling work, they analysed the relationship between species catch and oceanographic variables, and relationships between the presence of bycatch and target species. Here they used two different approaches. In one, they modelled the presence-absence of target species using the catch of bycatch species as explanatory variables in an attempt to better quantify the relationship between the two groups. A second approach involved modelling the catch of a select group of bycatch species in terms of oceanographic variables in order to study the effect of different environmental conditions on the catch of these species.

Data processing

A first generation analysis of the ETBF data was used as the basis for creating the ETBF reporting system database. In this process it was evident that a number of data processing, checking and quality control measures could be automated. These were implemented and feedback was given to AFMA database managers regarding common problems encountered in the data. During the course of the project we met with representatives from AFMA to create a long term plan for integration of our reporting system into their planned Data Warehouse and Reporting System. Problems regarding species codes and details of fishing activity were presented at a RAG meeting. The ensuing disussions resulted in answers to most problems and suggestions for dealing with recurring problems.

In general the reports are generated completely from information available in the AFMA database, but in one instance there is a need to calculate a location datum that deserves explanation: in order to assign a shot to a zone, a center point is calculated for each shot by determining the intersection of the vector between the first set and the first haul and the vector between the last set and the last haul (Figure 2).

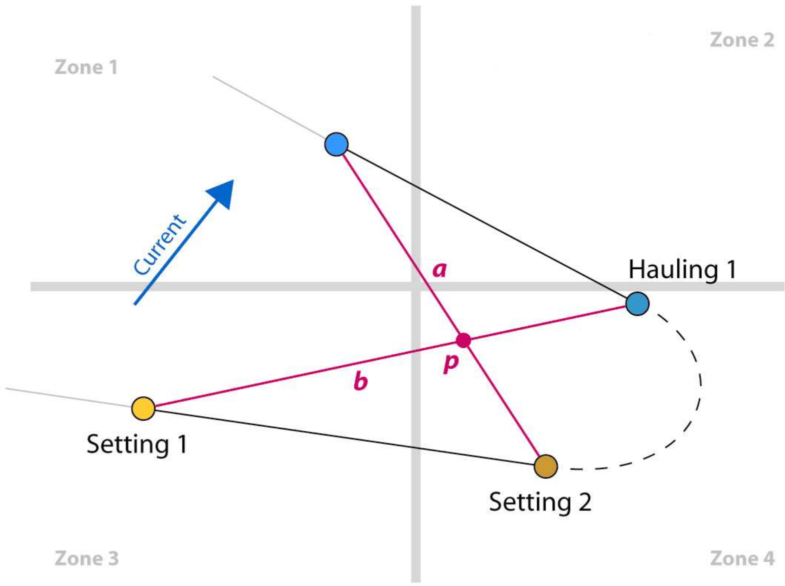


Figure 2 – Determination of the centrepoint of a longline fishing shot.

Results/Discussion

Reporting System

Two separated report modes were created for this project 1) an annual report with fixed spatial boundries and time periods, and 2) a specialised report with user-defined spatial boundaries and time periods (see Appendix 5 for an example of this report type used in the submitted paper [Hobday et al., *In review*]). For each mode, two separate reports are generated; one has catch estimates with 95% confidence intervals based on both the simple random sampling estimator of variance, and the other with catch estimate confidence intervals based on the cluster estimator of variance.

In the first report we adopted spatial boundaries previously defined by Bravington et al. (2002). Zone 1 is a region north of 22° S, and zone 5 is the region south of 34° S. Between these two latitudes are three intermediate zones, consisting of zone 2, which is the western inshore region extending out to 155° E, zone 3, a middle region from 155° E to 158° E (approximately the Lord Howe Rise) and zone 4 the eastern region east of this to about 170° E (east of Norfolk Island) (Figure 3). Each zone is divided into four seasonal strata: July - September, October – December, January – March and April – June, thus giving 20 reporting strata in each annual report. For the specialised report, users define discrete spatial boundaries and time periods. In both reports, however, the general structure of the resulting output is identical.

Below, the annual 2005/2006 report is presented as an example of a typical annual report. For each report there are two figures and six tables; for either the annual or specialized report, the report output is identical. In general the report shows the distribution and intensity of effort in the ETBF, as well as the corresponding coverage of AFMA observers (Figure 3, Table 1). Additionally, the catch locations of sensitive species are mapped (Figure 4). For the major commercial species there is a statistical comparison of retained catch reported in ETBF logbooks and that estimated from AFMA observer data for both commercial species (Tables 2 and 3) and all species in the fishery (Table 6). The remaining tables give the total estimated catch in the ETBF (Table 4), and the life status of discarded catch (Table 5).

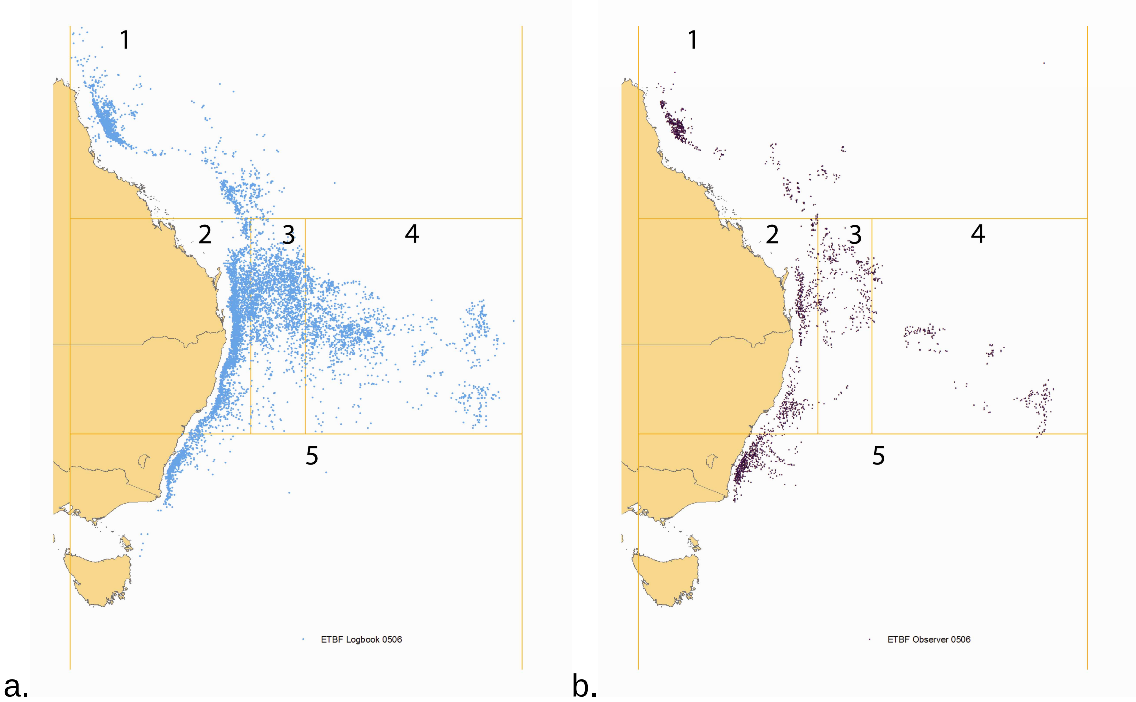


Figure 3. Location of longline sets in Australia's Eastern Tuna and Billfish Fishery, July 2005–June 2006, recorded in (a) fisherman logbooks, and (b) by AFMA observers; boxed areas denote spatial zones 1–5.

Table 1. Observer coverage of fishing operations in the ETBF, July 2005-June 2006.

		Number of sets		Number of hooks		Coverage	
Year	Zone	Logbook	Observed	Logbook	Observed	Sets	Hooks
July-Sept. 2005	1	324	21	165,179	6,535	6.48%	3.96%
	2	1,109	74	1,145,003	78,579	6.67%	6.86%
	3	388	13	400,510	12,874	3.35%	3.21%
	4	457	15	500,889	17,481	3.28%	3.49%
	5	103	63	104,910	62,090	61.17%	59.18%
Oct.-Dec. 2005	1	333	52	205,113	31,178	15.62%	15.20%
	2	444	26	450,744	24,637	5.86%	5.47%
	3	556	27	603,038	27,472	4.86%	4.56%
	4	477	19	554,432	25,446	3.98%	4.59%
	5	215	41	204,790	43,812	19.07%	21.39%
Jan.-Mar. 2006	1	357	22	264,484	14,323	6.16%	5.42%
	2	465	17	462,827	16,471	3.66%	3.56%
	3	397	8	413,698	8,213	2.02%	1.99%
	4	399	26	450,219	31,404	6.52%	6.98%
	5	214	16	194,910	12,068	7.48%	6.19%
Apr.-June 2006	1	495	46	520,966	42,010	9.29%	8.06%
	2	714	26	789,726	31,160	3.64%	3.95%
	3	599	38	759,718	46,102	6.34%	6.07%
	4	201	8	226,174	8,843	3.98%	3.91%
	5	143	21	135,316	15,896	14.69%	11.75%
Total		8,390	579	8,552,646	556,594	6.9%	6.5%

Table 2. Estimated number of fish in retained catch as percent of catch reported in logbooks for ETBF, July 2005–June 2006. ALB: albacore tuna, BET: bigeye tuna, DOL: dolphin fish, MLS: striped marlin, SWO: swordfish, YFT: yellowfin tuna.

Quarter	Zone	ALB	BET	DOL	MLS	SWO	YFT
July-Sept. 2005	1	9%	207%	22%	*	55%	258%
	2	66%	57%	74%	79%	109%	120%
	3	81%	114%	361%	61%	72%	147%
	4	169%	96%	264%	151%	96%	159%
	5	123%	162%	101%	51%	131%	72%
Oct.-Dec. 2005	1	21%	56%	85%	29%	54%	97%
	2	173%	141%	29%	87%	55%	68%
	3	22%	114%	166%	163%	134%	94%
	4	128%	54%	80%	101%	68%	146%
	5	103%	117%	95%	*	151%	62%
Jan.-Mar. 2006	1	23%	346%	162%	*	151%	139%
	2	162%	93%	169%	67%	87%	40%
	3	58%	29%	30%	138%	82%	84%
	4	42%	21%	117%	363%	154%	46%
	5	9%	160%	243%	98%	31%	135%
Apr.-June 2006	1	122%	53%	105%	33%	70%	75%
	2	149%	65%	80%	123%	21%	165%
	3	98%	102%	152%	173%	123%	102%
	4	130%	50%	53%	256%	34%	192%
	5	119%	25%	26%	52%	80%	363%
Total		94%	84%	113%	124%	96%	125%

*: None caught and retained in this strata.

Table 3. Comparison of estimated number of fish in retained catch, and catch reported in logbooks for ETBF, July 2005–June 2006. ALB: albacore tuna, BET: bigeye tuna, DOL: dolphin fish, MLS: striped marlin, SWO: swordfish, YFT: yellowfin tuna.

	ALB	BET	DOL	MLS	SWO	YFT
Observer estimate	92,354	13,873	17,187	6,596	24,104	64,747
Logbook total	98,111	16,514	15,245	5,331	25,020	51,939
Observer estimate as percent of logbook	94%	84%	113%	124%	96%	125%

Table 4. Total estimated catch (retained and discarded) in ETBF, July 2005–June 2006.

Group	Common name	Catch percent of total for all species	Estimated catch	95% CI percent of estimate
Tuna	Albacore Tuna	27%	95,655	12%
	Yellowfin tuna	20%	71,771	14%
	Bigeye Tuna	6%	20,344	12%
	Skipjack Tuna	2%	7,259	26%
	Southern Bluefin Tuna	0.5%	1,740	47%
	Northern Bluefin Tuna	0.04%	131	70%
Billfish	Broad Billed Swordfish	8%	26,619	8%
	Striped Marlin	2%	7,324	19%
	Black Marlin	1%	2,227	21%
	Blue Marlin	0.2%	782	29%
	Indo-Pacific Sailfish	0.04%	145	75%
	Marlins, Sailfishes nei	0.02%	55	87%
Other fishes	Longnose Lancetfish	8%	29,216	13%
	Dolphinfish	5%	18,199	20%
	Black Oilfish	5%	17,231	18%
	Snake Mackerel	3%	9,603	20%
	Ray's Bream	3%	9,484	53%
	Mixed fish	1%	3,802	18%
	Sunfish	1%	3,537	27%
	Shortnose Lancetfish	1%	3,086	25%
	Wahoo	1%	1,990	15%
	Shortbilled Spearfish	1%	1,963	25%
	Opah	0.2%	742	39%
	Oilfish	0.2%	660	85%
	Great Barracuda	0.1%	475	32%
	Toadfishes	0.1%	338	47%
	Rudderfish	0.1%	208	58%
	Pickhandle Barracuda	0.04%	150	73%
	Barracouta	0.03%	104	64%
	Malabar Grouper	0.02%	86	77%
	Big-scale Pomfret	0.02%	66	103%
	Hector's Lanternfish	0.02%	58	194%
	Sand Whiting	0.01%	50	70%
	Dealfish species	0.01%	45	132%
	Black Bream	0.01%	42	106%
	Frostfishes	0.01%	29	135%
	Blue-toothed Tuskfish	0.01%	25	162%
	Yellowtail Kingfish	0.01%	19	157%
	Rough Pomfret	0.005%	16	180%
	Blue Warehou	0.004%	14	199%
	Jackass Morwong	0.004%	14	199%
	Black Conger Eel	0.001%	5	196%
	Driftfishes	0.001%	3	117%
Sharks and Rays	Blue Shark	2%	6,384	15%
	Shortfin Mako	1%	2,384	14%
	Pelagic Stingray	0.5%	1,698	24%
	Bronze Whaler	0.4%	1,499	29%
	Oceanic Whitetip Shark	0.2%	847	31%
	Tiger Shark	0.2%	773	34%
	Bigeye Thresher	0.2%	605	42%

Group	Common name	Catch percent of total for all species	Estimated catch	95% CI percent of estimate
Sharks and Rays	Silky Shark	0.2%	595	42%
	Manta Ray	0.1%	314	47%
	Hammerhead Shark	0.1%	299	56%
	Pelagic Thresher	0.1%	243	63%
	Dusky Shark	0.1%	235	68%
	Thresher Shark	0.1%	225	67%
	Basking Shark	0.1%	221	135%
	Common Blacktip Shark	0.05%	162	145%
	Whaler Shark	0.04%	133	57%
	Crocodile Shark	0.04%	133	93%
	Scalloped Hammerhead	0.03%	99	157%
	Porbeagle	0.02%	60	100%
	Great Hammerhead	0.01%	25	205%
	Smooth Hammerhead	0.01%	25	208%
	Cookie-cutter Shark	0.01%	24	164%
	Longfin Mako	0.004%	15	101%
	Spurdog	0.004%	14	199%
	Australian blacktip shark	0.004%	12	162%
Sea birds	Flesh Footed Shearwater	0.01%	34	116%
	Wandering Albatross	0.01%	27	127%
	Albatrosses	0.004%	15	193%
	Black Browed Albatross	0.002%	9	146%
	Great Winged Petrel	0.001%	5	194%
	Wedge Tailed Shearwater	0.001%	5	196%
	Grey Headed Albatross	0.0005%	2	118%
Turtles	Leatherback turtle	0.03%	97	83%
	Pacific (Olive) Ridely turtle	0.02%	75	102%
	Loggerhead turtle	0.01%	48	124%
	Green turtle	0.01%	18	173%
	Turtles	0.005%	16	180%
Marine mammals	Australian Fur Seal	0.002%	9	146%
	Humpback whale	0.01%	31	177%
Total			352,724	

Table 5. Observed life status of discarded catch of fish and total catch of sea birds, turtles and marine mammals in the ETBF, July 2005–June 2006.

Group	Common name	Number observed	Dead	Just-alive	Sluggish-alive	Vigorous-alive
Tuna	Yellowfin tuna	490	44%	1%	7%	47%
	Bigeye Tuna	333	25%	3%	11%	61%
	Albacore Tuna	205	92%	2%	0%	6%
	Skipjack Tuna	97	96%	0%	0%	4%
	Southern Bluefin Tuna	92	5%	5%	11%	78%
Billfish	Black Marlin	218	39%	2%	7%	51%
	Broad Billed Swordfish	126	48%	6%	10%	37%
	Blue Marlin	47	32%	4%	15%	49%
	Striped Marlin	41	34%	0%	10%	56%
	Marlins, Sailfishes nei	5	20%	0%	20%	60%
	Indo-Pacific Sailfish	3	33%	0%	0%	67%
Sharks and Rays	Blue Shark	406	6%	4%	13%	77%
	Pelagic Stingray	90	2%	14%	10%	73%
	Bronze Whaler	81	6%	2%	14%	78%
	Shortfin Mako	52	10%	0%	13%	77%
	Tiger Shark	49	2%	0%	12%	86%
	Bigeye Thresher	36	56%	6%	3%	36%
	Oceanic Whitetip Shark	34	3%	6%	6%	85%
	Silky Shark	26	12%	0%	15%	73%
	Manta Ray	20	0%	5%	15%	80%
	Hammerhead Shark	13	31%	8%	0%	62%
	Pelagic Thresher	13	46%	8%	23%	23%
	Dusky Shark	10	20%	0%	10%	70%
	Thresher Shark	10	60%	10%	10%	20%
	Whaler Shark	10	0%	0%	0%	100%
	Basking Shark	5	20%	0%	0%	80%
	Crocodile Shark	5	0%	0%	20%	80%
	Common Blacktip Shark	4	0%	0%	0%	100%
	Scalloped Hammerhead	4	75%	0%	0%	25%
	Longfin Mako	3	33%	0%	0%	67%
	Porbeagle	3	100%	0%	0%	0%
	Australian blacktip shark	1	0%	0%	0%	100%
	Spurdog	1	0%	100%	0%	0%
Other fishes	Longnose Lancetfish	1,638	86%	5%	3%	7%
	Snake Mackerel	624	89%	2%	4%	5%
	Sunfish	242	1%	2%	13%	85%
	Shortnose Lancetfish	240	59%	5%	3%	34%
	Mixed fish	111	64%	5%	2%	29%
	Black Oilfish	108	38%	10%	25%	27%
	Dolphinfish	47	21%	4%	4%	70%
	Shortbilled Spearfish	47	72%	13%	0%	15%
	Great Barracuda	46	35%	9%	15%	41%
	Wahoo	16	81%	0%	0%	19%
	Toadfishes	14	0%	14%	43%	43%
	Malabar Grouper	13	15%	0%	0%	85%
	Oilfish	13	31%	8%	8%	54%
	Pickhandle Barracuda	12	33%	33%	17%	17%
	Barracouta	8	38%	13%	25%	25%
	Ray's Bream	7	29%	0%	0%	71%
	Rudderfish	6	83%	0%	17%	0%
	Hector's Lanternfish	4	0%	25%	0%	75%

Group	Common name	Number observed	Dead	Just-alive	Sluggish-alive	Vigorous-alive
Other fishes	Big-scale Pomfret	3	33%	33%	33%	0%
	Opah	3	33%	0%	67%	0%
	Dealfish species	2	100%	0%	0%	0%
	Frostfishes	2	0%	0%	50%	50%
	Sand Whiting	1	100%	0%	0%	0%
Sea birds	Flesh Footed Shearwater	2	100%	0%	0%	0%
	Wandering Albatross	2	50%	0%	0%	50%
	Albatrosses	1	100%	0%	0%	0%
	Black Browed Albatross	1	0%	0%	0%	100%
	Great Winged Petrel	1	100%	0%	0%	0%
	Grey Headed Albatross	1	100%	0%	0%	0%
	Wedge Tailed Shearwater	1	100%	0%	0%	0%
Turtles	Leatherback turtle	6	0%	0%	0%	100%
	Pacific (Olive) Ridely turtle	3	0%	33%	33%	33%
	Loggerhead turtle	2	0%	0%	0%	100%
	Green turtle	1	0%	0%	0%	100%
	Turtles	1	0%	0%	0%	100%
Marine mammals	Australian Fur Seal	1	0%	0%	0%	100%
	Humpback whale	1	0%	0%	0%	100%

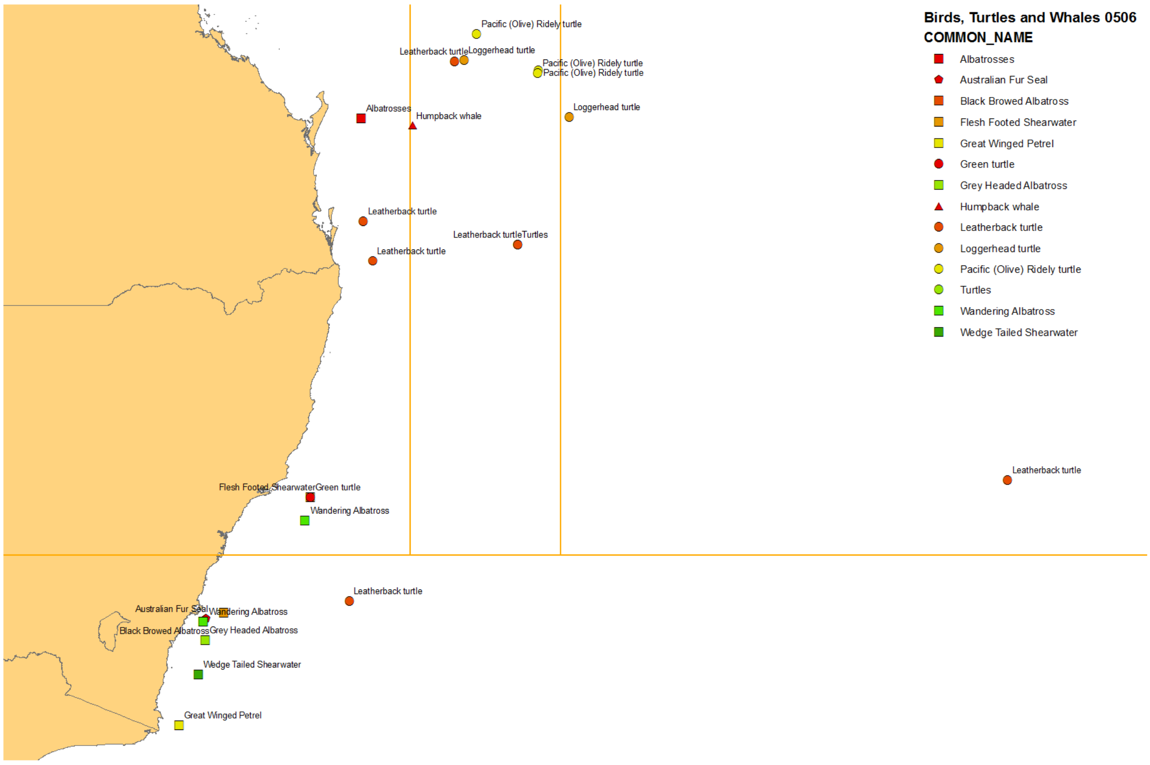


Figure 4. Location of seabirds, turtles, and marine mammals observed in longline catch of the ETBF, July 2005–June 2006.

Table 6. Comparison of catch (retained and discarded) estimated from observer data and catch reported in logbooks for the ETBF, July 2005–June 2006. Comparisons that are significantly different (α = 0.05) are in bolded font.

Group	Common name	Estimate catch	+/- 95% confidence interval as percent of estimated catch	Logbook reported catch	Percent estimate of logbook
Tuna	Albacore Tuna	95,655	12%	99,695	96%
	Yellowfin tuna	71,771	14%	55,395	**130%**
	Bigeye Tuna	20,344	12%	18,372	111%
	Skipjack Tuna	7,259	26%	2,619	**277%**
	Southern Bluefin Tuna	1,740	47%	375	**464%**
	Northern Bluefin Tuna	131	70%	77	170%
	Dogtooth Tuna			1	
Billfish	Broad Billed Swordfish	26,619	8%	26,077	102%
	Striped Marlin	7,324	19%	5,682	**129%**
	Black Marlin	2,227	21%	641	**347%**
	Blue Marlin	782	29%	732	107%
	Indo-Pacific Sailfish	145	75%	94	154%
	Marlins, Sailfishes nei	55	87%		
Sharks and Rays	Blue Shark	6,384	15%	3,347	**191%**
	Shortfin Mako	2,384	14%	2,144	111%
	Pelagic Stingray	1,698	24%		
	Bronze Whaler	1,499	29%	1,185	127%
	Oceanic Whitetip Shark	847	31%	813	104%
	Tiger Shark	773	34%	501	**154%**
	Bigeye Thresher	605	42%		
	Silky Shark	595	42%	100	**595%**
	Manta Ray	314	47%	26	**1208%**
	Hammerhead Shark	299	56%	293	102%
	Pelagic Thresher	243	63%		
	Dusky Shark	235	68%	155	152%
	Thresher Shark	225	67%	223	101%
	Basking Shark	221	135%		
	Common Blacktip Shark	162	145%		
	Whaler Shark	133	57%		
	Crocodile Shark	133	93%	57	233%
	Scalloped Hammerhead	99	157%		
	Porbeagle	60	100%	22	274%
	Great Hammerhead	25	205%		
	Smooth Hammerhead	25	208%		
	Cookie-cutter Shark	24	164%	2	1182%
	Longfin Mako	15	101%	2	744%
	Spurdog	14	199%		
	Australian blacktip shark	12	162%		
	Shark other			0	
	Grey Nurse			1	
	Stingray			29	
	Blacktip sharks			183	
	Skates and Rays			619	
Other fishes	Longnose Lancetfish	29,216	13%		
	Dolphinfish	18,199	20%	15,521	117%

Group	Common name	Estimate catch	+/- 95% confidence interval as percent of estimated catch	Logbook reported catch	Percent estimate of logbook
Other fishes	Black Oilfish	17,231	18%	8,432	**204%**
	Snake Mackerel	9,603	20%		
	Ray's Bream	9,484	53%	15,914	**60%**
	Mixed fish	3,802	18%	38	**10006%**
	Sunfish	3,537	27%	1	**353651%**
	Shortnose Lancetfish	3,086	25%		
	Wahoo	1,990	15%	1,880	106%
	Shortbilled Spearfish	1,963	25%	1,222	**161%**
	Opah	742	39%	929	80%
	Oilfish	660	85%	494	134%
	Great Barracuda	475	32%		
	Toadfishes	338	47%		
	Rudderfish	208	58%	15,738	**1%**
	Pickhandle Barracuda	150	73%		
	Barracouta	104	64%	332	**31%**
	Malabar Grouper	86	77%		
	Big-scale Pomfret	66	103%		
	Hector's Lanternfish	58	194%		
	Sand Whiting	50	70%		
	Dealfish species	45	132%		
	Black Bream	42	106%		
	Frostfishes	29	135%		
	Blue-toothed Tuskfish	25	162%		
	Yellowtail Kingfish	19	157%	6	317%
	Rough Pomfret	16	180%		
	Blue Warehou	14	199%		
	Jackass Morwong	14	199%		
	Black Conger Eel	5	196%		
	Driftfishes	3	117%		
	Butterfly Mackerel			1	
	Luvaru			1	
	Oarfish			1	
	Pomfret			1	
	Spanish Mackerel			1	
	Trevallies and jacks			1	
	Dealfish			2	
	Diodontidae – undifferentiated			2	
	Green jobfish			3	
	Coral trout			4	
	Southern Frostfish			4	
	Cardinal Fish			5	
	Reef ocean perch			6	
	Tetraodontidae – undifferentiated			6	
	Black Kingfish			17	
	Sphyrna zygaena			157	
	Ocean Sunfish			656	
	Mackerel			4,313	
	Lancet fish			37,017	
Sea birds	Flesh Footed Shearwater	34	116%		
	Wandering Albatross	27	127%		
	Albatrosses	15	193%		

Group	Common name	Estimate catch	+/- 95% confidence interval as percent of estimated catch	Logbook reported catch	Percent estimate of logbook
Sea birds	Black Browed Albatross	9	146%		
	Great Winged Petrel	5	194%		
	Wedge Tailed Shearwater	5	196%		
	Grey Headed Albatross	2	118%		
Turtles	Leatherback turtle	97	83%		
	Pacific (Olive) Ridely turtle	75	102%		
	Loggerhead turtle	48	124%		
	Green turtle	18	173%		
	Turtles	16	180%		
Marine mammals	Humpback whale	31	177%		
	Australian Fur Seal	9	146%		

Bycatch modelling

Giannini and Lawrence (Appendix 4) found results of cluster analysis and PCA were both inconclusive. No significant evidence for groupings among the species data was revealed through either cluster analysis or PCA. In the modelling work, for six target species considered in the analysis, only one, swordfish, was found to have a significant relationship between its catch and that of a bycatch species. Thus it was concluded that in the ETBF there were no general or useful relationships between the catch rate of target and bycatch species.

The analysis of bycatch and oceanographic variables by Giannini and Lawrence was more hopeful. For a select number of bycatch species with high discard rates (i.e., sunfish, black marlin, lancetfish, and snake mackerel) there were a number of oceanographic variables that were significant predictors of catch rates (i.e., sea surface temperature, Australian altimetry, and SEP07_Cluster). The variable SEP07_Cluster, is a division of the ETBF into seasonal habitat clusters that was developed by Hobday et al. (*In review*, see Appendix 5). These habitat clusters, when used as strata in our reporting system, were found to increase the precision of ETBF catch estimates (see Hobday et al., [*In review*], see Appendix 5).

Data processing

The following is a list of data quality isues which were encountered and the solution which was employed.

- The Observer and Logbook data sets use a variety of different taxon codes such as fao_species_code, spc_id and spc_code. A translation table was created which maps codes between systems. Nonetheless, there were still uncertainties which had an effect on the data quality. Some of these problems were discussed and and solved at an ETBF RAG meeting. Consultation with fish taxonomists resolved the remaining differences. New data for which there is no matching entry in the translation table is manually checked and a new entry created.

- Tables with missing keys and indices - Indices are created. Records without keys are excluded from further processing.

- Missing values and values of unexpected field type – Data records in which required values are missing or of another type than expected are ignored.

- Orphaned records exist in some tables due to missing referential integrity at the time of data entry. These records are excluded from further processing.

- Longline sets outside the designated zones 1-5 or user-defined zone in the specialised report – A VBA program checks the position of each set, marks those which were not within zone coordinates as outliers and excludes them from further processing.

- Longline sets on land - All sets are plotted in ArcGIS (Figure 5). Sets on land, visually identified and recorded in an outlier table, excluding them from further processing.

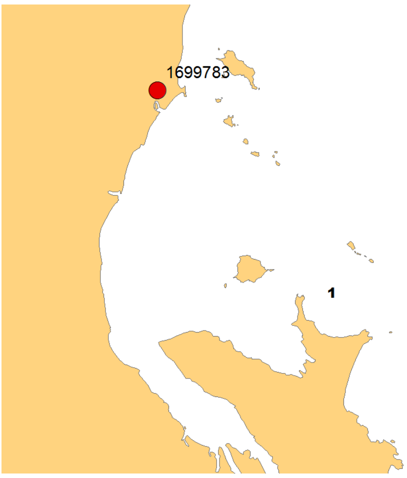


Figure 5 – Example of a longline set location on land

- In longline sets (Figure 6): insufficient set or haul records, set of haul distance greater than 500 km or an unacceptable number of hooks per set (0 or > 3000) – records are written to an erroneous records table and excluded from processing. A warning is issued if the number of bad records is excessive.

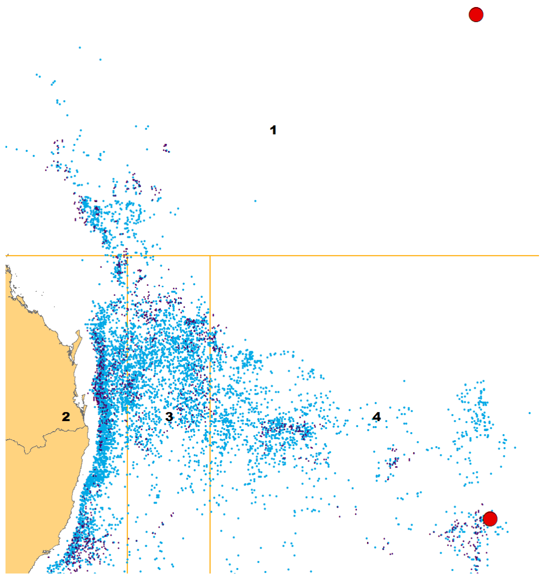


Figure 6 – Impossibly large distance between haul positions.

- Tables of excessive size for timely analysis – Depending on the amount of fields required from the table, either additional indices on seach fields are created (ie AFMOBS_BIO_COLLECTION) to reduce search time for each record or summary tables are created which include only relevant information (i.e., logbook pages table).

In addition to the above problems, a detailed list of problems and concerns regarding AFMA logbooks for the ETBF can be found in Appendix 6.

Benefits and adoption

This work proposed to benefit the ETBF fishery, and reports generated by this work are being used in great detail in annual meetings of the ETBF RAG. The reports allow for critical feedback to be given to the AFMA observer program to better allocate observer coverage and improve data quality. Catch and bycatch estimates provide a basis for discussion of ETBF management by the RAG members.

The reports are being used on an international level in management meetings such as the annual National Fishery Report for the Scientific Committtee of the WCPFC.

Data and evaluation techniques from the reports provided a contribution to the testing of Hobday et al.'s (*in review*) definition of dynamic pelagic habitats in the waters surrounding the ETBF.

Further Development

Further development of this work awaits completion of and integration with a proposed AFMA Data Warehouse and Reporting System (Figure 7). In the first step an Observer transaction database is to be constructed. This is to be followed by a transaction warehouse. Observer data will be entered in the field into electronic notebooks, and raw data files (dfl in Figure 7) will be uploaded into the Data Warehouse. Finally, old data files (i.e., Tuna I & Tuna II) will be entered. Future additions to the Data Warehouse are most likely CDR, VMS and licensing information.

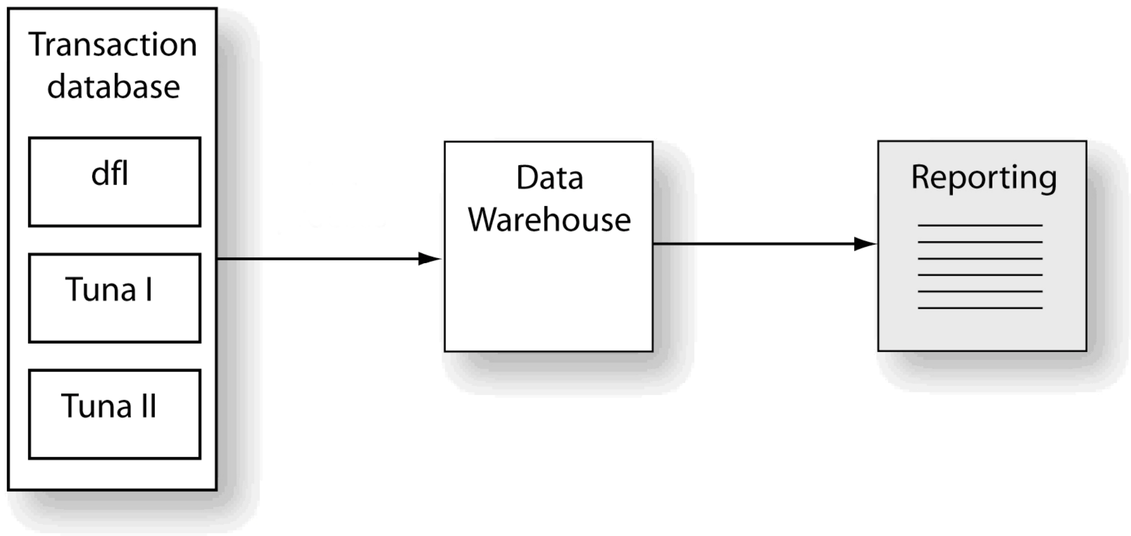


Figure 7 – AFMA Data Warehouse and Reporting System.

The ETBF report will be integrated with the AFMA system through a web-based user interface to specify the spatial and temporal boundaries for the specialized report. It will also be possible to use the statistical characteristics of the catch estimates from user-defined regions to understand how allocation of observer effort between regions affects the precision of catch estimates, thus providing the basis for improved observer effort allocation across the ETBF. The code required to generate data tables for this analysis has been completed but integration into the web-based interface is pending the completion of the AFMA Data Warehouse and Reporting System.

The Data Warehouse and Reporting System has undergone a substantial change in structure since the completion of our reporting system. Presently, the new AFMA database structure is incompatible with our reporting system. Future integration will require either rewriting of code and extensive object renaming, a task requiring considerable time and resources, or for AFMA to revert back to the original database structure.

Planned outcomes

This project was intended to facilitate the provision of information to support the monitoring and management of Australian fisheries. Specifically, this project had the goal of achieving:

1. Cost-effective management of Australian Commonwealth fisheries through optimization of outputs andrelevance of fishery observer activity and deployment and logbook programs.

   This planned outcome has been partially achieved through the development and adoption of standardized summaries and analysis of oberserver and logbook data for the ETBF. An analytical framework was developed that allowed for the rapid provision of catch estimates for the ETBF that has been used on an annual basis for RAG meetings and National Fishery Report for the Scientific Committtee of the WCPFC.

   We created a centralised database and flexible software interface for integrated queries and analyses of observer data, thus permitting regular and ad hoc summaries of observer data to AFMA, DEWHA, Research Advisory Groups and the Ecological Risk Assessment Team. Queries are designed to permit summaries based on selected areas and times, thus allowing for specific cross referencing to existing dependent data sets.

   Our reporting system detects data errors in a semi-automated process which has enabled a more streamlined and rapid method of preparing raw AFMA data for analysis. Moreover, this project identified ways to improve the usefulness of observer and logbook databases methods. This information was conveyed to AFMA and may be implemented in their planned Data Warehouse and Reporting System.

2. Improved understanding of impacts of case study fisheries on bycatch populations.

   Our reporting system produces statistical estimates of the catch and for all species caught in the ETBF. Additionally, it gives the fate and life status of all species encountered in the fishery. This information is particularly relevant and fundamental to re-occurring discussions at national and international fishery management meetings.

3. Improved information base for mitigation of bycatch of threatened and endangered species.

The provision of statistical estimates of bycatch provides the basis for management discussion on the conservation of threatened and endangered species. Our analysis of associations between targeted and bycatch species, as well as associations with environmental variables can be used to inform and thus improve mitigation efforts to minimize bycatch of selected species.

4. Effective and adaptive deployment of effort in AFMA observer programs.

Our reporting system provides an analysis of observer coverage in the ETBF, and therefore creates the basis for improved allocation of observer effort in the fishery. The statistical analyses of catch and bycatch can be used to guide an adaptive approach to allocation of observer effort across the ETBF, this final step however, requires completion of, and integration with, the AFMA Data Warehouse and Reporting system.

References

AFMA. 2001. Australia's tuna fisheries bycatch action plan: background paper. Australian Fisheries Management Authority, Canberra.

Babcock, E.A., E.K. Pikitch, and C.G. Hudson. 2003. How much observer coverage is enough to adequately estimate bycatch? Oceana, Washington DC, http://www.oceana.org/uploads/BabcockPikitchGray2003FinalReport.pdf.

Bravington, M., C. Burridge and P. Toscas. 2002. Design of Observer Program to monitor bycatch species in the Eastern Tuna and Billfish Fishery. Report to AFMA, Canberra.

Cochran, W. G. 1977. Sampling techniques, 3rd edition. Wylie and Sons, NY.

Dambacher, J.M., T.A. Patterson, J.S. Gunn, and T.I. Carter. 2003. Southern bluefin tuna by-catch in the Eastern Tuna and Billfish Fishery. Confidential Final Report-Project No. R00/1126. CSIRO Marine Research. Hobart.

Hobday, A., Stobutzki, I. Webb, H. and ERA Team. 2004. Ecological Risk Assessment for Effects of Fishing: Eastern Tuna and Billfish Fishery. Report for the Australian Fisheries Management Authority, Canberra.

Hobday, A. J., A. Smith, H. Webb, R. Daley, S. Wayte, C. Bulman, J. Dowdney, A. Williams, M. Sporcic, J. Dambacher, M. Fuller, T. Walker. 2007. Ecological Risk Assessment for the Effects of Fishing: Methodology. Report R04/1072 for the Australian Fisheries Management Authority, Canberra.

Hobday, A. J., Young, J. W., Dambacher, J. M., Moeseneder, C. Defining dynamic pelagic habitats on the east coast of Australia. *In review*.

Sainsbury, K., Haward, M., Kriwoken, L., Tsamenyi, M. and Ward, T. 1997. Multiple use management in the Australian marine environment: principles, definitions and elements. Oceans Planning and Management Issues Paper 1. Department of the Environment, Canberra. 42 pp.

Appendices 1-6

Appendix 1

Intellectual Property

There is no intellectual property and/or valuable information arising from this research project.

Appendix 2

Staff Engaged on Project

Jeffrey M. Dambacher[1], Christian H. Moeseneder[2], Scott P. Cooper[3], Robert A. Campbell[3], Fiona M. Giannini[4], Emma K. Lawrence[2*, 4], Alistair J. Hobday[3], Jock W. Young[3].

[1]CSIRO Mathematics, Informatics and Statistics, GPO Box 1538, Hobart TAS 7001, Australia.
[2]CSIRO Marine and Atmospheric Research, 233 Middle Street, Cleveland QLD 4163, Australia.
[3]CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart TAS 7001, Australia.
[4]Fisheries and Marine Sciences, Bureau of Rural Sciences, GPO Box 858, Barton ACT 2601, Australia.
* Present address.

Appendix 3

Data work document v. 2.0

Christian H. Moeseneder

Data sources

All **LOGBOOK** data is in area **TUNA** in the pealgic DB in Oracle (hostname: aqua.hba.marine.csiro.au, SID: aqua, username: MOE009, password: tuna. In SQLDeveloper click on Pelagic DB, Other users, TUNA)

All **OBSERVER** data is in area **AFMAOBS** in the pealgic DB in Oracle (hostname: aqua.hba.marine.csiro.au, SID: aqua, username: MOE009, password: tuna. In SQLDeveloper click on Pelagic DB, Other users, AFMAOBS)

General notes

The name of tables created by query outputs do not change from one year to the next so that other queries which use these tables will still work. To interprete the data in these tables the currently active season in table REPORT_PARAMETERS can be used. Query output tables which are used in analyses and publications are copied and renamed with a season appendix SSSS (ie 0506)

Output tables should never be renamed since this may change queries. Output tables (such as those ending in "…SSSS") should be copied to new tables (with names such as "…0506" to indicate the season).

Usage of species codes

- Taxa in the observer data are looked up in the ETBF_SPECIES_GROUPS table using FAO_SPECIES_CODE
- Taxa in the logbook data are looked up in the ETBF_SPECIES_GROUPS table using SPC_ID

Concept

These procedures import data from the original Oracle tables and prepare the data for use in all subsequent queries.

When

Perform this procedure upon notification from Scott Cooper of the Data Centre that new AFMA logbook and observer data has been received. Usually this happens twice a year.

Run order

- Open table REPORT_PARAMETERS and check/set field ACTIVITY for the year for which you want to run the queries
- Logbook positions (not year-specific)
- Observer positions (not year-specific)
- Voyages (not year-specific)
- Gear items (not year-specific)
- Bio collection (not year-specific)
- Catch aggregate (not year-specific)
- Combine species (not year-specific)
- MSS codes (not year-specific)
- Create logbook trips (not year-specific)
- Create logbook trips per strata (year-specific)
- Logbook effort per year, quarter and zone (year-specific)
- Logbook effort per species, year, quarter and zone (year-specific)
- Logbook vessels (year-specific)
- Create observer sets (year-specific)
- Create observer sets per year, quarter and zone (year-specific)
- Prepare species groups

- Figure 1a
- Figure 1b
- Table 1
- Trips and hooks per strata for cluster estimate
- Table 2
- Table 2 totals
- Create observer null catch records
- Create observer null catch records for cluster estimate
- Observer species catch per shot with variance
- Observer species catch per voyage with variance
- Create species 95 percentile
- Create species 95 percentile for cluster estimate
- Table 4 using Ratio Estimate
- Table 4 using Cluster Estimate
- Table 6 using Ratio Estimate
- Table 6 using Cluster Estimate
- Table 3
- Table 5
- Figure 3
- Additional Info for Jeff

Logbook positions

Prerequisites
None

Tables used
- TUNA_AFZ_FOP_AUS
- TUNA_AFZ_FOP_AUS_OUTLIERS
- TUNA_AFZ_FOP_AUS_REMOTE
- ETBF_LOG_POS

Queries used
- ETBF_LOG_POS_1
- ETBF_LOG_POS_2
- TUNA_AFZ_FOP_AUS_CREATE

Concept
TUNA_AFZ_FOP_AUS_REMOTE is a link to the original Oracle table. (Connect with service name: AQUA, username: moe009, password: tuna). Query TUNA_AFZ_FOP_AUS_CREATE creates a local copy of the table named TUNA_AFZ_FOP_AUS (for performance reasons) which includes all years but only records of form_type “AL05”, ACTIVITY_CODE = 274, FISHING_METHOD_ID = 236 and geographic coordinates within the ETBF. Table TUNA_AFZ_FOP_AUS_OUTLIERS contains the records which have been identified as not being within the ETBF area. Query ETBF_LOG_POS_1 selects all records from TUNA_AFZ_FOP_AUS which are not in TUNA_AFZ_FOP_AUS_OUTLIERS. Query ETBF_LOG_POS_2 selects fields FOP_ID , the date and the longitude and latitude from query ETBF_LOG_POS_1.

Procedure
1. Ensure new data has been loaded into Oracle table TUNA_AFZ_FOP_AUS by Scott Cooper of the Data Centre
2. Run query TUNA_AFZ_FOP_AUS_CREATE to create a new local table
3. Open table TUNA_AFZ_FOP_AUS and
   3.1. create a primary key on field FOP_ID to increase performance
   3.2. create an index named “book” on fields FORM_BOOK_NO and FORM_SEQ_NO
4. Run query ETBF_LOG_POS_2 to create table ETBF_LOG_POS
5. Plot lats and longs from table ETBF_LOG_POS in ArcMap to check for outliers
6. Identify outliers and enter the FOP_ID of the outliers in table TUNA_AFZ_FOP_AUS_OUTLIERS
7. Run query ETBF_LOG_POS_2 again
8. This produces the final version of table ETBF_LOG_POS

Final output

- ETBF_LOG_POS

Observer positions

Prerequisites
None

Tables used

- AFMAOBS_ACTIVITY_LOG
- AFMAOBS_ACTIVITY_LOG_OUTLIERS
- AFMAOBS_ACTIVITY_LOG_REMOTE
- ETBF_OBS_POS

Queries used

- ETBF_OBS_POS_1
- ETBF_OBS_POS_2
- AFMAOBS_ACTIVITY_LOG_CREATE

Concept
AFMAOBS_ACTIVITY_LOG_REMOTE is a link to the original Oracle table. (Connect with service name: AQUA, username: moe009, password: tuna). Query AFMAOBS_ACTIVITY_LOG_CREATE creates a local copy of the table named AFMAOBS_ACTIVITY_LOG (for performance reasons) which includes only records with geographic coordinates within the ETBF. Table AFMAOBS_ACTIVITY_LOG_OUTLIERS contains the records which have been identified as not being within the ETBF area. Query ETBF_OBS_POS_1 selects all records from AFMAOBS_ACTIVITY_LOG which are not in AFMAOBS_ACTIVITY_LOG_OUTLIERS. Query ETBF_OBS_POS_2 selects the key fields and lats and longs from ETBF_OBS_POS and creates a new table

Procedure

1. Ensure new data has been loaded into Oracle table AFMAOBS_ACTIVITY_LOG by Scott Cooper of the Data Centre
2. Run query AFMAOBS_ACTIVITY_LOG_CREATE to create a new local table
3. Run query ETBF_OBS_POS_2 to create table ETBF_OBS_POS
4. Plot lats and longs from table ETBF_OBS_POS in ArcMap to check for outliers
5. Identify outliers and enter the VOYAGE_ID, DAY_DATE_TIME, VESSEL_ACTIVITY_CODE and SHOT_NO of the outliers in table AFMAOBS_ACTIVITY_LOG_OUTLIERS
6. Run query ETBF_OBS _POS_2 again
7. This produces final version of table ETBF_OBS_POS

Final output

- ETBF_OBS_POS

Voyages

Prerequisites
None

Tables used

- AFMAOBS_VOYAGE_REMOTE
- AFMAOBS_VOYAGE

Queries used

- AFMAOBS_VOYAGE_ CREATE

Concept
This procedure creates a local copy of the AFMAOBS_VOYAGES table. The table is used to determine the number of unique boats in Table 1.

Procedure

1. Run query AFMAOBS_VOYAGE_ CREATE. This creates table AFMAOBS_VOYAGE.

Final output
- AFMAOBS_VOYAGE

Gear items

Prerequisites
None

Tables used
- AFMAOBS_VOYAGE_GEAR_ITEMS_REMOTE
- AFMAOBS_VOYAGE_GEAR_ITEMS

Queries used
- AFMAOBS_VOYAGE_GEAR_ITEMS_CREATE

Concept
Table AFMAOBS_VOYAGE_GEAR_ITEMS_REMOTE is the linked Oracle table. Query AFMAOBS_VOYAGE_GEAR_ITEMS_CREATE creates a local copy by copying all records from AFMAOBS_VOYAGE_GEAR_ITEMS_REMOTE into table AFMAOBS_VOYAGE_GEAR_ITEMS.

Procedure
1. Run query AFMAOBS_VOYAGE_GEAR_ITEMS_CREATE
2. Create an index on field VOYAGE_ID for performance reasons: Open table AFMAOBS_VOYAGE_GEAR_ITEMS in design mode, click on View/Indexes and add an index named “VOYAGE” on field “VOYAGE_ID” (Primary: No, Unique: No, Ignore Nulls: No).

Final output
- AFMAOBS_VOYAGE_GEAR_ITEMS

Bio collection

Prerequisites
None

Tables used
- AFMAOBS_BIO_COLLECTION
- AFMAOBS_BIO_COLLECTION_REMOTE

Queries used
- AFMAOBS_BIO_COLLECTION_CREATE

Concept
This procedure creates a local copy of linked table AFMAOBS_BIO_COLLECTION_REMOTE

Procedure
1. Run query AFMAOBS_BIO_COLLECTION_CREATE
2. Open table AFMAOBS_BIO_COLLECTION and create an index to improved access speed. Open table in design mode, click Index button, create index with name “VOYAGE_SHOT” on fields VOYAGE_ID and SHOT_NO.

Final output
- AFMAOBS_BIO_COLLECTION

Catch aggregate

Prerequisites
None

Tables used
- TUNA_AFZ_CATCH_AGGREGATE_AUS
- TUNA_AFZ_CATCH_AGGREGATE_AUS_REMOTE

Queries used
- TUNA_AFZ_CATCH_AGGREGATE_AUS_CREATE

Concept
This procedure copies the Oracle table TUNA_AFZ_CATCH_AGGREGATE_AUS to the local database.

Procedure
1. Run query TUNA_AFZ_CATCH_AGGREGATE_AUS_CREATE (duration possibly more than 10 minutes)
2. Create an index on field FOP_ID for performance reasons: Open table TUNA_AFZ_CATCH_AGGREGATE_AUS in design mode, click on View/Indexes and add an index named “FOP” on field “FOP_ID” (Primary: No, Unique: No, Ignore Nulls: No)

Final output
- TUNA_AFZ_CATCH_AGGREGATE_AUS

Combine species

Prerequisites
- Procedure “Catch aggregate”
- Procedure “Bio collection”

Tables used
- AFMAOBS_BIO_COLLECTION
- TUNA_AFZ_CATCH_AGGREGATE_AUS

Concept
This procedure combines species in observer and logbook data. It reads sequentially through both files and replaces certain FAO_SPECIES_CODE and SPC_ID values with others. The replacement parameters are in procedure ETBF_COMBINE_SPECIES and need to be manually edited if changes are required.

Procedure
1. Run procedure ETBF_COMBINE_SPECIES

Final output
Rewritten species codes in tables AFMAOBS_BIO_COLLECTION and TUNA_AFZ_CATCH_AGGREGATE_AUS

MSS codes

Prerequisites
None

Tables used
AFMAOBS_MSS_CODES_DETAIL
AFMAOBS_MSS_CODES_DETAIL_REMOTE

Queries used
AFMAOBS_MSS_CODES_DETAIL_CREATE

Concept
This procedure creates a local version of the AFMAOBS_MSS_CODES_DETAIL table

Procedure
1. Run query AFMAOBS_MSS_CODES_DETAIL_CREATE

Final output
AFMAOBS_MSS_CODES_DETAIL

Create logbook trips

Prerequisites
None

Tables used
TUNA_AFZ_TRIPS
TUNA_AFZ_TRIPS_REMOTE
TUNA_AFZ_TRIPS_COMPACTED

Queries used
TUNA_AFZ_TRIPS_CREATE
TUNA_AFZ_TRIPS_COMPACTED_CREATE

Concept
This procedure creates a local version of the TUNA_AFZ_TRIPS table. Only records with FISHERY_CODE = "ECT" are selected. The procedure also copies all records from TUNA_AFZ_TRIPS into TUNA_AFZ_TRIPS_COMPACTED. This table receives the records sorted by logbook serial and page numbers and creates a unique index. This dramatically improves the run time required for the following procedure "Create logbook trips per strata".

Procedure
1. Run query TUNA_AFZ_TRIPS_CREATE
2. Open table TUNA_AFZ_TRIPS and create
    a. a primary key on field TRP_ID
    b. an index named "start" on START_LOG_BOOK_SERIAL_NO and START_LOG_BOOK_PAGE_NO
    c. an index named "end" on END_LOG_BOOK_SERIAL_NO and END_LOG_BOOK_PAGE_NO
3. Open table TUNA_AFZ_TRIPS_COMPACTED and delete all records
4. Run query TUNA_AFZ_TRIPS_COMPACTED_CREATE

Final output
TUNA_AFZ_TRIPS
TUNA_AFZ_TRIPS_COMPACTED

Create logbook trips per strata

Prerequisites
Procedure "Create logbook trips"

Tables used
TUNA_AFZ_FOP_AUS
TUNA_AFZ_TRIPS
TUNA_AFZ_TRIPS_COMPACTED
ETBF_LOG_TRIPS

Queries used
ETBF_LOG_TRIPS_PER_STRATA_PRE
ETBF_LOG_TRIPS_PER_STRATA

Concept
This procedure reads all records in TUNA_AFZ_FOP_AUS for the current season set in REPORT_PARAMETERS and creates a new record in ETBF_LOG_TRIPS. Each new record is assigned a year, quarter and zone (for stratification). For each record a matching record in TUNA_AFZ_TRIPS_COMPACTED is searched for using the logbook page and serial number. Where a match is found, the new record in ETBF_LOG_TRIPS receives a trip number and a vessel id. Query ETBF_LOG_TRIPS_PER_STRATA creates a group by year, quarter and zone and counts the number of unique trips. This is used in the Cluster Estimate in Table 4.

Procedure
1. Ensure that the correct season is checked in table REPORT_PARAMETERS
2. Run procedure CREATE_ETBF_LOG_TRIPS. Note: This procedure may last longer than one hour.

Final output
Query ETBF_LOG_TRIPS_PER_STRATA

Logbook effort per year, quarter and zone

Prerequisites
Procedure “Logbook positions”

Tables used
- REPORT_PARAMETERS
- TUNA_AFZ_FOP_AUS
- ETBF_LOG_EFFORT_YEAR_QTR_ZONE

Concept
This procedure creates a table with logbook sets and hooks per year, quarter and zone.

Procedure
1. Ensure that the correct SEASON is checked in table REPORT_PARAMETERS
2. Ensure that procedure “Logbook positions” has run.
3. Run function CALC_ETBF_LOG_EFFORT_YEAR_QTR_ZONE. This fills table ETBF_LOG_EFFORT_YEAR_QTR_ZONE

Final output
- ETBF_LOG_EFFORT_YEAR_QTR_ZONE (in conjunction with ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE this table replaces old table TOTAL_CATCH_AND_EFFORT_2005_2006_YEAR_QTR_ZONE)

Logbook effort per species, year, quarter and zone

Prerequisites
Procedure “Logbook positions”

Tables used
- TUNA_AFZ_FOP_AUS
- ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE
- ETBF_SPECIES_GROUPS

Concept
This procedure creates a table with the logbook catch, weight, discard and catch records per species, year, quarter and zone. For each record a FAO_SPECIES_CODE is retrieved from ETBF_SPECIES_GROUPS.

Procedure
1. Ensure that the correct SEASON is checked in table REPORT_PARAMETERS
2. Ensure that procedure “Logbook positions” has run
3. Run function CALC_ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE

Final output
- ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE (in conjunction with table ETBF_LOG_EFFORT_YEAR_QTR_ZONE this table replaces old table TOTAL_CATCH_AND_EFFORT_2005_2006_YEAR_QTR_ZONE)

Logbook vessels

Prerequisites
Pocedure “Logbook Positions”

Tables used
- ETBF_LOG_VESSELS_YEAR_QTR_ZONE

Queries used

- ETBF_LOG_VESSELS_YEAR_QTR_ZONE_GROUPED
- ETBF_LOG_VESSELS_YEAR_QTR_ZONE_TOTAL

Concept
This procedure reads all records in table TUNA_AFZ_FOP_AUS in the same season as determined by table REPORT_PARAMETERS. It determines the quarter and zone of each record and writes the vessel id into table ETBF_LOG_VESSELS_YEAR_QTR_ZONE. Queries ETBF_LOG_VESSELS_YEAR_QTR_ZONE_GROUPED and ETBF_LOG_VESSELS_YEAR_QTR_ZONE_TOTAL create the unique number of vessels per strata. These numbers are then joined with other tables to create Table 1.

Procedure
1. Ensure that the correct SEASON is checked in table REPORT_PARAMETERS
2. Ensure that procedure “Logbook Positions” has run
3. Run function CALC_ETBF_LOG_VESSELS_YEAR_QTR_ZONE. This creates table ETBF_LOG_VESSELS_YEAR_QTR_ZONE

Final output
- Query ETBF_LOG_VESSELS_YEAR_QTR_ZONE_TOTAL

Create observer sets

Prerequisites
Pocedure “Observer Positions”

Tables used
- REPORT_PARAMETERS
- ETBF_OBS_POS
- ETBF_OBS_SETS
- ETBF_OBS_SETS_BAD

Concept
This procedure reads all shot records in ETBF_OBS_POS which are within the season specified in REPORT_PARAMETERS and creates one entry per cruise and shot in table ETBF_OBS_SETS. Caculated are average shot position, zone, number of sets and hauls as well as the number of hooks. Shots which have no set or haul records or shots where the set or haul distance is greater than 500 km are not written to ETBF_OBS_SETS but an entry is made for these records in ETBF_OBS_SETS_BAD. Number of hooks is determined by summarizing all hook values in table AFMAOBS_VOYAGE_GEAR_ITEMS of code TOTHKSST.

Procedure
1. Run function CALC_ETBF_OBS_SETS
2. Check number of entries in table ETBF_OBS_SETS_BAD to ensure that the number of bad records is not excessive.

Final output
- ETBF_OBS_SETS (replaces old tables ETBF_SETS_2005_2006 and ETBF_2005_2006_SETS_AND_HOOKS)
- ETBF_OBS_SETS_BAD

Create observer sets per year, quarter and zone

Prerequisites
- Pocedure “Observer positions”
- Procedure “Create sets”

Tables used
- ETBF_OBS_SETS
- ETBF_OBS_SETS_YEAR_QTR_ZONE

Queries used
- ETBF_OBS_SETS_YEAR_QTR_ZONE_CREATE

Concept
Creates a summarized table of ETBF_OBS_SETS per year, quarter and zone. This simplifies many of the queries which use this data.

Procedure
1. Run query ETBF_OBS_SETS_YEAR_QTR_ZONE_CREATE. This creates table ETBF_OBS_SETS_YEAR_QTR_ZONE

Final output
- ETBF_OBS_SETS_YEAR_QTR_ZONE

Prepare species groups

Prerequisites
Procedure “Logbook effort per species, year, quarter and zone”

Tables used
- ETBF_SPECIES_GROUPS
- ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE
- ETBF_SPECIES_GROUPS_MISSING

Functions used
- CHK_SPC_ID_IN_AGG

Concept
This procedure ensures that a translation entry is present in ETBF_SPECIES_GROUPS for all species in the Logbook. ETBF_SPECIES_GROUPS is a lookup table for translating AFMA species codes to SPC_ID and vice versa. Entries in this table were produced manually to begin with. Species which did not have an SPC_ID have been given an autogenerated SPC_ID starting with number 9000.

Procedure
1. Ensure that procedure “Logbook effort per species, year, quarter and zone” has completed.
2. Run function CHK_SPC_ID_IN_AGG. This function checks that there is an entry in ETBF_SPECIES_GROUPS for every species which is in ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE.
3. Open table ETBF_SPECIES_GROUPS_MISSING. If there are any entries create a new record for each species in table ETBF_SPECIES_GROUPS (use tables OBSERVER_SPECIES, TUNA_SPECIES and maybe AFMAOBS_SPECIES_CODE to find the correct species_code). Make sure that a new entry in this table has a value in the field COMBINE_TABLE6. Use a unique value in the field to calculate catch (and CI) for this species on its own or use the same number in multiple species to combine catches (and CIs) for Table 6.

Final output
- ETBF_SPECIES_GROUPS_MISSING

Figure 1a
Location of longline sets in Australia’s Eastern Tuna and Billfish Fishery recorded in fisherman logbooks

Content
Create map of logbook positions

Prerequisites
Procedure “Logbook positions”

Tables used
REPORT_PARAMETERS
ETBF_LOG_POS
ETBF_LOG_SSSS
ETBF_LOG_[SSSS]

Queries used
ETBF_LOG_PER_SEASON

Concept
This procedure extracts the logbook positional records from table ETBF_LOG_POS by selecting only those records which are within the season checked as active in table REPORT_PARAMETERS. A table with these records is created and the records are subsequently displayed in ArcGIS.

Procedure
1. Ensure that procedure “Logbook positions” has run.
2. Ensure the correct season is set active in table REPORT_PARAMETERS
3. Run query ETBF_LOG_PER_SEASON. The query creates table ETBF_LOG_SSSS.
4. **Copy** table ETBF_LOG_SSSS to ETBF_LOG_[SSSS] where SSSS stands for the season, ie 0506. Replace destination table if it already exists.
5. Create logbook map in ArcMap using W:\Mini_Projects\Eastern Tuna and Billfish Fishery Observer Data\ETBF.mxd:
    1. Import table ETBF_LOG_[SSSS] and copy display attributes from another “ETBF Logbook…” layer.
    2. Export this layer into the personal geodatabase “ETBF_GIS”.
    3. Delete the imported layer.
    4. Set all other data layers to inactive.
    5. Save map as ETBF Logbook Positions [SSSS].png

Final output
ETBF_LOG_[SSSS]
Graphic “ETBF Logbook Positions [SSSS].png”

Figure 1b
Location of longline sets in Australia’s Eastern Tuna and Billfish Fishery recorded by AFMA observers

Content
Create map of observer positions

Prerequisites
Procedure “Observer positions”

Tables used
REPORT_PARAMETERS
ETBF_OBS_POS
ETBF_OBS_SSSS
ETBF_OBS_[SSSS]

Queries used
ETBF_OBS_PER_SEASON

Concept
This procedure extracts the observer positional records from table ETBF_OBS_POS by selecting only those records which are within the season checked as active in table REPORT_PARAMETERS. A table with these records is created and the records are subsequently displayed in ArcGIS.

Procedure
1. Ensure that procedure “Observer positions” has run.
2. Ensure the correct season is set active in table REPORT_PARAMETERS
3. Run query ETBF_OBS_PER_SEASON. The query creates table ETBF_OBS_SSSS.
4. **Copy** table ETBF_OBS_SSSS to ETBF_OBS_[SSSS] where SSS stands for the season, ie 0506. Replace destination table if it already exists.
5. Create logbook map in ArcMap using W:\Mini_Projects\Eastern Tuna and Billfish Fishery Observer Data\ETBF.mxd:
    1. Import table ETBF_OBS_[SSSS] and copy display attributes from another “ETBF Observer Logbook…” layer.
    2. Export this layer into the personal geodatabase “ETBF_GIS”.
    3. Delete the imported layer.

4. Set all other data layers to inactive.
5. Save map as ETBF Observer Positions [SSSS].png

Final output
ETBF_OBS_SSSS
ETBF_OBS_[SSSS]
Graphic “ETBF Observer Positions [SSSS].png”

Table 1
Observer coverage of fishing operations in the ETBF

Prerequisites
Procedure “Observer positions”
Procedure “Voyages”
Procedure “Create sets”
Procedure “Gear Items”
Procedure “Logbook effort per year, quarter and zone”
Procedure “Logbook vessels”
Procedure “Catch aggregate”

Required tables
- ETBF_OBS_POS
- REPORT_PARAMETERS
- ETBF_OBS_SETS
- AFMAOBS_GEAR_ITEMS
- ETBF_LOG_EFFORT_YEAR_QTR_ZONE
- ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_[SSSS]
- ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_SSSS

Required queries
- ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE
- ETBF_OBS_VESSELS_PER_STRATA_GROUPED
- ETBF_LOG_VESSELS_YEAR_QTR_ZONE

Concept
This procedure combines sets, hook counts, unique vessel numbers and number of trips from the observer and logbook data into one table per year, quarter and zone. In addition the NC value (left side of equation) is calculated which is used in calculation 6.9 for table 4 and other tables.

Procedure
1. Ensure the correct season is set active in table REPORT_PARAMETERS
2. Ensure procedure “Observer Positions” has run and created table ETBF_OBS_POS
3. Ensure procedure “Voyages” has run and created table AFMAOBS_VOYAGE
4. Ensure procedure “Create Observer Sets” has run and created table ETBF_OBS_SETS
5. Ensure procedure “Gear Items” has run and created table AFMAOBS_GEAR_ITEMS
6. Ensure procedure “Logbook effort per year, quarter and zone” has run and created table ETBF_LOG_EFFORT_YEAR_QTR_ZONE
7. Ensure procedure “Logbook vessels” has run and created table ETBF_LOG_VESSELS_YEAR_QTR_ZONE
8. Ensure procedure “Catch Aggregate” has run and created table TUNA_AFZ_CATCH_AGGREGATE_AUS
9. Run query ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE to create table ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_SSSS.
10. **Copy** table ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_SSSS to ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_[SSSS]. Replace destination table if it exists.
11. Copy data to Excel spreadsheet as “Table 1”

Final output
- ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_SSSS (replaces old table ETBF_2005_2006_TABLE1_FINAL)
- ETBF_OBSLOG_SETS_HOOKS_YEAR_QTR_ZONE_[SSSS]

Trips and hooks per strata for Cluster Estimate

Concept
This procedure combines trips, hook counts and unique vessel numbers from the observer and logbook data into one table per year, quarter and zone. This is the same procedure as procedure “Table 1” but instead of sets, values are calculated for trips. In addition the NC value (left side of equation) is calculated which is used in calculation 6.9 for table 4 and other tables.

Procedure
1. Run query ETBF_OBSLOG_TRIPS_HOOKS_YEAR_QTR_ZONE
2. Copy table ETBF_OBSLOG_TRIPS_HOOKS_YEAR_QTR_ZONE_SSSS to ETBF_OBSLOG_TRIPS_HOOKS_YEAR_QTR_ZONE_[SSSS] where SSSS is the correct season

Final output
- Table ETBF_OBSLOG_TRIPS_HOOKS_YEAR_QTR_ZONE_SSSS
- Table ETBF_OBSLOG_TRIPS_HOOKS_YEAR_QTR_ZONE_[SSSS]

Table 2
Estimated number of fish in retained catch as percent of catch reported in logbooks for ETBF

Prerequisites
Procedure “Bio collection”
Procedure “Prepare species groups”
Procedure “Create observer sets”
Procedure “Create observer sets per year, quarter and zone”
Procedure “Logbook effort per year, quarter and zone”
Procedure “Logbook effort per species, year, quarter and zone”

Tables used
AFMAOBS_BIO_COLLECTION
ETBF_SPECIES_GROUPS
ETBF_OBS_SETS
ETBF_OBS_SETS_YEAR_QTR_ZONE
ETBF_LOG_EFFORT_YEAR_QTR_ZONE
ETBF_LOG_EFFORT_SPECIES_YEAR_QTR_ZONE

Queries used
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_FATE
ETBF_OBS_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM
ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM
ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_CROSS

Concept
This procedure created a crosstab matrix of the number of commercial fish in retained catch (fate code “R”) as percent of catch reported in logbooks. Rows are by year, quarter and zone. Columns are commercial species.

Procedure
1. Run query ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_CROSS
2. Copy/paste data from crosstab query into Excel spreadsheet (since result of a crosstab query in Access cannot create a table) as “Table 2”.

Final output
ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_CROSS

Table 2 totals

Prerequisites
Procedure “Table 2”

Queries used
ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_TOTAL

Concept
This procedure creates the totals for table 2.

Procedure
1. Run query ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_TOTAL
2. Append the totals to the bottom of table 2 making sure to match the total to the correct species.

Final output
ETBF_OBSLOG_CATCH_YEAR_QTR_ZONE_SPECIES_FATE_R_COM_TOTAL

Create observer null catch records

Prerequisites
- Procedure “Bio collection”
- Procedure “Create observer sets”

Concept
This procedure adds records with zero catch for each species in a voyage and shot which was caught in each stratum. Calculation 6.9 for table 4 ratio estimate depends on this.

Procedure
1. Run procedure CALC_ETBF_OBS_NULL_CATCH

Final output
Table ETBF_OBS_NULL_CATCH

Create observer null catch records for cluster estimate

Prerequisites
- Procedure “Bio collection”
- Procedure “Create observer sets”

Concept
This procedure adds records with zero catch for each species in a voyage which was caught in this strata. Calculation 6.9 for table 4 cluster estimate depends on this.

Procedure
1. Run procedure CALC_ETBF_OBS_NULL_CATCH_CLUSTER

Final output
Table ETBF_OBS_NULL_CATCH

Observer species catch per shot with variance

Prerequisites
- Procedure “Create observer sets per year, quarter and zone”
- Procedure “Create observer sets”

Queries used
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE_CREATE

Concept
This procedure creates a table of observer data per shot (inlcudes year, quarter and zone) and species. Catch per species and set, observed hooks per set, obsefved hooks per strata and catch per species and strata are included and variance is calculated for later use in table 4 using Ratio Estimate.

Procedure

1. Run query ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE_CREATE

Final output
- Table ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE

Observer species catch per voyage with variance

Prerequisites
- Procedure “Create observer sets per year, quarter and zone”
- Procedure “Create observer sets”

Queries used
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE
ETBF_OBS_CATCH_VOYAGE_SHOT_SPECIES_VARIANCE_CREATE

Concept
This procedure creates a table of observer data per voyage (inlcudes year, quarter and zone) and species. Catch per species per voyage, observed hooks per voyage, observed hooks per strata and catch per species and strata are included and variance is calculated for later use in table 4 using Cluster Estimate.

Procedure
1. Run query ETBF_OBS_CATCH_VOYAGE_SPECIES_VARIANCE_CREATE

Final output
- Table ETBF_OBS_CATCH_VOYAGE_SPECIES_VARIANCE

Create species 95 percentile

Prerequisites
- Procedure “Observer species catch per shot with variance”

Queries used
- ETBF_OBS_SPECIES_95PERC_CREATE
- ETBF_OBS_SPECIES_95PERC

Concept
This procedure calculates the 95 percentile per species

Procedure
1. Run query ETBF_OBS_SPECIES_95PERC_CREATE

Final output
- ETBF_OBS_SPECIES_95PERC

Create species 95 percentile for cluster estimate

Prerequisites
- Procedure “Observer species catch per voyage with variance”
- Procedure “Trips and hooks per strata for cluster estimate”

Queries used
- ETBF_OBS_SPECIES_95PERC_CLUSTER_CREATE
- ETBF_OBS_SPECIES_95PERC_CLUSTER

Concept
This procedure calculates the 95 percentile per species for Table 4 with cluster estimate

Procedure
1. Run query ETBF_OBS_SPECIES_95PERC_CLUSTER_CREATE

Final output

- ETBF_OBS_SPECIES_95PERC_CLUSTER

Table 4 using Ratio Estimate
Total estimated catch (retained and discarded) in ETBF

Prerequisites

- Procedure “Create observer sets”
- Procedure “Logbook effort per year, quarter and zone”
- Procedure “Create observer sets per year, quarter and zone”
- Procedure “Table1”
- Procedure “Observer species catch per shot with variance”
- Procedure “Create species 95 percentile”

Queries used
ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED_CREATE
ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL

Tables used
ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED

Procedure

1. Ensure that procedure “Create observer sets” has run
2. Ensure that procedure “Logbook effort per year, quarter and zone” has run
3. Ensure that procedure “Create observer sets per year, quarter and zone” has run
4. Ensure that procedure “Table1” has run
5. Ensure that procedure “Observer species catch per shot with variance” has run
6. Ensure that procedure “Create species 95 percentile” has run
7. Run query ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED_CREATE. This creates table ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED.
8. Run query OUTPUT_TABLE4_RATIO
9. Paste data into Excel spreadsheet as “Table 4 using Ratio Estimate”

Final output

- Table ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED

Table 4 using Cluster Estimate
Total estimated catch (retained and discarded) in ETBF

Prerequisites

- Procedure “Create observer sets”
- Procedure “Logbook effort per year, quarter and zone”
- Procedure “Create logbook vessels”
- Procedure “Create observer sets per year, quarter and zone”
- Procedure “Create logbook trips per strata”
- Procedure “Table1”
- Procedure “Observer species catch per voyage with variance”
- Procedure “Create species 95 percentile for cluster estimate”

Queries used
ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL_FORMATTED_CREATE
ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL

Tables used
ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL_FORMATTED

Concept
Note: Ratio estimate is based on voyages and shots. Cluster analysis is based on voyage. However a voyage may have sets in more than one zone. In this case, for cluster analysis, the voyage is assigned an arbitrary (first) zone that was visited in this voyage.

Procedure
1. Ensure that procedure “Create observer sets” has run
2. Ensure that procedure “Logbook effort per year, quarter and zone” has run
3. Ensure that procedure “Create logbook vessels” has run
4. Ensure that procedure “Create observer sets per year, quarter and zone” has run
5. Ensure that procedure “Create logbook trips per strata” has run.
6. Ensure that procedure “Table1” has run
7. Ensure that procedure “Observer species catch per voyage with variance” has run
8. Ensure that procedure “Create species 95 percentile for cluster estimate” has run
9. Run query ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL_FORMATTED_CREATE. This creates table ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL_FORMATTED.
10. Run query OUTPUT_TABLE4_CLUSTER
11. Paste data into Excel spreadsheet as “Table 4 using Cluster Estimate”

Final output
- Table ETBF_OBS_CATCH_ESTIMATE_PERCENT_TOTAL_FORMATTED

Table 6 using Ratio Estimate
Comparison of catch (retained and discarded) estimated from observer data and catch reported in logbooks for the ETBF

Prerequisites
- Procedure “Table 4”

Tables used
- ETBF_OBSLOG_CATCH_ESTIMATE_FORMATTED

Queries used
- ETBF_OBSLOG_CATCH_ESTIMATE
- ETBF_OBSLOG_CATCH_ESTIMATE_FORMATTED_CREATE

Concept
This procedure uses the observer catch estimates data created in procedure “Table 4 using the Ratio Estimate” (95 percentile based on sets) and adds a logbook catch estimate. Note that adding columns “Estimate Retained Catch” and “Estimate Discarded Catch” does not always equal column “Estimate Total Catch”. It may be off by +1 or -1 due to rounding errors since the estimates have decimal places and are rounded in one of the underlying queries.

Procedure
1. Ensure that procedure “Table 4” has run.
2. Run query ETBF_OBSLOG_CATCH_ESTIMATE_FORMATTED_CREATE. This creates table ETBF_OBSLOG_CATCH_ESTIMATE_FORMATTED
3. Run query OUTPUT_TABLE6_RATIO
4. Paste result into Excel spreadsheet as “Table 6 using Ratio Estimate”

Final output
- ETBF_OBSLOG_CATCH_ESTIMATE_FORMATTED

Changes
09.04.2008 Added columns “Estimate Retained Catch”, “Percent Retained Catch” and “Estimate Discarded Catch”.

Table 6 using Cluster Estimate

Prerequisites
- Procedure “Table 4”
- Procedure “Table 6”

Tables used
- ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER_FORMATTED
- ETBF_SPECIES_GROUPS

Queries used

- ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER_FORMATTED_CREATE
- ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER
- ETBF_OBS_CATCH_ESTIMATE_PERCENT_CLUSTER_TOTAL

Concept

This procedure uses the observer catch estimates data created in procedure “Table 4 using the Cluster Estimate” (95 percentile based on trips) and adds a logbook catch estimate.

Procedure

1. Ensure that procedure “Table 4 using Cluster Estimate” has run.
2. Run query ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER_FORMATTED_CREATE. This creates table ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER_FORMATTED
3. Paste result into Excel spreadsheet as “Table 6 using Cluster Estimate”

Final output

- ETBF_OBSLOG_CATCH_ESTIMATE_CLUSTER_FORMATTED

Table 3
Comparison of estimated number of fish in retained catch, and catch reported in logbooks for ETBF

Prerequisites

- Procedure “Logbook effort per year, quarter and zone”
- Procedure “Create observer sets”
- Procedure “Create observer sets per year, quarter and zone”

Tables used

- ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK_FORMATTED

Queries used

- ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK

Functions used

- FILL_ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK_FORMATTED

Concept

This procedure uses the final table created in procedure “Table 2”. It extracts commercial species, summarizes logbook catch and creates a percentage of observer catch to logbook catch.

Procedure

1. Ensure that procedure “Logbook effort per year, quarter and zone” has run
2. Ensure that procedure “Create observer sets” has run
3. Ensure that procedure “Create observer sets per year, quarter and zone” has run
4. Run function FILL_ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK_FORMATTED. This creates tables ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK_FORMATTED and OUTPUT_TABLE3.
5. Copy data from OUTPUT_TABLE3 to Excel spredsheet as Table 3.

Final output

- ETBF_OBSLOG_COM_CATCH_ESTIMATE_TO_LOGBOOK_FORMATTED
- OUTPUT_TABLE3

Table 5
Observed life status of discarded catch of fish and total catch of sea birds, turtles and marine mammals in the ETBF

Prerequisites

- Procedure “MSS codes”
- Procedure “Bio collection”

- Procedure “Create observer sets”

Queries used
- ETBF_OBS_SPECIES_LIFESTAGE_PERCENT_CREATE
- OUTPUT_TABLE5

Tables used
- ETBF_OBS_SPECIES_LIFESTAGE_PERCENT

Concept
This procedure produces a table of life stages of observed catch by species. Included are catch of groups Billfish, Other Fishes, Sharks and Rays and Tuna with fate other than “R” and groups Sea bids, Turtles and Whale with any fate code. Life status is queried to be 1, 2, 3 , 4 or 5.

Procedure
1. Run query ETBF_OBS_SPECIES_LIFESTAGE_PERCENT_CREATE
2. Run query OUTPUT_TABLE5

Final output
- ETBF_OBS_SPECIES_LIFESTAGE_PERCENT

Figure 3
Location of seabirds, turtles, and marine mammals observed in longline catch of the ETBF

Prerequisites
- Create observer sets
- Bio collection

Required tables
- ETBF_OBS_BIRDS_TURTLES_WHALES_SSSS
- ETBF_OBS_BIRDS_TURTLES_WHALES_[SSSS]

Required queries
- ETBF_OBS_BIRDS_TURTLES_WHALES_CREATE

Concept
Locations of observed seabirds, turtles and whales are extracted for maping in GIS.

Procedure
1. Run query ETBF_OBS_BIRDS_TURTLES_WHALES_CREATE. This creates table ETBF_OBS_BIRDS_TURTLES_WHALES_SSSS.
2. **Copy** table ETBF_OBS_BIRDS_TURTLES_WHALES_SSSS to ETBF_OBS_BIRDS_TURTLES_WHALES_[SSSS], where SSS is the correct season. Replace destination table if it exists.
3. Create map in ArcMap using W:\Mini_Projects\Eastern Tuna and Billfish Fishery Observer Data\ETBF.mxd:
    1. Import table ETBF_OBS_BIRDS_TURTLES_WHALES_[SSSS] and copy display attributes from another “ETBF Birds Turtles and Whales” layer.
    2. Enable “Label features in this layer” and set the font to Arial 6 point.
    3. Export this layer into the personal geodatabase “ETBF_GIS”.
    4. Delete the imported layer.
    5. Set all other data layers to inactive.
    6. Save map as “ETBF Birds Turtles and Whales [SSSS].png”

Final output
- ETBF_OBS_BIRDS_TURTLES_WHALES_SSSS
- ETBF_OBS_BIRDS_TURTLES_WHALES_[SSSS]

Additional Information

Prerequisites

All procedures up to and including procedure “Table 6 using Cluster Estimate”

Concept
The output of in this procedure is recquired by Jeff for additional analysis

Procedure

1. Ensure that all procedures up to and including “Table 6 using Cluster Estimate” have run

2. Run Query ADD_INFO_TOTAL_RATIO_ESTIMATE
3. Copy and paste the content of the query into a new Excel sheet named “Total Ratio Estimate”

4. Run Query ADD_INFO_TOTAL_CLUSTER_ESTIMATE
5. Copy and paste the content of the query into a new Excel sheet named “Total Cluster Estimate”

6. Run Query ADD_INFO_RATIO_ESTIMATE_PER_STRATA
7. Copy and paste the content of the query into a new Excel sheet named “Ratio Estimate per Strata”

8. Run Query ADD_INFO_CLUSTER_ESTIMATE_PER_STRATA
9. Copy and paste the content of the query into a new Excel sheet named “Cluster Estimate per Strata”

Final output

- Query ADD_INFO_TOTAL_RATIO_ESTIMATE
- Query ADD_INFO_TOTAL_CLUSTER_ESTIMATE
- Query ADD_INFO_RATIO_ESTIMATE_PER_STRATA
- Query ADD_INFO_CLUSTER_ESTIMATE_PER_STRATA

Appendix 4

An investigation into the relationship of bycatch species with target species and oceanographic environments in the Eastern Tuna and Billfish Fishery (ETBF)

Fiona M. Giannini and Emma K. Lawrence

INTRODUCTION

Australian Fisheries Management Authority (AFMA) observer programs provide an important role in the provision of data to fisheries research and in enabling improved scientific advice to fisheries stakeholders and managers. In the Eastern Tuna and Billfish Fishery (ETBF) observer coverage of 5% of logbooks sets is aimed for, as guided by the seabird Threat Abatement Plan (TAP).

The Ecological Risk Assessment for Effects of Fishing Project (ERA) highlighted a large number of gaps in the data and information available for assessing the sustainable nature of fishing practices (Hobday, A. et al., 2004). These data gaps include a lack of information on the biology of many of the species caught as well as a lack of understanding of the relationship between catch rates of bycatch species with fishing gear and practices, and surrounding oceanographic conditions. Exploratory analysis was conducted on observer data from 2001-05 from the ETBF with the aim of using this data along with information on oceanographic conditions during the time of fishing to enable a better understanding of the species' distribution and to provide a means of supporting management strategies of minimising bycatch where appropriate.

The result of analysis was two approaches to modelling species catch with the view of better understanding the relationship between target species, bycatch and oceanographic variables. The first approach was to model the presence-absence of target species using the catch of bycatch species as explanatory variables in an attempt to better quantify the relationship between the two groups. The second method involved modelling the catch of a select group of bycatch species in terms of oceanographic variables in order to study the effect of different environmental conditions on the catch of these species.

Other methods of investigating the relationship between target species and bycatch species were also considered in the exploratory stage of this project and these methods have been outlined with the aim of aiding future study of this data.

DESCRIPTION OF DATA

Observer data for the ETBF, provided by the AFMA for the years 2001–05 was used in this analysis. There were a total of 1802 sets observed with catch information on 115 different target and bycatch species. Appendix A gives a list of species names with their corresponding code names which will be used in this report. Oceanographic information on a set level was also provided (by CSIRO). The oceanographic variables are defined in Table 1. Fishing gear variables were considered for inclusion in the analysis but the data contained too many gaps to be of use. The percentage of zero catch sets for each of the species was found (see Appendix B). This table shows that a lot of the catch data for species was heavily zero inflated which was a consideration in the modelling process. Table 2 shows the breakdown of sets over years. The area in which these sets fall is from about 15 degrees south to 38 degrees south, and from 146 degrees east to 168 degrees east. Hook numbers ranged from 12 to 1940.

Table 1:Oceanographic variables used in analysis

Variable	Description
SEP07_CLUSTER	September output from Alistair Hobday
A	AGSO Bathymetry
B	GA 9sec 05 Bathymetry
SST	Filled from 1 to 6 day
C	CSIRO SST: 1 Day Composite
D	CSIRO SST: 3 Day Composite
E	CSIRO SST: 6 Day Composite
F	synTS CSIRO Temperature at Depth (100 m)
G	Australia Altimetry
H	synTS Eastward Geostrophic velocity (u)
I	synTS Northward Geostrophic velocity (v)
J	NCEP Wind Speed
K	SeaWiFS k490 Turbidity Data
L	SeaWiFS Chlorophyll A Concentration
M	AAsia MLD - synTS (CARS MLD definition)
N	SST Frontal Density

Table 2: Number of observed sets by year.

Year	***2001***	***2002***	***2003***	***2004***	***2005***
Sets	***45***	***410***	***385***	***425***	***537***

EXPLORATORY WORK

In investigating the relationship between species and oceanographic variables, a number of methods were explored. Cluster analysis and Principle Component analysis were applied to the data in order to look for group structure among the species as well as a relationship between species catch and oceanographic variables. This exploratory work is outlined as a reference for further study on this data.

CLUSTER ANALYSIS

Cluster analysis looks to discover underlying structure in data by sorting objects into groups with the association between two objects being maximal if they belong to the same group and minimal if they don't (Hill, T. et al., 2006). Different clustering algorithms were applied to the species catch data to look for structure in the data that would suggest a natural segregation of species into groups. Grouping the species would be useful in fisheries management as for a particular target species clusters may exist that mean it is more likely that bycatch species that belong to the same cluster as the target species will be caught. Relating these species clusters to habitat or oceanographic variables could also be helpful in better understanding the biology of the species.

Cluster analysis was performed on the data using simple hierarchical clustering algorithms with different distance measures and linkage rules. Most of these algorithms begin with each species being in their own group or cluster. At each step of the algorithm, the two most similar clusters are amalgamated until the last step where all objects or species are joined in one cluster. To decide whether two clusters are similar or not, distance measures are used and a linkage rule is agreed on. Single linkage refers to the method of linking two clusters when any two objects in the two clusters are closer together according to the distance measure than any objects belonging to other clusters. Complete linkage refers to using the distance measure of the two objects in the two clusters that are furthest away as defining the similarity between two clusters (Hill, T. et al., 2006). There are other linkage rules possible. The result of most of these clustering algorithms is a tree structure with each level relating to a different number of clusters from the total number of species to one cluster including all species. The tree should highlight some natural grouping if there is structure within the data. The analysis was conducted on both the raw and standardised data and in each case there was no significant evidence of clustering among the species catch data. The most promising result was obtained when using the Ward algorithm. The resulting tree seemed to suggest that there were three clusters of species but this result did not appear in any of the other analyses and so was not deemed significant.

Another method of cluster analysis, known as k-means clustering was also investigated. This method of cluster analysis assumes knowledge of how many clusters there are present in the data before conducting the analysis. The algorithm begins with k (the number of clusters in the data) randomly chosen clusters and then species are moved between these clusters with the goal to minimise variability within clusters and maximise variability between clusters. This analysis was attempted using various cluster numbers and was run several times for each k to mitigate any error due to the beginning chosen clusters. This method also did not produce any significant results.

The final method considered in cluster analysis was a model-based clustering algorithm which uses Expectation Maximisation (EM). This clustering algorithm computes probabilities of cluster memberships based on one or more probability distributions. The EM algorithm attempts to maximise the overall probability or likelihood of the species data given the clusters. This algorithm also found no significant results when looking to cluster species. As an exploratory step, the algorithm was rerun to consider clusters of sets. The algorithm found 16 to be an optimal number of clusters for segregating the sets. This was further investigated in the possibility that these clusters would identify a better way to stratify the fishery. Temporal and spatial ways of defining the clusters were attempted but there was no obvious definition apparent.

PRINCIPLE COMPONENT ANALYSIS

Principle Component Analysis (PCA) looks to transform the data to a new coordinate system such that the greatest variance of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on (Venables, W.N et al., 2002). The usefulness in PCA in the context of our species data is in finding structure in the data that suggests significant relationships between the catch of one species to another.

PCA was implemented on the presence-absence data of catch for all species. The first approach to PCA involved treating the species as the “variables” and each set as an “observation”. Results of PCA give the proportion of variance in the data explained by each principle component with the first principle component showing the largest proportion of variance. This analysis showed that even the first principle component only accounted for 2.16% of the variance. This suggests that there is little structure in the data and as a result viewing the data in terms of a smaller number of dimensions or components is not likely to have much benefit in cluster analysis. This was verified by creating a biplot of the data in terms of the first and second principle component.

As an exploratory step following from the investigation using EM clustering on a set level of the data, PCA was also applied to the presence-absence data treating the sets as variables and each species as an observation. As may have been expected, there was more structure evident when the data was viewed in this way. The first principle component accounted for 41.3% of the variance, with the first 22 components (out of 115) accounting for 90% of the variance. This seems to support the need to stratify sets when analysing catch data.

MODELLING

Two different approaches to modelling the data were considered. Both approaches aimed to increase knowledge of bycatch species, in particular in terms of relationships between species presence and oceanographic variables, and relationships between bycatch species presence with that of target species presence. The first approach looked at the latter relationship. The presence-absence data of 6 target species was modelled in terms of oceanographic factors as well as the catch rate of a selected group of bycatch species. The second method involved modelling a select group of bycatch species in terms of oceanographic variables.

TARGET SPECIES

Generalised Additive Models (GAMs), with a Binomial error distribution, were used to model the relationship between the presence of 6 different target species and the catch rate of other species along with 10 oceanographic variables (see Table 1 for a full list and definition of these variables).

The correlation of variables was investigated prior to modelling. It was found that there were three pairs of oceanographic variables that correlated with each other by more than 60%. For these three pairs only one variable from each was chosen to be considered in modelling to prevent errors occurring in the modelling process. The variables A and B both relate to bathymetry so only A was considered. K and L, relating to turbidity and chlorophyll concentration were highly correlated so L was chosen as a potential explanatory variable. Out of the variables SST (sea surface temperature) and F (temperature at 100m), SST was used in the modelling process. All other oceanographic variables were used in fitting models to the target species data. These other variables related to altimetry and wind speed. One categorical environmental variable was considered in modelling. SEP07_CLUSTER is an eight level variable where each level relates to a different set of environmental conditions or habitat. Each set or observation has a cluster level associated with it.

The target species modelled were Yellowfin Tuna (YFT), Bigeye Tuna (BET), Albacore Tuna (ALB), Broadbill Swordfish (SWO), Striped Marlin (MLS) and Southern Bluefin Tuna (SBF). The correlations between the catch of these 6 species with the catch data of all other species was calculated and those species associated with a correlation of greater than or equal (in absolute terms) of 0.1 were grouped with each target species (see Table 3).

Table 3: Target species and associated species in terms of correlated catch. Species with a correlation of greater than or equal (in absolute terms) to 0.1 are identified as being possibly associated to the target species.

Target Species	Correlated Species
ALB	BET, BSH, DOL, SBF, SMA, STI
BET	ALB, BTH, SWO
MLS	DUS, RMB, SKJ
SBF	ALB, BSH, BSK, EGD
SWO	ALX, BET, BSH, DKK, MOO, RIB, STI, YFT
YFT	ALX, BSH, LEC, MAA, SKJ, SWO

The catch rates of the species grouped with each of the target species were then used as a variable considered in modelling the presence of the target species.

The models were fitted using a forward stepwise method to add explanatory variables to the model from the full set of oceanographic and species catch rate variables available. The first step in this method is to model each potential explanatory variable individually and look at the "proportion deviance explained" statistic to determine which variable explains the variation in the target species presence data best. This proportion compares the deviance of the null model (the model with just an intercept term) to the deviance of the fitted model. The variable with the highest proportion deviance explained becomes the first explanatory variable of the model. Following steps involve looking at the proportion deviance explained of all possible models that combine the intermediary model with remaining potential explanatory variables and picking the best of these variables to add to the model. This process is terminated when no remaining potential variables add more than 0.005 to the proportion deviance explained to the intermediary model. The final model only keeps the significant variables (those with a p-value of less than 0.05). Table 4 shows the significant variables when modelling each of the target species.

Table 4: Significant variables for target species

Target Species	Significant Explanatory Variables
ALB	M, SST, BET_CR, A, SEP07_CLUSTER, I, H
BET	ALB_CR, A, SST, M, I, SEP07_CLUSTER
MLS	SST, A, L
SBF	SEP07_CLUSTER, ALB_CR, G, A
SWO	YFT_CR, J, BET_CR, BSH_CR, SEP07_CLUSTER, SST, L, I
YFT	L, A, M, I

Table 4 shows that only SWO has the catch of a bycatch species (BSH) as being a significant variable in the resulting model. ALB, BET and SBF all have the catch of another target species as being significant in the model. These results show that though this may be a useful approach to understanding the relationship between target species and bycatch species for some combinations of species, this approach may not be useful for all species.

It should also be noted that this approach as only been used to determine whether there is a significant relationship between target and bycatch species catch rates. Using bycatch or other target species as explanatory variables when estimating trends in abundance is not recommended due to the potential of trends in the explanatory variable abundance indices to distort the trend in the target species.

BYCATCH SPECIES

The main difficulty faced with when modelling bycatch species is that the catch data are largely made up of zeros. In some cases over 99% of the data is zeros (see Appendix B). In these cases it is difficult to get a good model to fit to the data, even when looking at the data as a Binomial set of presence and absences rather than modelling the actual catch count of a particular species. For this reason, a select group of bycatch species was modelled where the percentage of zeros in the catch data was not higher than 94%. For these 4 species a two-stage model was fitted which involves modelling the chance of a non-zero catch and the catch rate separately. As there is generally little non-zero data, there are restrictions on the number of variables added to the model and specifically the degrees of freedom used to model the catch rate of a particular species to prevent the model being overfitted.

The species modelled fit the criteria of having less than 94% zeros and each of the species had high levels of discard. The species modelled were Sunfish (MOP), Black Marlin (BLM), Lancetfish (ALX) and Snake Mackeral (GES). The two models fit to each of the bycatch species were GAMs where the model fit to the presence-absence species data had a Binomial error distribution and the model fit to the catch count (where catch is non-zero) used a truncated Poisson error distribution that takes into account the absence of zeros from the data. The method used to model the data was similar to the modelling of the target species. The same 10 oceanographic variables were considered as potential explanatory variables and a forward step-wise approach was used to fit the models. Only significant terms (see Table 5) were kept and in the case of the truncated Poisson error model, the degrees of freedom of the resulting model were monitored to prevent overfitting.

Table 5: Significant variables for bycatch species

Bycatch Species	Binomial Model	Truncated Poisson Model
MOP	SST, J, G, A	SEP07_CLUSTER, SST
BLM	SST, SEP07_CLUSTER, A	G, SST, M, H
ALX	I, A, G, N, H, J	SST, A, G, J, I
GES	J, SEP07_CLUSTER, SST	SST, G, I, SEP07_CLUSTER

Table 5 shows that many variables such SST, G and SEP07_CLUSTER, are consistently significant in the catch rates of bycatch species.

CONCLUDING REMARKS

This report outlines two approaches to modelling using the catch of bycatch species with the aim of gaining a better understanding of these species' biology and behaviour in support of fishery management. The first approach models the presence-absence of target species in terms of correlated bycatch species and oceanographic variables in order to define relationships between the species that are targeted in fishing and those caught as a consequence. The second approach looks at the relationship between bycatch species with high levels of discard with oceanographic variables. Both approaches have merit in providing management with information on bycatch species that could be used in support of strategies to minimise bycatch.

REFERENCES

Hill, Thomas and Lewicki, Pawel (2006). Statistics: Methods and Applications, First Edition, StatSoft.

Hobday,A., Stobutzki,I.m Webb, H. and ERA team (2004), Ecological Risk Assessment for Effects of Fishing: eastern Tuna and Billfish Fishery. Draft report 17 May 2004.

Nurmi, Petteri. Mixture Models. http://www.cs.helsinki.fi/u/salmenki/lda-seminaari04/mixturemodels.pdf

Venables, W.N and Ripley, B.D (2002). Modern Applied Statistics with S, Fourth Edition, Springer

APPENDIX A

Species' names relating to code names used in report.

Code Name	Species Name
ALB	Albacore Tuna
ALX	Lancetfish
BET	Bigeye Tuna
BLM	Black Marlin
BSH	Blue Shark
BSK	Basking Shark
BTH	Bigeye Thresher
DKK	Leatherback Turtle
DOL	Dolphinfish
DUS	Dusky Shark
EGD	White Cardinalfish
GES	Snake Mackeral
LEC	Black Oilfish
MAA	Blue Mackerel
MLS	Striped Marlin
MOO	Razor Trevally Moonfish
MOP	Sunfish
RIB	Ribaldo
RMB	Manta Ray
SBF	Southern Bluefish Tuna
SKJ	Skipjack Tuna
SMA	Shortfin Mako
STI	Dwarf Black stingray - Pelagic stingray
SWO	Broadbill Swordfish
YFT	Yellowfin Tuna

APPENDIX B

Percentage of zero catches for each species.

Species	*%*	*Species*	*%*	*Species*	*%*	*Species*	*%*	*Species*	*%*
ACK	***99.83***	***BSK***	***99.94***	***DOL***	***59.21***	***MOP***	***78.80***	***SHW***	***99.94***
ACU	***99.94***	***BTH***	***96.50***	***DUS***	***97.95***	***OCS***	***93.40***	***SKJ***	***81.58***
AKB	***99.89***	***BUM***	***92.12***	***EGD***	***99.94***	***OIL***	***87.46***	***SKX***	***99.94***
ALB	***28.19***	***CBA***	***99.83***	***FAL***	***98.61***	***OTH***	***99.50***	***SMA***	***68.70***
ALK	***99.89***	***CCA***	***99.94***	***GBA***	***99.00***	***PCW***	***99.78***	***SNK***	***99.67***
ALO	***92.29***	***CCB***	***99.94***	***GEM***	***99.94***	***PDM***	***99.33***	***SPI***	***99.94***
ALS	***99.94***	***CCL***	***99.89***	***GES***	***90.07***	***PFC***	***94.78***	***SPL***	***99.89***
ALV	***97.00***	***CCU***	***99.72***	***GLO***	***99.94***	***PFG***	***99.89***	***SPN***	***95.28***
ALX	***47.95***	***CEO***	***96.28***	***HUW***	***99.94***	***PFT***	***99.50***	***SQQ***	***99.94***
ALZ	***99.89***	***CSK***	***99.94***	***ISB***	***99.56***	***PFZ***	***99.61***	***SSP***	***90.57***
BAC	***99.00***	***CUP***	***99.89***	***KAW***	***99.78***	***POA***	***86.40***	***STI***	***86.35***
BAR	***98.95***	***CUX***	***99.83***	***LAG***	***97.84***	***POR***	***99.61***	***SWO***	***40.34***
BAU	***99.72***	***CVX***	***98.61***	***LAN***	***99.94***	***PSK***	***99.00***	***SXX***	***99.94***
BDL	***99.94***	***DAC***	***99.78***	***LEC***	***48.45***	***PTH***	***98.17***	***TAL***	***99.67***
BEH	***99.33***	***DCO***	***99.89***	***LKV***	***99.78***	***PTZ***	***99.89***	***TIG***	***94.45***
BET	***47.17***	***DCR***	***99.94***	***LMA***	***99.22***	***PUX***	***99.45***	***TRP***	***98.34***
BFT	***99.33***	***DCU***	***99.83***	***MAA***	***99.67***	***REO***	***99.94***	***TTH***	***99.94***
BIL	***99.78***	***DIB***	***99.94***	***MAN***	***99.94***	***RIB***	***99.94***	***TTL***	***99.83***
BLM	***92.51***	***DIC***	***99.89***	***MAR***	***99.00***	***RMB***	***94.40***	***TUG***	***99.56***
BRA	***99.56***	***DIM***	***99.56***	***MEW***	***99.94***	***RSK***	***99.94***	***UNK***	***86.13***
BRO	***94.17***	***DIP***	***99.94***	***MKT***	***99.72***	***RUZ***	***99.94***	***WAH***	***96.28***
BRU	***99.89***	***DIX***	***99.61***	***MLS***	***67.81***	***SBF***	***90.68***	***YFT***	***12.04***
BSH	***68.53***	***DKK***	***99.06***	***MOO***	***99.89***	***SFA***	***98.61***	***YTC***	***99.50***

Appendix 5

Defining dynamic pelagic habitats on the east coast of Australia

A. J. Hobday[1]
J. W. Young[1]
J. M. Dambacher[1,2]
C. Moeseneder[3]

1. Wealth from Oceans National Research Flagship, CSIRO Marine and Atmospheric Research, Hobart, Tasmania, 7001
2. CSIRO Mathematical and Information Sciences, Hobart, Tasmania, 7001
3. CSIRO Marine and Atmospheric Research, Cleveland, Queensland

In review for **Deep Sea Research** EAC issue

Abstract

Although many species in the pelagic ocean are widespread, they are not randomly distributed. These species may have associations with particular water masses or habitats, but to best understand patterns in the ocean, these habitats must be identified. Previous efforts have produced static or seasonal climatologies, which still represent smearing over habitats. The eastern tuna and billfish longline fishery (ETBF) targets a range of high trophic level species throughout the east coast of Australia. In this study, dynamic ocean habitats in the ETBF region were identified for each month based on cluster analysis of a set of five oceanographic variables averaged at a monthly time scale and a spatial scale of 0.5 degrees for the period 1997–2006. A total of seven persistent habitats were identified with intra and interannual variation in size and location, indicating the importance of spatial and temporal variation in the dynamics of the ETBF region. The degree to which these dynamic habitats were distinguished was tested using (i) stable isotope analysis of top fish predators caught in the region, and (ii) estimates of variation in estimated abundance generated from catch data from the fishery. More precise estimates of isotopic values based on samples from swordfish, yellowfin tuna and albacore were obtained using the dynamic habitat groupings, which indicates that the habitats identified had underlying differences in ecosystem structure. With respect to estimation of fish abundance, using estimates based on habitats produced more precise estimates for seven of eight species. These findings could be used to guide development of effective monitoring strategies that can distinguish between patterns due to environmental variation, including in the longer term, climate change.

Keywords: habitat characterization, habitat association, isotope signal, abundance estimation

Introduction

While fishermen have always ranged widely in search of particular species and high catches, technology now allows humans to visit every portion of the planet, and to extract resources from those regions. Not surprisingly, fishing impacts are evident even in the pelagic ocean far from the coast (Baum et al., 2003; Halpern et al., 2008). As a result, there is a growing need to understand the impacts and to reduce activities that threaten sustainability. If marine environments are to be systematically monitored for the effects of human activities, then identification of the types of marine habitats and the communities they contain, and delineation of their boundaries is required (Roff et al., 2003). Human impacts on these environments can then be assessed, health of these environments and the dependent ecological communities can be monitored, and management strategies designated (Roff et al., 2003; Roff and Evans, 2007).

However, as each animal or plant in the ocean has a particular requirement, and so in reality each has a unique habitat, an intermediate level of complexity in identifying habitats is required. As on land, a useful habitat definition might be based on a collection of similar processes or structures. In the ocean, benthic habitats have been described on the basis of substrate types (Williams and Bax, 2001; Roff et al., 2003), while pelagic habitats for a number of species have been described according to values of one or more physical variables such as temperature, salinity, oxygen content, and chlorophyll (e.g., Lynn and Simpson, 1987; Etnoyer et al., 2004; Zainuddin et al., 2006; Hobday and Hartmann, 2006).

In recognition that the ocean is not homogenous, many workers have shown that spatially coherent properties do allow definition of persistent regions, for example, the classification by Longhurst et al. (1995) recognized a number of regions in the worlds' oceans. Most schemes, however, are coarse representations only, as the ocean does not follow straight lines, and the distribution of the habitats is also likely to vary over time. More natural classification schemes have recently been developed with boundaries that more closely follow oceanographic boundaries and recognize seasonal differences (Devred et al., 2007). Around Australia, a number of static bioregionalisation schemes have been developed for the coastal zone, for example, the IMCRA bioregionalisation (http://www.environment.gov.au/coasts/mbp/imcra/nmb.html) and more broadly for the offshore bioregional provinces (Condie and Dunn, 2006).

Studies of the distribution of biological patterns in the ocean by Etnoyer et al. (2004) and Palacios et al. (2006) have also recognized that biological pattern within these so-called habitats differ from those in surrounding waters. Most bioregionalisations to date, however, have provided an average description of the habitat distribution over time, and ignored the temporal component to the spatial distribution. In regions with highly seasonal dynamics, however, the same fixed position in the ocean may be covered by different habitats, depending on the time of year. On the east coast of Australia, for example, this variation in location of water masses is driven by the north-south seasonal expansion and contraction of the East Australia Current (Ridgway and Godfrey, 1997; Ridgway, 2007). Habitat classifications that allow the dynamic nature of the ocean to be included in the classification approach are needed, particularly when trying to understand biological pattern. Defining these habitats is important, as species have particular affinities for different marine habitats (e.g., Williams and Bax, 2001; Block et al., 2003; Polovina et al., 2004; Sydeman et al., 2006). If dynamic habitats can be described, then habitat-specific management strategies can be developed, which allows for more efficient management (e.g., Hobday and Hartmann, 2006; Sydeman et al., 2006; Hobday et al., 2009).

With regard to eastern Australia, and the eastern tuna and billfish fishery (ETBF), there is recognition that some fish species are less abundant than in the past, and that fishers must travel further to maintain high catch rates (Campbell and Hobday, 2003; Campbell 2008). Management regulations, such as the discarding of sharks and trip limits for other species such as dolphinfish, have been implemented as a way to reduce impacts of fishing in the ETBF. While laudable, these management regulations have been implemented without a clear understanding of the impacts on the exploited ecosystem. Clear definition of the exploited ecosystem(s) is in fact lacking. A key research question is determining the number of ecosystems in the ETBF region, where ecosystems can be defined as a closely interacting set of species and ecological processes (O'Neil, 2001).

Here for the east coast of Australia, we develop an approach to describe the time-varying distribution of pelagic habitats that are utilized by the fishers of the ETBF. This will allow impacts of fishing to be better resolved from the complex physical variation in the system. A habitat description based on physical characteristics can then be used to classify species patterns, and determine the level of habitat association in this region. Accordingly we present a classification of pelagic habitats relevant to the ETBF, the results of which are tested in two ways (i) precision of estimates of stable isotope concentrations of ^{15}N and ^{13}C in pelagic fishes, and (ii) estimation precision of the total abundance of pelagic species captured in the region of interest. Identification of these dynamic habitats may also inform the design of monitoring strategies to resolve patterns of natural system variation and to follow the progress and effect of management regulations.

Methods

To identify pelagic habitats, the eastern Australia region was divided into grid cells of size 0.5° latitude and 0.5° longitude in the region enclosed by 150–160°E and 20–38°S (**Figure 1**). This region was selected on the basis of available physical data, and the core location of the ETBF fishery. Only cells with an average depth of more then 100 m were retained, as this is the minimum depth that longlines are set in this fishery. A variety of marine datasets were considered (**Table 1**), although not all were used in the final habitat descriptions (see Results). These data were accessed from the Spatial Dynamics Ocean Data Explorer, a tool which assembles a range of oceanographic data at CSIRO (Hobday et al., 2006). Data exist at a range of spatial and temporal scales, and so were aggregated to the common spatial and temporal scale (monthly).

To define habitats we used cluster analysis (Bridge 1993). Cluster analysis is a technique frequently used in climatology for grouping cases to define classes (synoptic types or climate regimes, for example), or for grouping grid points to define regions (e.g., Mimmack et al., 2001). Average values for each month for the years 1997–2006 in each 0.5 degree grid cell were generated for all variables, and then using a subset of the variables values, the cells were allocated to clusters using several purpose-built MATLAB® functions, that relied on inbuilt MATLAB® codes for cluster identification (e.g., cluster.m). A variety of combinations of data sets were explored in defining the clusters, some of which are time invariant (e.g., bathymetry), or annual climatologies (e.g., nitrate, phosphate). To avoid biasing the clusters by including too many climatological data sets (**Table 1**), only a single climatology was used in the final habitat determination. Cluster analysis is based on some form of distance matrix, and the most commonly used metric in the climatological field has been Euclidean distances (Mimmack et al 2001). Several other approaches to classification were explored (PCA and non-PCA based classification), and data were transformed to produce variables on a similar numerical scale, which is important for cluster algorithms (Bridge 1993).

Pelagic habitats were identified using cluster analysis for each month in the period 1995–2006. The start year was dictated by the availability of environmental data, and when the ETBF data to use for testing the habitat classification are considered reliable (Campbell and Hobday, 2003). The end year (2006) was chosen based on availability of complete environment datasets (Hobday et al., 2006). The variation in size, location and persistence of the final habitat set is described, before examining if biological pattern reflects this dynamic physical subdivision of the pelagic ocean.

Relationship to biological pattern

To test if the division into physical habitats was reflected in the biology of the region, two tests were conducted. The basic premise in each was to see if classification of biological samples into these spatially and temporally variable “habitats” produced more precise results than classification into a northern and southern habitat, based around an average and constant latitude. If so, then the biology is reflecting the physical pattern. If not, this indicates that for the biological examples chosen, the physical habitats are not limiting the biological patterns.

The first test involved examination of the precision of stable isotope estimates in large pelagic fishes captured in the region. Stable isotope analysis (SIA) has developed as a useful method by which changes within and between adjacent environments of can be assessed relatively quickly (Hobson and Wassenaar 1999). A number of ^{15}N and ^{13}C samples were taken from a range of species in the ETBF over the period 2004-2006 (Young et al 2009; Revill et al 2009) and were used to generate a test data set. The three species with 30 or more ^{15}N and ^{13}C isotope samples were yellowfin tuna, albacore tuna, and swordfish. Limited SIA data confined this test to two dynamic groups, cluster set 1 [3,4,5, 6] and 2 [1,2,7], and comparison was made to a northern and southern static group based on the average latitude of separation between the clusters with the most data (cluster 3 and 7) (33.75°S). The precision of the estimated mean valu es for each group was used to determine if cluster allocation reduced the variation in the isotopic data, indicating biological coherence within these oceanic habitats.

In the second test, observed catch from a large number of species captured in the ETBF fishery and appearing in logbook and observer records were allocated to each of the appropriate dynamic clusters and to static northern and southern clusters, as for the SIA test. As most data fell into the same two habitat clusters (3 and 7), analysis was also restricted to a two-group comparison. Levels of catch in cluster set 1 [3,4,5, 6] and 2 [1,2,7] were generated according to a ratio estimator, based on fish caught per number of hooks deployed, and variances according to variability of catch in long-line sets within a trip (Cochran 1977, i.e., this is formally known as a cluster estimate of variance, where long-line sets within a trip form a statistical sampling cluster, not to be confused with clusters based on similar ocean conditions discussed elsewhere in text). The precision of these catch estimates was used to examine if the clustering approach led to more precise estimates of fish abundance than using the static north-south habitat division at the mean latitude separating cluster 3 and 7 (33.75°S). Specifically, the precision (CV) from the estimates of abundance based on the two dynamic cluster sets was compared to the precision based on abundance estimates from static grouping of the fish within the ocean region, using the combined logbook (to provide the number of hooks and sets) and observer data (number of individuals caught per hook).

Results

Defining ocean habitats

The five variables used to define the final oceanic habitat clusters were bathymetry (time-invariant), sea surface temperature, temperature at 250 m, chlorophyll, and nitrate climatology (**Table 1**). The excluded variables were highly correlated with variables that were included (e.g., turbidity and chlorophyll), or were annual climatologies. Use of more than one climatology data set reduced the interannual signal in the habitat distribution (data not shown), and so nitrate was selected as a representative and biologically important variable.

Clusters based on PCA analysis (the principal components were used as the variables for clustering) and on the transformed variables themselves were very similar, and so for ease of interpretation the standard (non-PCA) approach (Mimmack et al., 2001) is presented. An example of the clustering output from the first stage of the standard approach is shown in **Figure 2** (the PCA-approach was identical for this step). PCA clusters based on the same five variables as the standard approach were very similar to those based on all 11 variables, indicating the correlation between environmental variables (**Figure 3**). This similarity underpinned the choice of just five variables for the standard clustering approach. In exploratory analyses, the number of clusters was not specified, but after examination of a range of cutoffs, and the distribution of the resulting patterns, up to 10 clusters were initially allowed by the algorithm. Only clusters composed of more than 10 cells were retained, cells from smaller clusters were allocated to the surrounding cluster. Thus, for each year-month, between 1 and 10 clusters could result. Spatial mapping of the resultant clusters showed coherence between closely located cells and for the majority of habitats, over consecutive months (**Figure 4**).

After analysis for all months and years, resulting in 144 habitat classifications, a total of seven persistent clusters (habitats) were identified for the ETBF study region (**Table 2)**. Because of the temporal variation in the distribution of these habitats we show as a summary the centre of these habitats for all months and years in **Figure 5**. Variation in the location of these habitats as indicated by the movement of the centre of these habitats is notable, supporting the development of non-static habitat descriptions for this dynamic ocean region. An alternative view is a probability map for each habitat over the time period (**Figure 6).** This shows that some regions of the ocean are always classified as one habitat type (e.g., the north-east portion of the study region is always classified as Cluster 2, West Pacific), while in other regions (e.g., south-east Australia, Cluster 4 is present less than 7% of the year-month combinations), habitats are quite ephemeral. The dynamic overlap between habitats can also be visualized in this way.

Values (e.g., mean temperature) representative of each habitat are not illustrative, as the clustering is based on relationships between variables each year-month combination. Thus, a typical temperature of say, cluster 2 in the north, cannot be defined. This can be further explained by recognizing that the same cluster is present in both summer and winter, and so a mean value of temperature would not be representative of either the summer or the winter temperature. With the exception of bathymetry, the same is true for the other variables.

Habitat persistence over time

Not all seven habitat clusters were equally present over time (**Figure 6**, **Figure 7**). The Tasman Sea and West Pacific habitats were present in all months but one, while the southern eddy habitats (habitat 4 and 5) were the most infrequent habitats both within and across years. The variation in the latitudinal extent and the size of the habitats over time is shown for the Coral Sea in **Figure 8**. This habitat was in the northern part of the study region, and so the northern boundary was generally at the maximum latitude of the study region. The southern boundary varied by over 10 degrees of latitude between different months – evidence of the dynamic nature of the habitats in the region. There was also an apparent decline in area of the Coral Sea habitat over time. This is due to a reduction in the longitudinal extent of the habitat, as the latitudinal bounds did not display a northern or southern change (Figure 8). Conversely, the size of the adjacent habitat to the east (e.g., West Pacific) increased over time (data not shown).

Seasonal patterns in habitat size

If a strong seasonal pattern in the location of the habitats exists, then a seasonal average could be a useful tool for management, and reduce the need for generating a habitat description each month. If there is not a strong seasonal pattern, then fisheries managers (for example) would need to consider using specific month-year combinations, and would need these descriptions in near real-time, as for the SBT habitat prediction used by AFMA (Hobday and Hartmann, 2006; Hartog et al, this volume). The analysis of the seasonal signal in the location of the northern and southern boundaries, as for the interannual analyses, showed high variation (**Figure 9**). For example, while the northern boundary of Habitat 2 (West Pacific) was always at the top of the study region (20°S), there was a seasonal trend in the location of the southern end of this habitat. In summer, the southern boundary of this habitat was further south on average (30°S) than in winter (28° S). Considerable variation around the average location of the southern boundary is evident. This illustrates that using a single value for the location of boundaries is not appropriate.

Similarly, for southern habitats, the southern boundary was often at the edge of the study domain (38°S), while the north boundary varied in location. For most of the habitats, however, there was high variation and hence no strong seasonal signal (**Figure 9**). Thus, a seasonal average, while useful, would not be the best way to inform management of this region. Despite this, for a summary only, an average location of the habitat edges and boundaries may be useful for some purposes (**Table 3**). The extent of each boundary was defined as the mean of the most extreme 5% of cells comprising each habitat in each of four directions (N, S, E, and W). Thus, the mean boundary between the Coral Sea habitat (southern boundary) and the habitat to the south, the Central habitat is at 27.79°S. This is not strictly the converse (the mean northern boundary of the Central habitat is at 25.79°S), as the habitats also interact with other habitats to each side and interweave (i.e., not a straight boundary). The largest habitat was the West Pacific, followed by the Central, and Tasman Sea (**Table 3**). The ephemeral habitats (Eddies and the Southern Ocean) were smallest in average size.

Testing biological relationships to physical habitats

The first test of the usefulness of the dynamic description of ocean habitat was based on stable isotope data collected in the east coast region for large fishes. A total of 30, 34 and 43 ^{15}N and ^{13}C samples were available for swordfish, albacore tuna and yellowfin tuna respectively. The total CV for the mean ^{15}N and ^{13}C isotope values for each species were compared between the two classification approaches (dynamic and static) (**Figure 10**). The sample sizes are small for each species, however, the dynamic clusters showed a smaller total CV for ^{15}N all three species (8 – 45% improvement, mean 22%). For ^{13}C, the dynamic clustering resulted in more precision for albacore tuna (7% lower CV), and less precision for yellowfin (4% higher CV) and swordfish (26% higher CV). Overall, the dynamic clusters resulted in more precise estimates (lower CV) of mean isotope values for 4 of 6 comparisons.

The second test involved estimating abundance and variation in catch within regions of the east coast. The top eight species captured in the ETBF were considered, each with more than 1500 records for the period 2001-2006, together these represented over 91% of individuals observed in the ETBF longline fishery in the period 2001-2006 (**Figure 11**). For seven of the eight species, the precision of estimated abundance was greater when clusters were used compared with static division into a northern and southern group. Only yellowfin tuna (YFT) abundance estimates were less precise when the clusters were used, and the difference was only 3.8%. The improvement in precision (CV) for the other seven species ranged from 9% (Rudderfish and Black Oilfish) to 35% (bigeye tuna). Overall, the average improvement in precision of abundance estimates when the clusters were used across the eight species was 19%.

Discussion

In many marine studies, space is considered to be constant over time. For example in studies that seek to explain species distribution, space is often included as a single variable, such as latitude or longitude, while time is included as a second variable, such as month or year. This representation, which implicitly assumes a static ocean, is used in the absence of any suitable description of the habitat, and is common in fisheries oceanography where the goal is to determine influences on species abundance or distribution (e.g., Bigelow et al., 1999; Agenbag et al., 2003). We have shown that, based on cluster analysis of oceanographic variables, the east coast region is comprised of seven dynamic pelagic habitats. These are large and generally always present, but vary in size within and between years. Within each habitat, there may be biological assemblages and foodweb structures that differ from surrounding habitats. Thus, analysis of biological pattern in the fisheries in this region (e.g., the ETBF), such as diet studies, catch analysis, and management strategies should consider accounting for habitat type as we have described.

Cluster-based identification of ocean habitats

The advantage of defining habitats for each month is that the description is spatially flexible, and recognizes the variability in the ocean. The disadvantages of using clustering approaches are that determining the number of clusters is qualitative (Bridge 1993). Although rules for cutoffs to the number of clusters can be created, in practice these rules generally are to inflexible over a wide temporal or spatial range. A second “disadvantage” is that clustering data will always give a difference – i.e., clustering identifies clusters, even if nothing in the data justifies separation. However, if clusters are spatially coherent, then the clusters are sensible and not just creating artificial divisions in the data (Mimmack et al 2001). This was the case for the pelagic habitats identified in the region of interest.

Cluster sensitivity to number of variables

Identification of the pelagic habitats was based on oceanographic variables. In preliminary exploration, increasing the total number above five variables did not lead to major changes in the emergent habitat patterns. In part, this is expected, as the pelagic habitats correspond to major water masses or features in the region, within which variables are correlated. More than a single variable is needed to distinguish these habitats, as similar values can occur in very different and remote water masses. For example, temperature can be warm due to surface heating, or to tropical origin.

Seven dynamic pelagic habitats were identified via cluster analysis in the region of the ETBF, and these habitats were present for between 14 and 143 months of 144 possible months. The three largest clusters were most persistent, and represent the dominant water masses in the region, the West Pacific, Tasman Sea and Central water mass. While seven clusters were identified, for some purposes, for example where data on biological distributions are limited, habitats may need to be combined, as for the two tests of the relationship to biological patterns. This combining may be done on the basis of proximity, or on distinctness.

Relationship of physical habitats to biological description

We showed in two cases that biological patterns reflected these dynamic habitats better than did static grouping. We used a measure of precision, rather than mean values for each grouping, as we were interested in determining if allocation to dynamic habitats reduced the variation in the estimated values, indicating that “like samples” were being grouped.

The first test, based on isotope samples, showed precision of estimates was greatest for dynamic habitats for 4 of 6 comparisons involving three large pelagic fishes. We have previously shown that off eastern Australia the ^{15}N signal of a range of top fish predators was distinguishable across a range of latitudes (Revill et al 2009), and this goes further to show that these signals might reflect habitats as opposed to a latitudinal gradient.

The second test of the utility of these dynamic pelagic habitats, in estimating precision of abundance estimates for the number of individuals captured in a longline fishery, showed that for most species considered, use of the dynamic habitat clusters resulted in more precise estimates of abundance than using static groups. The improvement is likely due to species having affinities for particular water masses, that are then reflected in the probability of capture in the longline fishery.

Value in identification of pelagic habitats

Overall, these results illustrate that the east coast is a dynamic oceanic region (Ridgway 2007), with recognizable ocean habitats showing seasonal variation in location and extent persistent over a number of years. This has important implications for ecosystem monitoring and attribution of change in the region. For example, changes in the abundance of a particular tuna species may be due to changes in exploitation, or to changes in the availability or location of the preferred habitat. Correct assignment of the cause of a change is crucial to an effective management response.

These pelagic regions could also be used in design or allocation of monitoring effort. For example, monitoring of ecosystem indicators such as stable isotopes or fatty acid signals that represent patterns in energy flow through the system (e.g., Jennings et al 2002), could be done within each habitat. Real-time identification and mapping of these pelagic habitats (e.g., Hobday and Hartmann 2006) could allow near-real time monitoring or sampling, while the historical description of these pelagic habitats can allow retrospective understanding of changes in the east coast system.

Identification of pelagic habitats may also be important for ecosystem understanding. For example, the move to ecosystem-based fishery management (EBFM) has seen increased importance placed on development of food web models (Link et al., 2002; Health 2005). Construction of both quantitative and qualitative ecosystem models requires information on species occurrence (e.g., from logbook data) and the relationship between those species (based on foodweb diet analyses) (Watters et al., 2003; Dambacher et al., 2003; Dambacher et al., 2008). We showed in the second validation example, that abundance of species varied between the two habitat types considered. Appropriate models could be developed for each of the major habitat types, just as has been recognized for inshore and benthic locations (e.g., Davenport and Bax, 2002; Heath 2005).

Conclusion

The methods used here to identify persistent but spatially variable habitats have implications for how the ocean might be monitored, managed, and greater understanding for pelagic ecosystems achieved. Physical patterns in ocean structure were reflected in the biological patterns, indicating that oceanic habitat structure matters. For realtime monitoring, the challenge for the future will be to generate these habitats in real time, however, for many purposes post-hoc identification of habitat location will be suitable. Future effort should be directed at explaining the variation in the size and extent of these habitats, as well as understanding the biological affinity for these habitats.

Acknowledgements

Data used for the environment-based habitat classification was sourced from a range of providers detailed in the SDODE data libraries (Hobday et al 2006). This research was supported by CSIRO and FRDC (project 2004/063).

References

Agenbag, J.J., Richardson, A.J., Demarcq, H., Freon, P., Weeks, S. and Shillington, F.A. (2003). Estimating environmental preferences of South African pelagic fish species using catch size- and remote sensing data. Progress in Oceanography, 59, 275–300.

Baum, J. K., R. A. Myers, D. G. Kehler, B. Worm, S. J. Harley and P. A. Doherty (2003). Collapse and Conservation of Shark Populations in the Northwest Atlantic. Science **299**: 389-392.

Bigelow, K.A., Boggs, C.H. and He, X. (1999). Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fisheries Oceanography, 8, 178–198.

Block, B.A., Costa, D.P., Boehlert, G.W and Kochevar, R.E. (2003). Revealing pelagic habitat use: the tagging of Pacific pelagics program. Oceanologica Acta, 5, 255–266.

Bridge, P. D. (1993). Classification. *In* Biological data analysis : a practical approach. (ed J. C. Fry). Oxford, IRL Press at Oxford University Press**:** 219-242.

Campbell, R.A. and Hobday, A.J. (2003). Swordfish - Environment - Seamount - Fishery interactions off eastern Australia. Report to the Australian Fisheries Management Authority, Canberra, Australia. Available from Robert.Campbell@csiro.au

Campbell, R. A. (2008). Summary of Catch and Effort Information pertaining to Australian Longline Fishing Operations in the Eastern Tuna and Billfish Fishery. Background paper to ETBF Resource Assessment Group meeting, 29-30 July 2008, Hobart. Available from Robert.Campbell@csiro.au

Cochran W.G. 1977. Sampling techniques, 3rd edition. Wiley & Sons, NY.

Condie, S.A. and Dunn, J.R. (2006). Seasonal characteristics of the surface mixed layer in the Australasian region: implications for primary production regimes and biogeography. Marine and Freshwater Research, 57, 569–590.

Dambacher, J.M., Young, J.W., Olson, R.J., Allain, V., Lansdell, M.J. and Cooper, S.P. (2008). Analysing pelagic food webs leading to top predators in the Pacific Ocean: a graph-theoretic approach. FRDC 2004/063 Final Report.

Dambacher, J. M., H. W. Li and P. A. Rossignol (2003). Qualitative predictions in model ecosystems. Ecological Modelling **161**: 79-93.

Davenport, S. R. and N. J. Bax (2002). A trophic study of a marine ecosystem off southeastern Australia using stable isotopes of carbon and nitrogen. Canadian Journal of Fisheries and Aquatic Sciences **59**(3): 514-530.

Devred, E., Sathyendranath, S. and Platt, T. (2007). Delineation of ecological provinces using ocean colour radiometry. Marine Ecology Progress Series, 346, 1–13.

Etnoyer, P., Canny, D., Mate, B. and Morgan, L. (2004). Persistent pelagic habitats in the Baja California to Bering Sea (B2B) ecoregion. Oceanography, 17, 90–101.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spaulding, M., Steneck, R.S. and Watson, R. (2008). A global map of human impact on marine ecosystems. Science, 319, 948–952.

Heath, M. R. (2005). Changes in the structure and function of the North Sea fish foodweb, 1973-2000, and the impacts of fishing and climate. ICES Journal of Marine Science **62**: 847-868.

Hobday, A.J. and Hartmann, K. (2006). Near real-time spatial management based on habitat predictions for a longline bycatch species. Fisheries Management and Ecology, 13, 365–380.

Hobday, A.J., Hartmann, K., Hartog, J. and Bestley, S. (2006). SDODE: Spatial dynamics ocean data explorer. User Guide. CSIRO Marine and Atmospheric Research, Hobart. 8 pp.

Hobday, A. J., N. Flint, T. Stone and J. S. Gunn (2009). Electronic tagging data supporting flexible spatial management in an Australian longline fishery. Tagging and Tracking of Marine Animals with Electronic Devices II. Reviews: Methods and Technologies in Fish Biology and Fisheries. J. Nielsen, J. R. Sibert, A. J. Hobdayet al. Netherlands, Springer. **9:** 381-403.

Hobson, K. A. and L. I. Wassenaar (1999). Stable isotope ecology: an introduction. Oecologia **120**(3): 312-313.

Jennings, S., K. J. Warr and S. Mackinson (2002). Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. Marine Ecology Progress Series **240**: 11-20.

Link, J. S., J. K. T. Brodziak, S. F. Edwards, W. J. Overholtz, D. Mountain, J. W. Jossi, T. D. Smith and M. J. Fogarty (2002). Marine ecosystem assessment in a fisheries management context. Canadian Journal of Fisheries and Aquatic Sciences **59**(9): 1429-1440.

Longhurst, A., Sathyendranath, S., Platt, T. and Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. J. Plankton Res., 17, 1245–1271.

Lynn, R.L. and Simpson, J.J. (1987). The California Current system: The seasonal variability of its physical characteristics. Journal of Geophysical Research, 92, 12947–66.

Mimmack, G. M., S. J. Mason and J. S. Galpin (2001). Choice of Distance Matrices in Cluster Analysis: Defining Regions. Journal of Climate **14**: 2790-2797.

O'Neill, R. V. (2001). Is it time to bury the ecosystem concept? (with full military honors, of course!) The Robert H. MacArthur Award Lecture. Ecology **82**(12): 3275-3284.

Palacios, D.M., Bograd, S.J., Foley, D.G. and Schwing, F.B. (2006). Oceanographic characteristics of biological hot spots in the North Pacific: A remote sensing perspective. Deep-Sea Research II, 53, 250–269.

Polovina, J.J., Balazs, G.H., Howell, E.A., Parker, D.M., Seki, M.P. and Dutton, P.H. (2004). Forage and migration habitat of loggerhead (*Caretta caretta*) and olive ridley (*Lepidochelys olivacea*) sea turtles in the central North Pacific Ocean. Fisheries Oceanography, 13, 36–51.

Revill, A., J. W. Young and M. J. Lansdell (2009). Stable isotopic evidence for trophic groupings and bio-regionalization of predators and their prey in oceanic waters off eastern Australia. Marine Biology **156**: 1241-1253.

Ridgway, K.R. (2007). Seasonal circulation around Tasmania: An interface between eastern and western boundary dynamics. Journal of Geophysical Research, 112, C10016, doi:10.1029/2006JC003898,.

Ridgway, K.R. and Godfrey, J.S. (1997). Seasonal cycle of the East Australian Current. Journal of Geophysical Research, 102, (C10): 22921–22936.

Roff, J.C., Taylor, M.E. and Laughren, J. (2003). Geophysical approaches to the classification, delineation and monitoring of marine habitats and their communities. Aquatic Conservation: Marine and Freshwater Ecosystems, 13, 77–90.

Roff, J.C. and Evans, S.M.J. (2007). Frameworks for marine conservation - non-hierarchical approaches and distinctive habitats. Aquatic Conservation: Marine and Freshwater Ecosystems, 12, 635–648.

Sydeman, W.J., Brodeur, R.D., Grimes, C.B., Bychkov, A.S. and McKinnell, S. (2006). Editorial-Marine habitat ‘‘hotspots’’ and their use by migratory species and top predators in the North Pacific Ocean: Introduction. Deep-Sea Research II, 53, 247–249.

Williams, A. and Bax, N.J. (2001). Delineating fish-habitat associations for spatially based management: an example from the south-eastern Australian continental shelf. Marine and Freshwater Research, 52, 513–536.

Young, J. W., M. J. Lansdell, A. J. Hobday, J. M. Dambacher, S. P. Griffiths, S. Cooper, R. Kloser, P. D. Nichols and A. Revill (2009). Determining ecological effects of longline fishing in the Eastern Tuna and Billfish Fishery. FRDC Final Report 2004/063. 320 pp.

Zainuddin, M., Kiyofujia, H., Saitohb, K. and Saitoh, S.-I. (2006). Using multi-sensor satellite remote sensing and catch data to detect ocean hot spots for albacore (*Thunnus alalunga*) in the northwestern North Pacific. Deep-Sea Research II, 53, 419–431.

Table 1: Ocean data sets used to generate the habitat clusters. The product number represents the data set as designated in SDODE (Hobday et al., 2006). The data sets used for the final clustering are indicated in the last column.

Data code	Data set, preceded by SDODE product number	Used in final clustering
1	10: AGSO Bathymetry	Yes
2	112: CSIRO SST: 10 Day Composite Data	Yes
3	210: Australia Altimetry (sea surface height)	
4	400: SeaWiFS k490 Turbidity	
5	410: SeaWiFS Chlorophyll *a*	Yes
6	115: synTS CSIRO Temperature at Depth (250 m)	Yes
7	425: CARS Salinity Climatology	
8	430: CARS Oxygen Climatology	
9	435: CARS Nitrate Climatology	Yes
10	440: CARS Silicate Climatology	
11	445: CARS Phosphate Climatology	

Table 2: Summary of cluster names and the classification algorithm used to allocate clusters. Mean longitude and latitude were used in a set of decision rules to sequentially allocate the clusters to a consistent "name" for the habitat, in the order indicated in the table (cluster 1 through to 7).

Cluster name	Centre of cluster
1. Coral Sea	mean_lon <155 and mean_lat >-27
2. West Pacific	mean_lon ≥155 and mean_lat >-27
3. Tasman Sea	mean_lon ≥155 and mean_lat ≤-34
4. Eddy offshore	mean_lon ≥153 and mean_lon <155 and mean_lat ≤-34 and mean_lat >-36
5. Eddy inshore	mean_lon <153 and mean_lat ≤-32 and mean_lat >-36
6. Southern ocean	mean_lat ≤-36
7. Central/EAC	Remainder of centres

Table 3: Mean location of the seven habitats identified for the study region for the period 1995–2006. The number of months (maximum 144) that each habitat was identified is indicated, together with the mean centre and the mean eastern, western, northern, and southern extent of each habitat. The centre of each habitat is illustrated in **Figure 5**. Mean size of each habitat is in square degrees

Cluster	Months (n)	Mean size (deg)	Centre of habitat (°E, °S)		Mean eastern boundary (°E)	Mean western boundary (°E)	Mean northern boundary (°S)	Mean southern boundary (°S)
1. Coral Sea	106	30.14	153.26	22.84	154.82	151.65	20.35	27.79
2. West Pacific	143	200.08	156.92	23.26	159.99	153.40	20.04	27.89
3. Tasman Sea	143	146.89	156.30	35.94	159.70	151.69	33.14	37.97
4. Eddy offshore	14	48.07	154.14	34.66	157.98	152.11	33.01	36.20
5. Eddy inshore	48	12.10	151.33	34.89	154.64	150.37	31.62	37.48
6. Southern ocean	66	34.79	153.51	37.34	156.52	151.28	36.28	37.98
7. Central/EAC	133	196.51	156.61	30.29	159.83	153.27	25.79	34.27

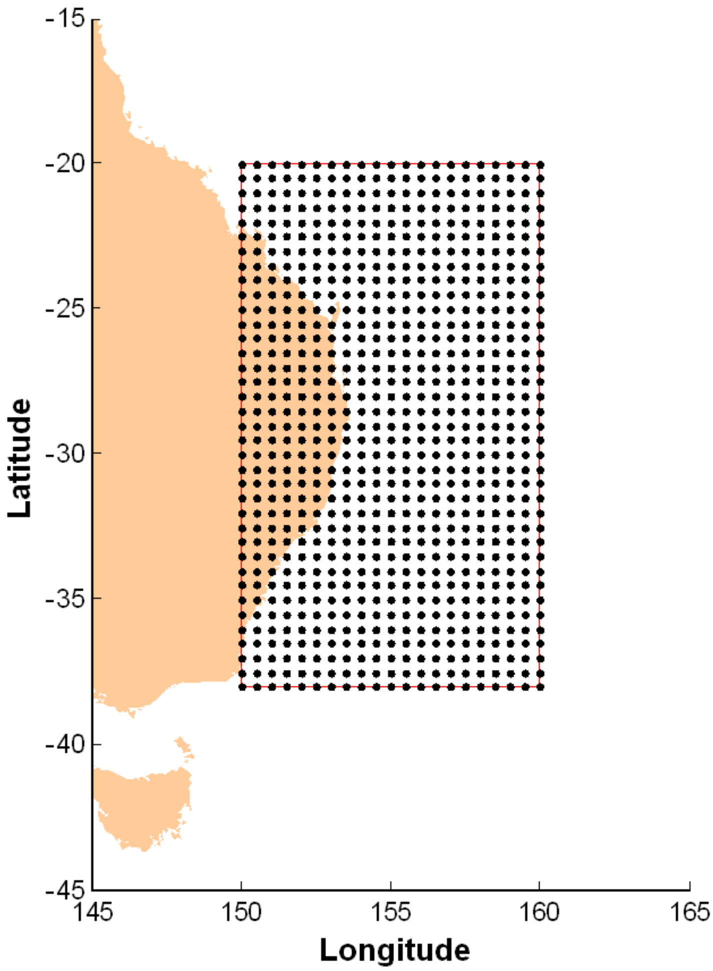


Figure 1: The study area was divided into 0.5 degree squares (grid cells), and a range of environmental data averaged for each grid cell. The centre of each grid cell is represented by a dot in this figure. Cells with an average water depth of less then 100 m were excluded during data extraction.

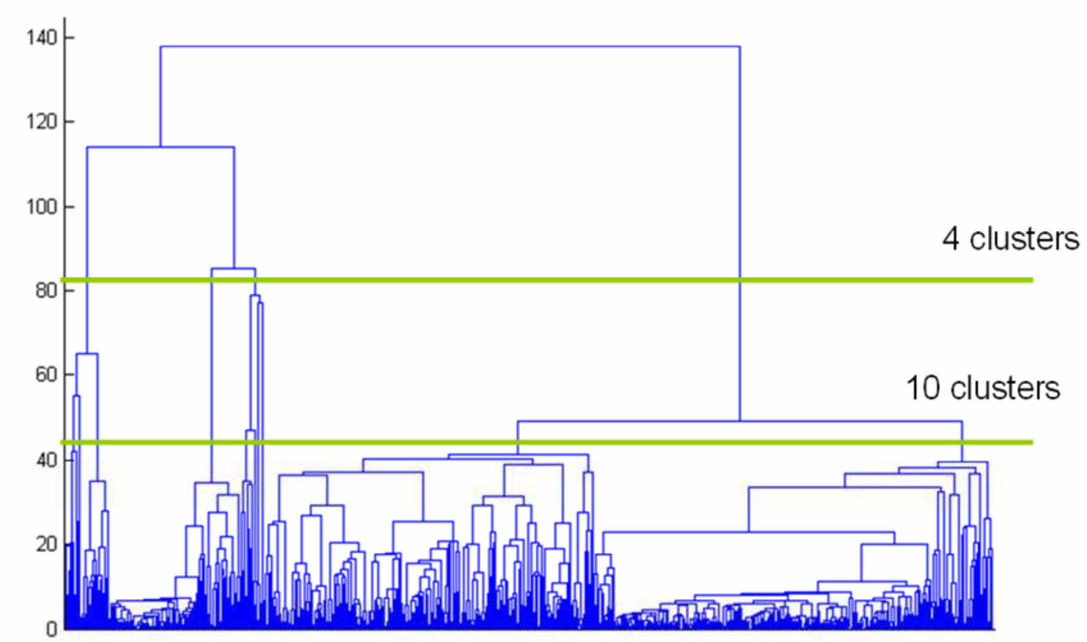


Figure 2: Example of the clustering output applied to the 0.5 degree boxes (arranged on the x-axis) based on five environmental variables averaged for each box at a monthly scale. Boxes with similar environmental characteristics were clustered together. The y-axis is a measure of difference – branches high on the y-axis indicate clusters that are distinct, while close to zero, lines connect boxes that are very similar. The two thick horizontal green lines indicate the groups that would arise if four or ten clusters were selected.

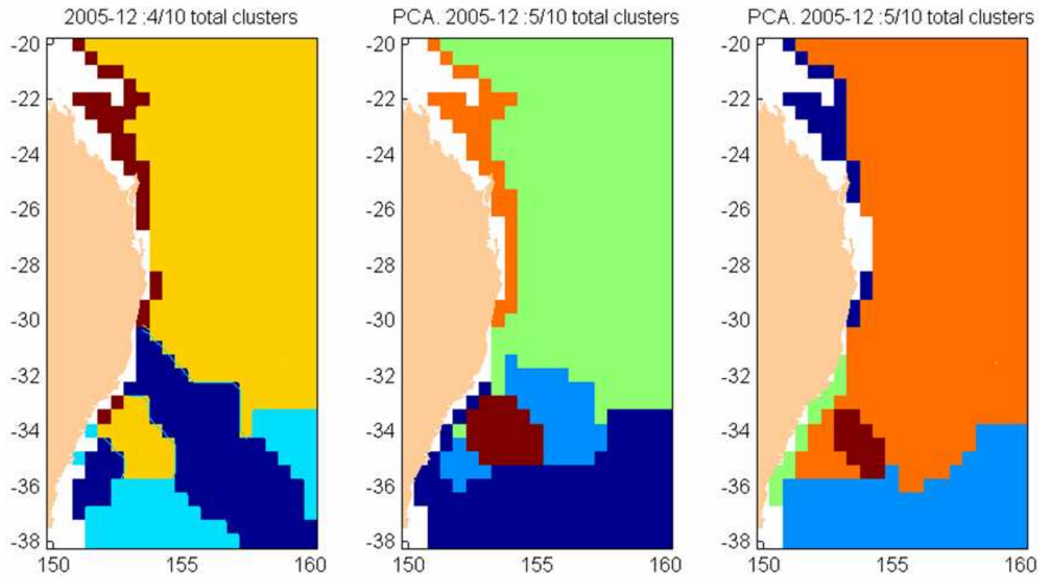


Figure 3: Example of the clustering output applied to the 0.5 degree boxes based on five environmental variables averaged for each box at a monthly scale for December 2005. The left panel represents the clustering algorithm applied used in analyses presented here, with the centre and right illustrations representing a PCA-based cluster analysis using just five (as in the left panel) and all 11 variables (right panel). Colours, automatically allocated, do not represent the same clusters in each figure.

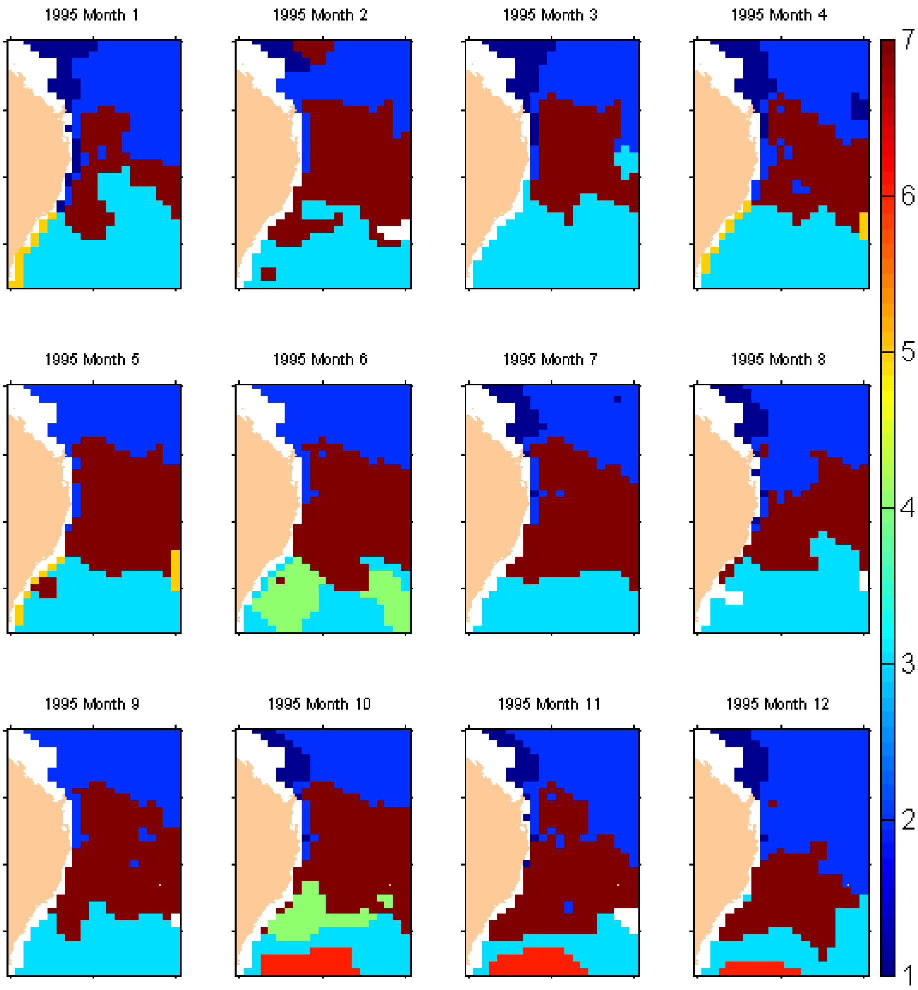


Figure 4. Example of the distribution of the dynamic ocean habitats over the year 1995. The integer values of the colorbar represent each of the seven habitats. Note that some of the habitats are relatively ephemeral, while the majority are persistent and coherent across months.

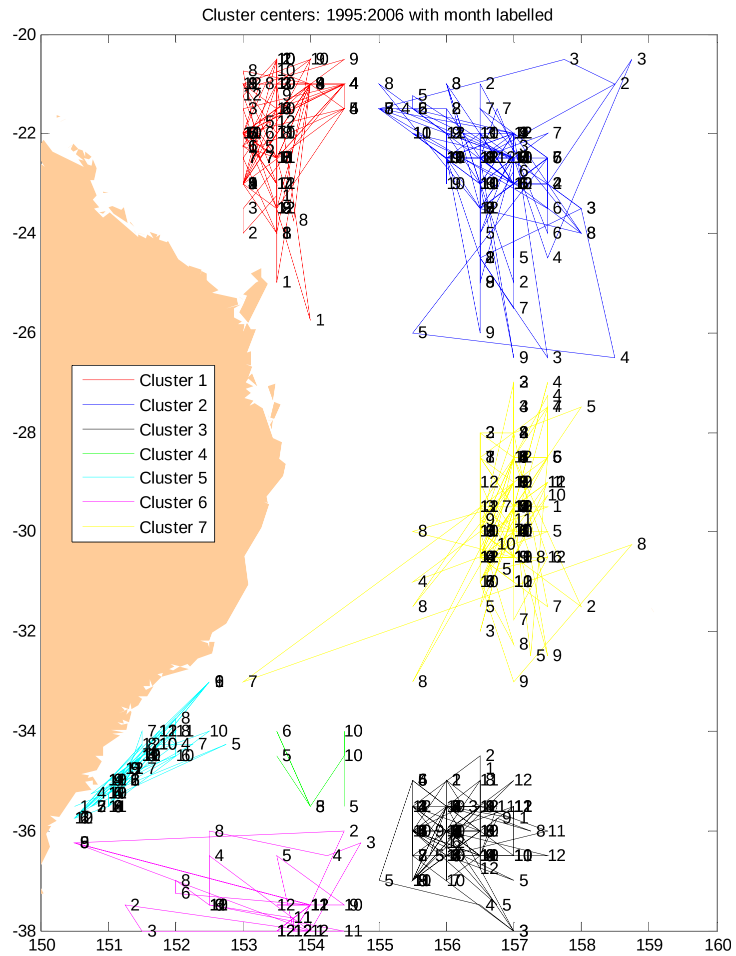


Figure 5: Distribution of the centres of the seven habitats identified for the study region for each month in the period 1995–2006. Cluster names are provided in **Table 2**. The month is indicated by the numerals (1–12), and sequential months are connected with straight lines. These colours for each cluster are consistent with colours in subsequent figures.

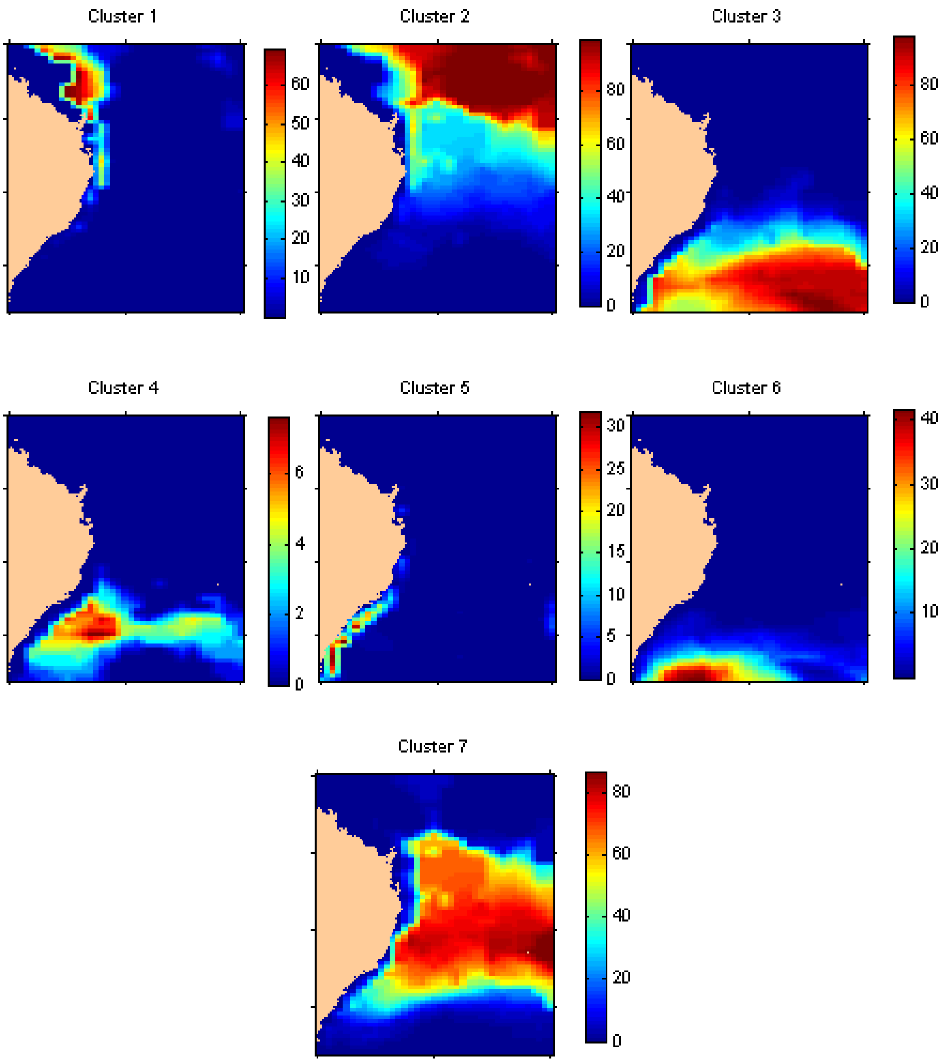


Figure 6. Probability maps for the seven pelagic habitats identified in eastern Australia for the years 1995-2006. Each map shows the percentage of the 144 month-year combinations that the habitat was present at each location. Note the different color scale for each habitat.

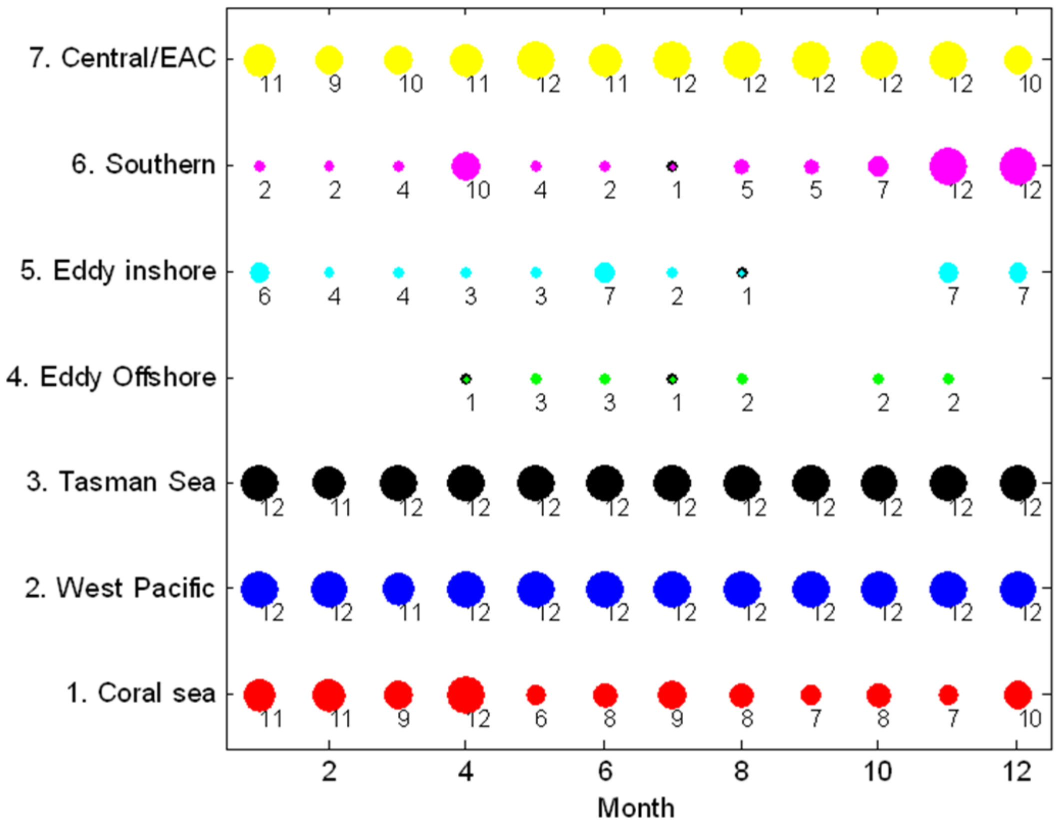


Figure 7: Summary of presence of pelagic habitats for the period 1995–2006 (12 years) for each month in the study region. The size of the symbols and the number above each symbol represents the number of months in this period that the habitat was present in the study region.

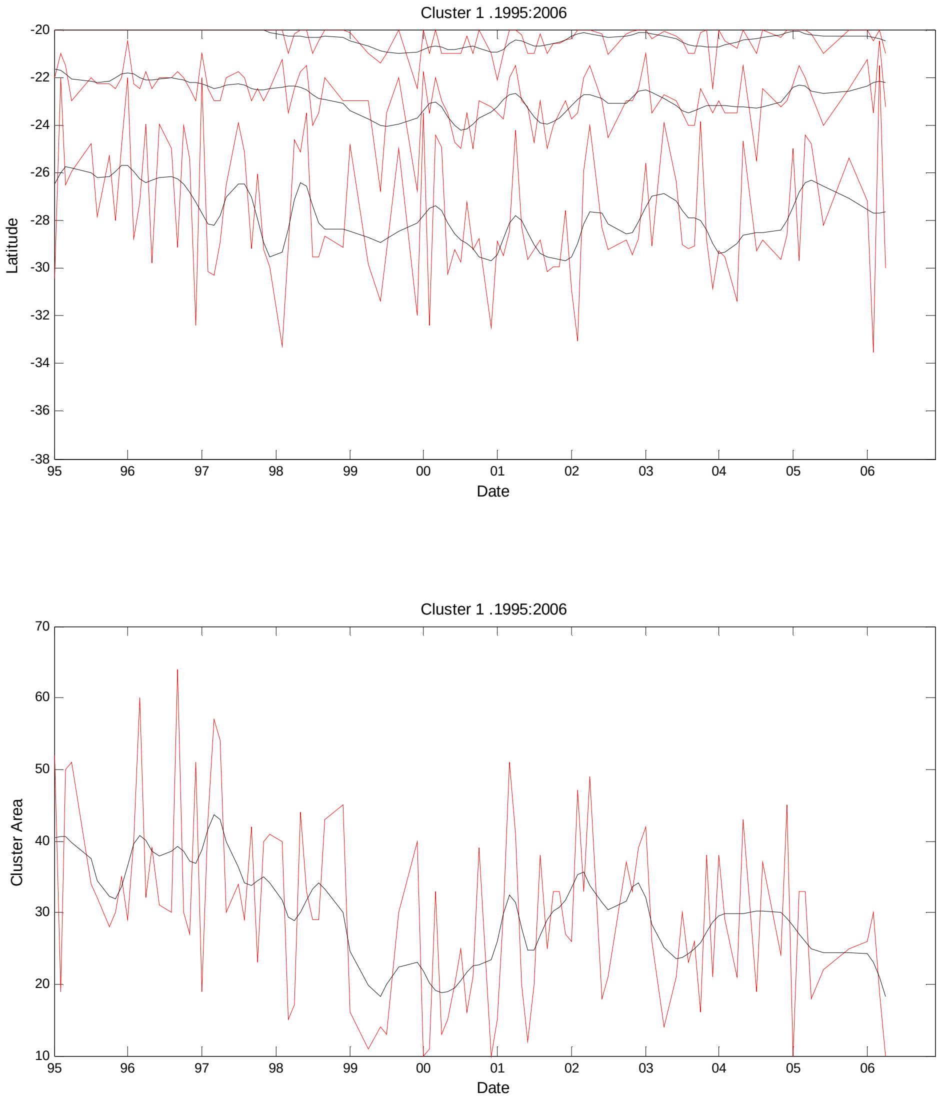


Figure 8. Variation in location of the northern (upper line) and southern boundaries (lower line), and the average centre (middle line) of Cluster 1 (Coral Sea) over the period 1995–2006 (upper panel) and the area of the habitat over time (lower panel). The solid lines in each panel represent a five month smooth to the data. Cluster area is measured in square degrees.

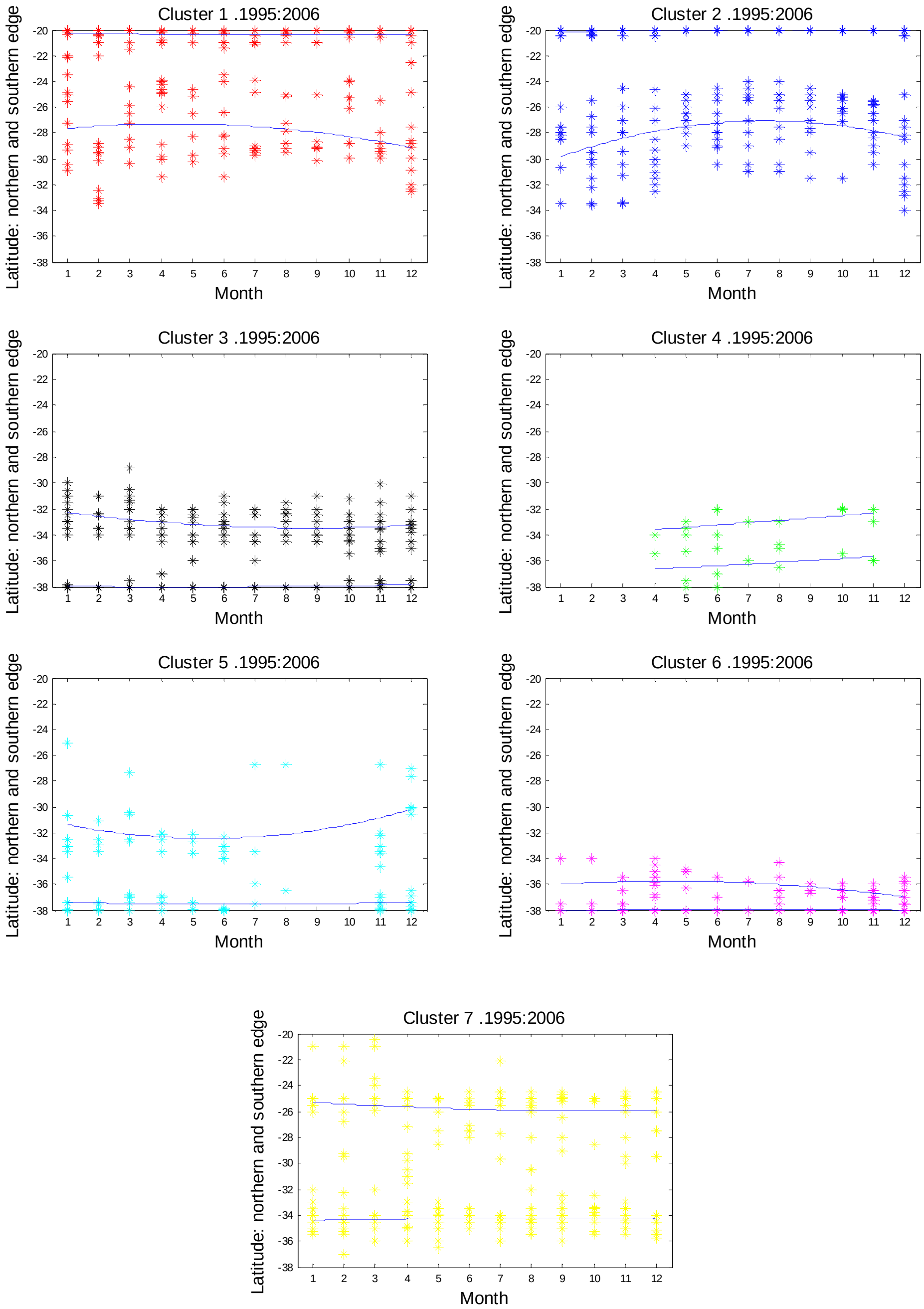
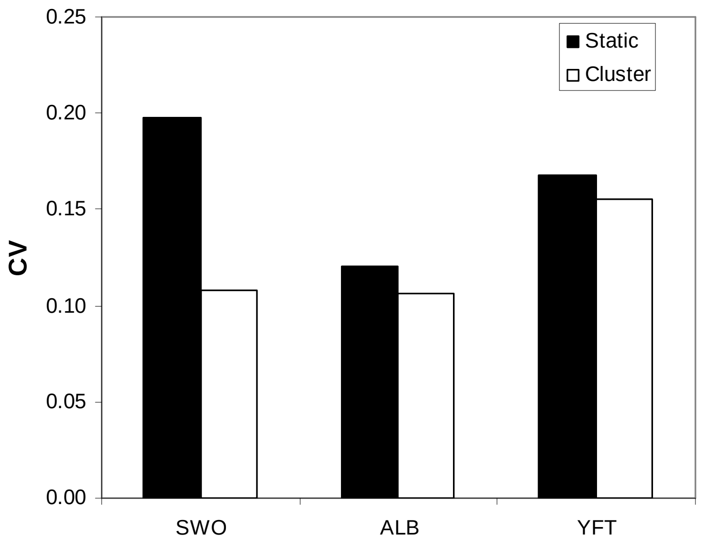


Figure 9. Seasonal cycle in the northern (upper line) and southern boundaries (lower line) for each habitat (cluster) over the period 1995–2006. Colours are the same as used in **Figure 5.** The solid lines in each panel represent second order polynomial fit to the data. Cluster names are provided in Table 2.

A.



B.

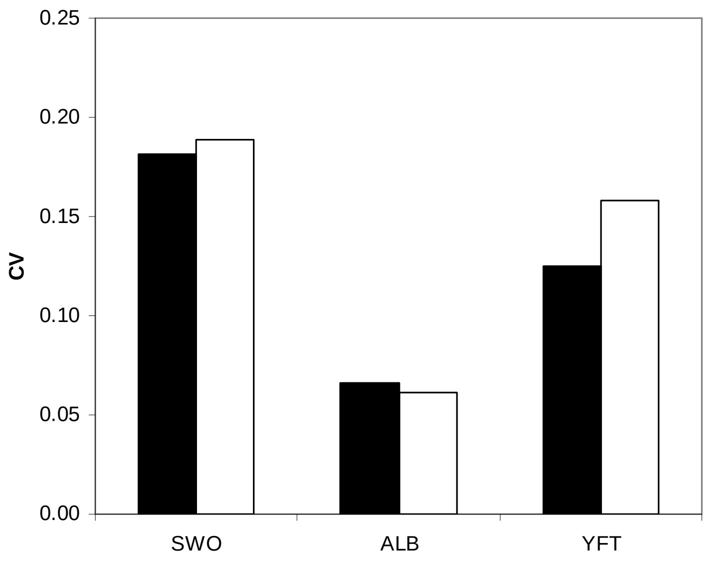


Figure 10. Precision (CV) of estimated mean isotope values for three fish species captured in the ETBF calculated using static and dynamic habitats. A. 15N B. 13C. Species codes represent SWO, swordfish; ALB, albacore; YFT, yellowfin tuna.

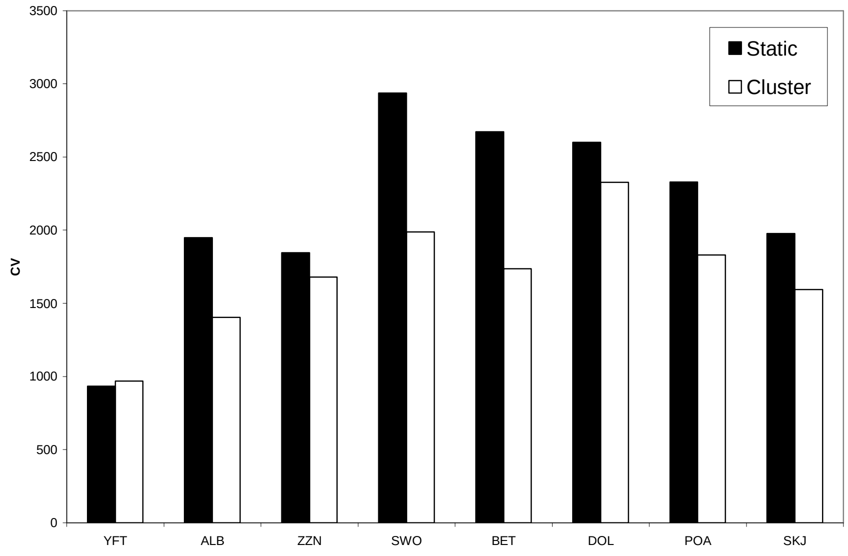


Figure 11. Precision (CV) of estimated abundance for the top eight species captured in the ETBF fishery for the period 2001-2006 using the static and time variable cluster allocation. Species codes: YFT, yellowfin tuna; ALB, albacore; ZZN, Rudderfish and Black Oilfish; SWO, swordfish; BET, bigeye tuna; DOL, dolphinfish; POA, Ray's Bream; SKJ, skipjack tuna.

Appendix 6

Notes on AFMA logbook data issues relating to the ETBF

Robert Campbell
CSIRO Marine and Atmospheric Research

September 2007

1. Introduction

A copy of the AFMA logbook data is stored within the Pelagic Database managed by CSIRO in Hobart. These data are used for undertaken research related to the Australian tuna and billfish fisheries. Over the years, a number of problems relating to these data have been identifed and these notes itemise some of the data issues encountered in compiling data related to fishing operations in the ETBF.

Step 1. Initial construction of ETBF data table.

Initially, all data relating to fishing operations in the ETBF (i.e., having Start_Set_Longitude ≥ 141 and having a logbook of the type AL0X or OT0X) is extracted from the table AFZ_FOP_AUS and stored in a new table ETBF_FOP. The last update (undertaken in June 2007) gave 151,340 records. Each record is indexed by a unique FISHING_OPERATION_NUMBER (FOP).

The number of records by logbook type is as follows:

LOGBOOK	NRECS
AL01	142
AL02	27250
AL03	8076
AL04	37206
AL05	71590
OT01	4116
OT02	2338
OT03	622
Sum	151340

The number of records by Activity Code is as follows:

CODE	ACTIVITY	NRECS
274	FSH	151289
275	WTH	17
276	PORT	20
277	BRK	3
278	STM	4
279	OTH	1
285	CONT	1
286	ELSE	5
Sum		151340

All records relating to non-fishing activities are deleted (i.e.CODE != 274). The number of records by Fishing Method is then as follows:

CODE	METHOD	NRECS
0		2461
236	LL	143645
237	PL	27
238	RR	961
239	TR	2834
240	HL	1152
300	XX	209
Sum		151289

ISSUE #1: Blank Method-Codes in Effort Table

A check of the latitude and longitude coordinates indicates that all records fall within the following region:

$$-49 \leq \text{Latitude} \leq -10$$
$$141 \leq \text{Longitude} \leq 172$$

Note: for 1 record the latitude is blank while the two records having a latitude<146 are doubtful.

Step 2 – Ascertain Fishing Method for each FOP

In order to ascertain the methods for those records in ETBF_FOP for which the fishing method is blank, we use the information in the table AFZ_CATCH_AGGREGATE_AUS. This table the catch (and in some instances the Method) for all FOPS in the AFZ_FOP_AUS table.

Initially, a listing of the number of associated catch data records, stratified by FISHING_METHOD_CODE, is extracted for all FOPS contained in EBF_FOP and stored in the table ETBF_CATCH. The number of records by method is as follows:

CODE	VALUE	NRECS	
0		39225	
93	PL	3	set CODE = 237
94	LL	36633	set CODE = 236
95	TR	104	set CODE = 239
96	RR	28	set CODE = 238
128	X	3	set CODE = 300
134	HL	10	set CODE = 240
236	LL	70697	
237	PL	26	
238	RR	77	
239	TR	255	
240	HL	396	
287	GN	1313	
300	XX	71	
301	DL	76	
Sum		148917	

ISSUE #2: Blank Method-Code in Catch Table (though as will be shown later, the Method was stored in the Catch Table only with AL04 and AL05 logbooks)

ISSUE #3: Multiple Codes for same Method in Catch Table (e.g., LL=94 and 236)

It was also noted that several FOPS have catches associated with more than one Fishing Method, with the number of different Methods recorded per FOP is as follows:

N_Methods	NFOPS
1	145776
2	1566
3	3
Sum	147345

ISSUE #4: Multiple catches associated with different fishing methods for some FOPS

For each FOP in the table ETBF_CATCH the corresponding method in the table ETBF_FOP is added. The number of records for each set of corresponding Methods is as follows:

METH_CAT	METH_FOP	NFOPS	
0	0	1965	
0	236	33689	
0	238	714	
0	239	2063	
0	240	592	
0	300	202	***
236 (LL)	236	107330	
237 (PL)	236	3	
237 (PL)	237	26	
238 (RR)	236	28	
238 (RR)	238	77	
239 (TR)	236	104	
239 (TR)	239	255	
240 (HL)	236	20	
240 (HL)	240	384	
240 (HL)	300	2	***
287 (GN)	0	8	
287 (GN)	236	1277	
287 (GN)	239	25	
287 (GN)	300	3	
300 (XX)	0	0	
300 (XX)	236	68	
300 (XX)	300	6	
301 (DL)	0	1	
301 (DL)	236	69	
301 (DL)	300	6	
Sum		148917	

ISSUE #5: Different Methods associated with some FOPS in FOP and CATCH tables.

Methods 287 (Gillnet), 300 (Other), and 301 (Drop-line) are not usually associated with the catch of tuna and billfish and so these records were deleted from the ETBF_CATCH and ETBF_FOP tables. (Note: a check of the species associated with these catches indicates no tuna or billfish – appears to consider mainly of a mix of demersal and shark species). Why these methods are associated with the AL0X logbooks remains unclear.

ISSUE #6: Non-tuna and billfish catch methods associated with some ETBF logbooks.

For some FOPs in the ETBF_CATCH table there are still multiple catch method records (NMETHS). The number of METH_CAT records for each FOP is added to each record in the ETBF_FOP table. The number of FOPs in ETBF_FOP grouped by NMETHS is:

NMETHS	NFOPs
1	146945
2	148
3	3
0	3841
Sum	150937

ISSUE #7: 3841 FOPs with no matching catch records –assume catch=0.

For those records where NRECS=1 the number of records for each set of corresponding Methods in the CATCH and FOP tables is as follows:

METH_CAT	METH_FOP	NFOPS
0	0	1965
0	236	33677
0	238	714
0	239	2063
0	240	592
236	236	107189
237	237	26
238	238	77
239	236	3
239	239	255
240	240	384
Sum		146945

So, for these records METH_CAT=METH_FOP and for some records where METH_CAT=0 it is possible to ascertain the method from METH_FOP. However, 1965 FOPs still have an unknown method.

STEP 3 – Separate LL and OT records

Create table ETBF_OT and insert all non-longline records from ETBF_LL (i.e., where METH_FOP>236 and logbook like 'OT%'). Delete corresponding records from ETBF_FOP and rename table to ETBF_LL.

The number of FOPS grouped by Logbook Type and Method for ETBF_OT is:

LOGBOOK	METHOD	NFOPS	
AL02	240 HL	10	
AL04	237 PL		3
AL04	238 RR	28	
AL04	239 TR	101	
AL04	240 HL	10	
AL05	237 PL		2
AL05	238 RR	48	
AL05	239 TR	95	
AL05	240 HL	310	
OT01	0	10	
OT01	238 RR	828	
OT01	239 TR	2456	
OT01	240 HL	613	
OT02	0	2338	
OT03	237 PL		25
OT03	238 RR	85	
OT03	239 TR	283	
OT03	240 HL	229	
sum		7474	

The number of FOPS grouped by Logbook Type and Method for ETBF_LL is:

LOGBOOK	METHOD	NOPS	
AL02	0	113	
AL02	236 LL		27122
AL03	236 LL		8040
AL04	236 LL		37205
AL05	236 LL		71135
sum		143615	

ISSUE #8: FOPs with no METHOD but AL0X logbook are assumed to be LL records.

STEP 4 – Add Vessel

Create table ETBF_VSL listing unique VSL_IDs in ETBF_LL – 677 records created – and add VSL_ID_AFMA, VSL_NAME and VSL_CALLSIGN from table VESSELS.

Number distinct VSL_ID	536	(0 null)
Number distinct VSL_ID_AFMA	496	(39 null)
Number distinct VSL_NAME	388	(0 null)
Number distinct VSL_CALLSIGN	514	(1 null)

ISSUE #9: Not a 1-1 relation between various VSL identifiers.

ISSUE #10: Some incorrect spellings of vessel identifiers such as VSL_NAME, eg.

VSL_ID	ID_AFMA	VSL_NAME	VSL_CALL
2804	11681	NOUVO GUISEPPE	LFB11950
14364		NUOVO GUISEPPE	LFB11950

Due to the above issues there are a number of problems in attempting to identify unique vessels. However, information in the table VESSELS and VESSEL_ATTRIBUTES (which is based on data sent to me directly by AFMA) are often different. Furthermore, some vessels having the same name but different identifiers (e.g., VSL_IDs, CALL_SIGNS) are found to be fishing during the same period.

A first pass through the data correcting some possible errors reduces the number of distinct VSL_NAMES in the ETBF_LL data to 378 (down from the 388 initially identified). These VSL_NAMES are added to the ETBF_LL table using the VSL_ID identifier. (Note, a listing of vessel identifiers is provided in Appendix A.)

STEP 5 – Check for duplicate FOPS

Each FOP should be uniquely identified by the combination of VSL_NAME, FOP_DATE and OP_NUM. Here we check this assumption. For each FOP record a new FOP_NUM is calculated as follows:

FOP_NUM = VSL_NUM*100000+JULIAN_DAY*10+OP_NUM

where

JULIAN_DAY = TO_NUMBER(TO_CHAR(FOP_DATE,'J')) – 2446125

and 2446125 = MIN(JULIAN_DAY)-1

Upon checking, 313 duplicate FOP_NUMs were found (626 records in total). For each duplicate set of records the START_TIME and LOCATION of the set were also compared as an additional check. This reduced the number of duplicate records to 86 (NB: the START_TIME was null for all these records). Finally the catch of yellowfin and bigeye tuna for each of these duplicate records was also compared and found to be the same. (NB: all duplicates were related to 2 vessels – Sensation (42) and Southern Gull II (44)). The conclusion reached was that these records were duplicates and each duplicate copy was deleted from the EBTF_LL table.

ISSUE #11: Duplicate records in the data.

For the above analysis, records where OP_NUM was not recorded were excluded. In order to incorporate these records, the analysis was repeated using START_TIME instead of OP_NUM. Again, a FOP_NUM was calculated for each record and duplicates found – 46 records were found to have duplicate FOP_NUM, LOCATION and OP_NUM (always zero). However, only 2 of these records were found to also have the same effort and catch information suggesting only a single duplicate record. This was deleted as before. The reasons for the duplication of the other FOP_NUMs remains uncertain but could be related to incorrect fields such as FOP_DATE.

(NB: there is only one record in the ETBF_LL table were both OP_NUM and START_TIME are null).

STEP 6 – Add Yellowfin Tuna Catch

The catch information relating to each FOP is stored in the AFZ_CATCH_AGGREGATE_AUS table. The catch (number of fish and weights – dressed and/or whole) is aggregated by species (SPC_ID) and grade

(SPC_GRADE). The steps to link the catch with each FOP is illustrated here using the yellowfin catch.

The yellowfin tuna catch (SPC_ID=38) was extracted from the catch table and stored in the following TEMP table:

FOP_ID	NUMBER(9)	
METHOD	NUMBER(3)	Fishing Method
GRADE	VARCHAR2(10)	Grade of fish
CNT	NUMBER(6)	Number of fish retained
DIS	NUMBER(6)	Number of fish discarded
WWT_CODE	NUMBER(3)	Whole weight code
WWT	NUMBER(6)	Whole weight
DWT_CODE	NUMBER(3)	Dressed weight code
DWT	NUMBER(6)	Dressed weight
TARGET	VARCHAR2(1)	Was it a target species? (Y/N)

The extraction is done is two steps: first, those FOPS in ETBF_LL for which there is only a single Method in the Catch table (as identified in Step 2), and second, those FOPS in ETBF_LL for which there are multiple Methods in the Catch table (for which we only select the longline catch).

Nine different Grades are identified in the yellowfin catch (with up to four grades per FOP). A summary of the number of records for each Grade is given below, where the Unique-FOPS column denotes the number of unique FOP_IDs for the records identified with each Grade:

Log	*Method*	*Grade*	*#FOPs*	*Non-Zero Records*				*Unique*
				CNT	*DIS*	*WWT*	*DWT*	*FOPS*
AL02	***Null***	***Domestic***	***14997***	***0***	***0***	***1287***	***13710***	***All***
	Null	***Export***	***10705***	***0***	***0***	***950***	***9755***	***All***
	Null	***Discard***	***1676***	***0***	***1676***	***0***	***0***	***All***
	Null	***Total***	***21304***	***21304***	***0***	***0***	***0***	***All***
AL03	***Null***	***Total-L***	***6387***	***6346***	***0***	***0***	***6361***	***All***
	Null	***Total-S***	***636***	***632***	***0***	***0***	***599***	***All***
	Null	***Discard-L***	***447***	***0***	***447***	***0***	***0***	***All***
	Null	***Discard-S***	***204***	***0***	***204***	***0***	***0***	***All***
AL04	***94***	***Null***	***28699***	***28250***	***1207***	***28296***	***0***	***27320***
AL05	***237***	***Null***	***59760***	***59157***	***6663***	***1003***	***58188***	***59035***
			144818					

It is obvious from the above table that different types of data have been collected with each successive logbook type. For example, with the AL02 logbook the total number of fish caught (and retained?) and the number of fish discarded were recorded together with the DWT or WWT weight of fish retained graded by Domestic or Export. On the AL03 logbook the number of fish retained and discarded together with DWT only were recorded, both graded by Small and Large. Finally, no grading has been used on both the AL04 and AL05 logbooks.

As the above table indicates, for both the AL04 and AL05 logbooks there are duplicate catches for the same FOP_ID. For the 1379 FOPs for the AL04 logbook which have duplicate catch records, each record is different and only contains information on the CNT, DIS and WWT. On the other hand, for the 724 FOPs for the

AL05 logbook which have duplicate catch records, each record is also different but for most FOPs (676) one record contains information on CNT, DIS, WWT_CODE and WWT and the other record contains information on CNT, DIS and DWT. There are 41 FOPs for which the two records both contain information on CNT, DIS and DWT only, while several others (7) have other combinations. The reasons for these duplicates remains unclear and as it is difficult to discern which are valid (one or both) then for the current purpose the catches for each FOP_ID are combined.

ISSUE #12: Duplicate CATCH records associated with some FOP_IDs.

Despite the fact that the logbooks request that the number of fish retained and the number of fish discarded be recorded separately, there are records in the catch table where CNT=DIS and total weight=0. In these situations it is assumed that all fish were discarded and so CNT was set equal to zero in these situations (28 FOPS).

ISSUE #13: Uncertain number of fish retained and discarded for some FOPs.

A summary by logbook type of the targeting information is as follows:

LOGBOOK	T	NRECS
AL02	N	48682
AL03 ,	N	5378
AL03	Y	2296
AL04	N	12458
AL04	Y	16241
AL05	T	39918
AL05		19842
		144815

This indicates that information on which species were targeted has only been recorded since the AL03 logbook was introduced. For the AL05 logbook it is assumed that where the Target field is blank that this denoted no targeting.

A new table ETBF_LL_YFT is created and the catch in the table TEMP is summed across each FOP_ID:

FOP_ID	NUMBER(9)
FREQ	NUMBER(1)
CNT	NUMBER(6)
DIS	NUMBER(6)
WWT	NUMBER(6)
DWT	NUMBER(6)
TARGET	NUMBER(1)

The extra field TOT_WT=DWT/0.865 + WWT is added to store the Total Whole Weight retained for each FOP. We assume that for yellowfin tuna the DWT-WWT ratio averages 86.5%. Upon inspection it is found that there are 1066 records where either CNT=0 or TOT_WT=0. (NB. Both are zero for 732 of these records indicating a zero retained catch – DIS is not zero for 649 of these records though DIS=0 for 83 records indicating a zero retained and zero discard catch)

ISSUE #14: For some FOPS only catch number is recorded whilst for other FOPS only catch weight is recorded.

ISSUE #15: Some catch records have a zero total catch.

In order to estimate the missing catch information, we calculate the mean weight of a yellowfin caught each year and use this value to convert numbers-to-weight and vice-versa for those FOPS where either is missing. Once done, the catch information is added to each corresponding FOP in the table ETBF_LL.

Step 7 – Add Trip Identifier

For many types of analyses it is important to identify the catch and effort to the trip level. Unfortunately the trip associated with each FOP is not identified in the FOP and CATCH tables. However, since the advent of the AL05 logbook, a separate TRIP table has been collated by AFMA. Within the CSIRO database this table has the name AFZ_TRIPS and records the departure date (and port) and arrival date (and port) for each vessel trip. The aim in this data step is to link each FOP in the table ETBF_LL with the corresponding trip (identified by a unique TRIP_ID) in the TRIP table.

The manner by which this can be achieved is as follows:

1) Listed with each FOP is the BOOK_NO and PAGE_NO on which the data was originally recorded by the skipper.
2) Also, each TRIP has associated with it the BOOK_NO and PAGE_NO corresponding to the first FOP and the BOOK_NO and PAGE_NO corresponding the last FOP.
3) Using the BOOK and PAGE number information it should be possible to associate each FOP with a unique TRIP.

A review of the data in the TRIP table indicates that eleven trips in the ECT fishery have a length (defined as the number of days between the departure and arrival dates) greater than 25 days, with three trips having a length of greater than one year (365 days). These trips appear to be greater than what is assumed in the fishery, with the latter obviously due to an error in one of the trip dates. Comparing the dates of FOPS with trip dates it is possible to correct some of these errors – i.e., incorrect month or year. – and this was done for the five trips having a trip length greater than 30 days (i.e., one month).

ISSUE #16: Incorrect trip dates for a few records in TRIP table.

The range of the BOOK_NOs and PAGE_Nos in the TRIP table were checked:

- START_BOOK: 1 to 999
- START_PAGE: 1 to 100
- END_BOOK: 0 to 999
- END_PAGE: 0 to 100

Two trips had an unidentified END_BOOK and one trip had an unidentified END_PAGE. Again, by comparing these trips with the information contained in the FOP table it was able to identify the correct values and these were fields were updated in the TRIP table.

ISSUE #17: Incomplete logbook information for a few records in TRIP table.

A new table ETBF_TRIPS is created with the following information:

TRP_ID	NUMBER(9)
VSL_ID	NUMBER(9)
DEPART	DATE
D_TIME	NUMBER(9)
D_PORT	NUMBER(4)
INIT_LOG_NO	NUMBER(9)
ARRIVE	DATE
A_TIME	NUMBER(9)
A_PORT	NUMBER(4)
LAST_LOG_NO	NUMBER(9)

where

INIT_LOG_NO = FIRST_BOOK_NO*1000 + FIRST_PAGE_NO
LAST_LOG_NO = LAST_BOOK_NO*1000 + LAST_PAGE_NO

A total of 19,565 records were created.

7a. Add VSL_NUM to ETBF_TRIPS table

Added in order to identify each vessel. It was noticed that 3 vessels and trips in the TRIP table could not be identified with a vessel in the ETBF_LL_VSL table. Upon inspection of these vessels it was found that they were associated with fishing operations west of 141E (i.e., in the WTBF) and as such they were eliminated from the TRIPS table.

ISSUE #18: Three WTBF trips incorrectly labelled as ECT in the TRIP table.

7b. Identify Trips With Same Departure Or Arrival Dates For Same Vessel

Investigation of the data in table ETBF_TRIPS indicates there are 74 (VSL_ID, DEPART) duplicates and 3 (VSL_ID, ARRIVE) duplicates. In all but two cases one of the trips has the same Depart and Arrive dates indicating that a vessel either did a short day trip then departed again on the same day or arrived on a given day then completed a short day time before arriving again on the same day . A listing of the last three duplicates is as follows:

Vsl_Id	Depart	D_Time	Init_Log	Arrive	A_Time	Last_Log
515	16/Mar/04	1	337076	18/Mar/04	1	337077
515	18/Mar/04	1200	337077	18/Mar/04	2359	337077
2804	10/Feb/01	1200	106037	10/Feb/01	1200	106037
2804	09/Feb/01	1200	106037	10/Feb/01	1200	106037
15307	06/Feb/05	0	571044	06/Feb/05	2359	571044
15307	04/Feb/05	1	571043	06/Feb/05	2359	571043

Whether or not such short trips are possible leading to either two arrivals or two departures on the same day remains unclear. For the first set of duplicate trips it is seen that the first trip recorded the FOPs on pages 76 and 77 whilst the second trip the FOP was only recorded on page 77. As each page has room to record two FOPs it is possible that the second trip recorded the data in the second column as the last FOP for the previous trip (despite the fact that the logbook instructions inform the skipper to begin a new page for each trip). For the second set of trips only the same single page number is used for each trip and is possible if the first column on that page is used for the first trip and the second column is used for the second trip. Finally, for the third set of trips different pages are used on each trip.

ISSUE #19: Unclear as to whether multiple departures or arrivals on the same date are possible or whether such short trips should be ignored.

In order to help clarify the above situation, the number of FOPS falling between the departure and arrival dates and the number of FOPS on both the departure and arrival date for each trip were calculated. Of the 77 trips with DEPART=ARRIVE 52 where found to have no FOPS on that date. As no fishing took place during these trips they were eliminated from the ETBF_TRIPS table. (Note: it is possible that vessels had to return to port on the day of departure during to some problem.) For each of these trips it was noted that either D_PORT or A_PORT was null while both these fields were completed for the corresponding duplicate trip. This situation was also noted for a further 21 trips and these were also eliminated. It was also found that two trips where DEPART=ARRIVE share a PAGE_NO with their respective duplicate trip but there is only a single FOP associated with this PAGE_NO. It is assumed that this FOP is associated with the trip where DEPART!=ARRIVE and hence these two trips are eliminated as before. Finally, the last two trips with a duplicate date were investigated and in each case it was possible to eliminate one trip. Hence, after completing this process no trips were found to have a duplicate DEPART or ARRIVE date. The number of trips remaining was 19,489.

7c. Identify Trips With Same Init_Log_No or Last_Log_No

This procedure checks different trips with FOPS recorded on the same page of the Logbook. Duplicates for either the initial page or last page are checked separately. On the first check 28 pages were found which are shared across trips. Upon checking 12 of these are found to be associated with single day trips for which the depart and arrival dates fell within the dates of another trip. Also, unlike the other trips, the arrival and departure times for these single day trips were always null. These trips were eliminated.

ISSUE #20: Single day trips with duplicate logbook pages and depart and arrival dates which fall within the dates of alternative trips found.

Inspection of the FOP dates and trip dates for the remaining 16 duplicates indicated that 7 were okay (i.e., the last FOP for one trip and the first FOP for the next trip recorded on the same logbook page) while for the other 9 duplicates one of the logbook pages was found to be in error and altered.

ISSUE #21: Log book Number and/or Page Number found to be in error for a few trips.

7d. Identify Trips with Init_Book_No different from Last_Book_No

There are likely to be trips with the initial logbook is completely filled during the trip and another logbook is commenced. Checking of the data indicates there are indeed 127 trips where this occurred. This feature will need to be considered when linking the FOPS with these trips.

ISSUE #22: Two logbooks used on some trips.

7e. Add TRP_ID to ETBF_LL

Given the issues raised above, this is undertaken in several stages.

1) For those trips where DEPART=ARRIVE (i.e., single day trips) identify the corresponding VSL_ID, FOP_DATE in ETBF_LL and append the TRP_ID. (230 records)
2) For each multiple day trip where INIT_BOOK=LAST_BOOK add the TRP_ID to the corresponding FOPs in the ETBF_LL table that satisfy the following criteria: matching VSL_ID, FOP_DATE between DEPART and ARRIVE and LOG_NO between INIT_LOG_NO and LAST_LOG_NO. One FOP_ID was found to be associated with two trips and upon further investigation the LAST_PAGE for one of the trips was altered and the FOP assigned to a single trip. In total 68394 FOPS were matched with a trip.
3) 122 FOPS were found that had not been matched with a trip in Step 2 but had a FOP_DATE that was between the DEPART and ARRIVE dates (i.e., these FOPS had a corresponding LOG_NO which was not between INIT_LOG_NO and LAST_LOG_NO. Further investigation suggested that the INIT_PAGE_NO or LAST_PAGE_NO for these FOPS was in error and so these were updated so that all FOPS were linked with the associated trips. In total 133 FOPS were matched in this Step.
4) For those trips where INIT_BOOK!=LAST_BOOK add the TRP_ID to the corresponding FOPs in the ETBF_LL table that satisfy the following criteria: matching VSL_ID and FOP_DATE between DEPART and ARRIVE. This matched 914 FOPS.
5) Finally, trips which shared the same INIT_LOG_NO or LAST_LOG_NO were matched with the corresponding FOPS using the same criteria as in Step 4. This matched 32 FOPS.

After completing each of the above five Steps, 1432 FOPS remained unmatched with any trip.

ISSUE #23: FOPS not able to be matched with any trip. In some instances this appears to be due to incorrect recording of logbook page numbers and for some trips it is possible to correct these.

There are also trips in the TRIP table that are not matched to any FOPS. Upon inspection of these TRIPS it was found that they were all associated with FOPS in the WTBF. As such, these trips were deleted from the ETBF_TRIPS table.

7f. Check Trip Dates

Using the list of FOP dates in the ETBF_LL table, the date of the first and last set, the number of sets and number of days fished was added to the record associated with each trip in the TRIPS table. Using the dates of the last set for a given trip and the dates of the initial sets for other trips, it was possible to match each trip to the next trip for the same vessel. Comparing the date of arrival in port with the date of departure for the next trip indicated 24 situations where the departure date was before the date of arrival.

ISSUE #24: For some vessels, the departure date for a given trip is before the arrival date of the previous trip.

The number of days between a vessel departing and the first set of the trip together with the number of days between the last set of the trip and the vessel arriving at port can also be calculated. Histograms of the number of trips versus these delays are shown below: Delays of more than a week are seen for some trips between departure and setting the first set while delays of up to 8 days are noted between the last set and arrival at the port. These seem to be unreasonably long and indicate that there are further errors in the data (probably with departure and arrival dates) not noted previously.

ISSUE #25: Unreasonably large delays noted in some trips between departure and first set and last set and arrival at port.

Number of days delay	*First-Set - Depart*	*Arrive – Last-Set*
0	7352	1335
1	8594	10916
2	2286	5078
3	767	1479
4	264	439
5	108	140
6	30	45
7	18	9
8	6	2
9	6	
10	4	
11	3	
12	1	
13	1	
14	2	
15	1	
Total	19443	19443

Note that in this appendix section "Listing of Vessel Information pertaining to VSL_IDs in ETBF_LL" (19 pages) has been omitted for brevity.

Contact Us
Phone: 1300 363 400
+61 3 9545 2176
Email: enquiries@csiro.au
Web: www.csiro.au

Your CSIRO