CSIRO MARINE RESEARCH

EXCHANGE AND ANALYSIS OF HISTORICAL SOVIET FISHERY SURVEY DATA FROM THE WATERS AROUND AUSTRALIA

J. Anthony Koslow, Peter Last,

Alan Williams and Paul Sabourenkov

Final Report to the

Fisheries Research Development Corporation

FRDC Project 93/239

February 1999

CONTENTS

NON-TECHNICAL SUMMARY

93/239	EXCHANGE AND ANALYSIS OF HISTORICAL SOVIET FISHERY SURVEY DATA FROM THE WATERS AROUND AUSTRALIA

Principal investigator: Dr. Tony Koslow

Address: CSIRO Division of Marine Research

GPO Box 11538

Hobart Tas 7001

Telephone: 03 62325 358 FAX: 03 62325 000

Objectives:

1. Computerize all fishery data from Soviet research cruises in Australian waters, 1963-1975: species composition by trawl, length frequency, maturity stage, and stomach content data
2. Validate and update species identifications for species in instances where the identification is in doubt, based on cruise material
3. Archive the data in the AFZIS database
4. Analyze the data for patterns of distribution and relative abundance for major species

NON-TECHNICAL SUMMARY

To deal effectively with issues, such as the impacts of climate change and of fishing, baseline data are needed, from which comparisons can be made. Consistent, long-term fishery-independent survey data are invaluable, but they are costly, particularly for a country such as Australia, with a vast coastline and relatively small fisheries. In the 1960s and 1970s, the Soviet Union carried out extensive fishery surveys on the shelf and upper slope of northern, western and southern Australia. The data set provides a unique record of the relative abundance of Australian fishes over an approximately ten-year period, but for many years the data were not accessible to researchers outside the Soviet Union. This changed, however, with the collapse of the Soviet Union.

Data were computerized from some 35 cruises carried out by Soviet research vessels to survey fisheries in Australian waters between 1964 and 1976, prior to Australia's declaration of the 200-mile EEZ. There were ~ 4000 demersal trawls carried out during this period with comparable gear. Most trawls were on the continental shelf, primarily in the Great Australian Bight, Northwest Shelf and Gulf of Carpentaria. On the other hand, coverage was not

comprehensive off eastern and southeastern Australia, except Bass Strait, and off central western Australia. Most survey effort was concentrated between 1965 and 1968.

Overall catch rates were analyzed by region and depth. Catch rates were generally highest in the Great Australian Bight and Southwest Cape regions and lowest in the tropics. Catches on the shelf were higher than over the slope, except off southeastern Australia, where relatively high catches were obtained over the upper slope. The Soviets did not explore depths greater than about 500 m.

The Soviet cruises consistently recorded catch composition by species, but neither weight nor numbers of individual species were consistently recorded, so most analyses were based on presence-absence data. The length frequency, maturity stages and diet of important commercial species were also recorded. The taxonomy of the Australian fish fauna was poorly known at the time of the surveys. We checked the identification of specimens collected during the cruises, but most specimens had been disposed of, so data analysis was restricted to species whose identity was highly certain, and we did not examine species diversity.

Nonhierarchical cluster analysis was used to examine community structure around the Australian continent. Ten distinct and robust clusters of stations were found, most of which corresponded to distinct geographic regions and depths and to a lesser degree with season. Results of the cluster analysis were largely consistent with a prior bioregionalization of the Australian marine fish fauna. Of particular note, confirmation was found for considering the North West Shelf and Gulf of Carpentaria to have distinct assemblages. Distinct assemblages were found characteristic of the fish fauna around Tasmania and the Great Australian Bight. On the other hand, distinct faunal communities were not observed off the west or east coasts of Australia, but this may be due to insufficient sampling.

Despite the limitations of the data, the data set provides a valuable historical benchmark for Australian fisheries. The clarification of Australia's marine bioregions should also assist in planning a national system of representative marine protected areas.

BACKGROUND AND NEED

To deal effectively with issues, such as the impacts of climate change and of fishing, it is essential to have baseline data from which comparisons can be made. Consistent, long-term fishery-independent survey data are invaluable for such purposes. Such data are costly to obtain, particularly for a country such as Australia, with a vast coastline and relatively small fisheries. As a result, Australia today finds itself with only a sporadic patchwork of fishery-independent survey data that can be used to assess the impacts of natural and anthropogenic change.

In the 1960s and 1970s before Australia instituted its 200-mile limit, the Soviet Union carried out extensive fishery surveys (35 in all) on the shelf and upper slope of northern, western and southern Australia (Table 1). The surveys were carried out and the data recorded in a consistent manner throughout the period. The data set thus provides a unique record of the relative abundance of Australian fishes around much of Australia over an approximately ten-year period.

Table 1. List of Soviet fishery survey cruises undertaken in the waters around Australia showing the vessel, dates, number of trawls, regions surveyed and whether the data were entered in Australia. GAB: Great Australian Bight; GOC: Gulf of Carpentaria; pelagic: pelagic trawls only

Vessel	Start Date	End Date	# of Trawls	NE	E	SE	GAB	SW	W	NW	GOC	Pelagic	Entered in Australia
BERG-1	13-Mar-65	07-May-65	144		3	3	3	3	3				3
BERG-2	08-Jan-66	22-Jun-66	649			3	3	3					3
SESKAR	18-Jan-66	17-Jun-66	281			3	3	3	3	3			
RADUGA	27-Aug-66	13-Mar-67	502			3	3	3	3	3			3
LIRA	18-Feb-67	02-Aug-67	672			3	3	3	3	3			3
BERG-3	26-May-67	20-Jul-67	372							3	3		3
KORIFEI	16-Feb-68	10-Jun-68	243			3	3	3		3			3
LIRA	04-Jun-68	30-Oct-68	226	3	3						3		3
SUTCHAN	19-Jul-68	16-Jan-69	163									3	
PR.DERUYGIN	21-Sep-68	21-Sep-68	1							3	3		
PROMETEY	18-Nov-68	17-Mar-69	410			3	3	3			3		3
SRTM 8-449	27-Mar-69	01-Jul-69	179			3	3	3			3		
ALBA	21-Sep-69	24-Sep-69	23								3		
PROMETEY	16-Feb-70	03-Jul-70	201							3	3		
ALBA	04-Sep-70	20-Sep-70	29			3	3						
ALBA	04-Mar-71	06-Apr-71	92			3	3		3	3			
POSEIDON	22-Jul-71	20-Aug-71	64				3		3	3			
EQUATOR	10-Sep-71	07-Oct-71	68	3			3		3	3			
RADUGA	19-Jun-	30-Nov-	389				3	3	3	3	3		

	72	72										
PR.DERUYGIN	29-Oct-72	27-Mar-73	34	3	3	3			3	3	3	
LIRA	25-Apr-73	05-Oct-73	316				3	3	3	3	3	
ALBA	17-Oct-73	29-Oct-73	31				3		3	3		
SHANTAR	30-Oct-73	02-Nov-73	7								3	
PR.DERUYGIN	05-May-74	07-May-74	11							3		
SHANTAR	05-May-74	28-Oct-74	37					3	3			
RADUGA	09-Mar-75	01-Jun-75	86	3	3						3	
BACAEVO	09-Jun-75	23-Jul-75	58	3	3							
PR.DERUYGIN	21-Dec-75	13-May-76	14			3			3	3		
KAMENSKY	01-Jul-76	19-Oct-76	36								3	
PR.DERUYGIN	05-Jan-77	10-May-77	32			3			3	3		
TICHOOKEANSCY	27-Mar-77	09-Apr-77	38			3						
POSEIDON	06-Apr-77	01-Sep-77	122								3	
PULK.MERIDIAN	03-Aug-77	08-Aug-77	24									3
TICHOOKEANSCY	02-Jan-78	09-Jan-78	29			3						
MYS TICHY	22-Mar-78	01-Apr-78	27									3

Historical background

Until recently the Soviet historical fishery data set was inaccessible to Australian researchers, who had sought to obtain it since at least the late 1980s. However with the collapse of the Soviet Union and virtual 'privatization' of research institutes, the data set was suddenly made available. Early in 1992, two Russian research scientists, Drs. Ernst Nosov, Head of the Laboratory of Resources of the South Pacific at the Pacific Research Institute of Fisheries and Oceanography (TINRO) and Konstantin Zgurovsky, a senior biologist from TINRO and fluent English speaker, came unannounced to CSIRO in Hobart to discuss a joint project to transfer the historical fishery data to Australia.

Based on this visit, Dr. Tony Koslow (CSIRO Marine Research) applied to the Department of Industry, Technology and Commerce (DITAC) for a travel grant to visit Russia to assess the feasibility of the project. Specifically, he proposed to inspect the data holdings at the TINRO laboratory in Vladivostok and the fish collections in Moscow and St. Petersburg, which held the fish taxonomic collections required to validate the identifications from the Australian cruises. With funding from DITAC, this trip was carried out in May 1993. The data, contained in the original hardbound notebooks, were in good order and had been entered using a consistent standard format throughout the survey period. Many fish collections from the surveys were lost during a move of TINRO to different quarters. However, remaining collections housed at the Zoological Institute of the Russian Academy of Sciences in St. Petersburg and the Zoological Museum of the University of Moscow appeared well-curated and sorted by species so they could be readily accessed. The remaining collections seemed sufficient to validate the identity of the major species from the surveys.

On the basis of the feasibility study, Dr. Koslow applied to FRDC for funds for the Russians to computerize the data holdings, validate the taxonomy, and analyze the data for patterns of distribution and relative abundance for major species.

It is not usual to detail the various hurdles that a project may face through the course of its life, but the difficulties experienced in this project are too extraordinary not to recount; they also explain why the lifetime of the project extended well beyond its original completion date. This history may also prove instructive to others undertaking collaborative projects with institutions not grounded in 'western' scientific traditions.

Although a Memorandum of Understanding regarding the project was signed by the Chief of CSIRO Division of Fisheries and Dr. Valery Akulin, Director of TINRO, there appears to have been little support for the project by the senior management of TINRO. From its inception, there appears to have been jealousy of the project leaders, and roadblocks were thrown up continuously in the way of the project. First, Dr. Zgurovsky was sacked from TINRO, apparently for having been too entrepreneurial. When a CSIRO database specialist, Mr. Aubrey Harris, went to Vladivostok to provide two computers for data entry and training in the use of Oracle for data entry, he was refused entry when he arrived at TINRO. Dr. Peter Young, Chief of Division, intervened successfully, but until then, training was carried out offsite. Although the project provided funds in excess of normal salaries to carry out the project, TINRO gradually withdrew support for it. One of the trained computer operators was transferred from the project shortly after Mr. Harris left TINRO. In the end, prior commitments notwithstanding, all personnel associated with the project were sacked and TINRO's Laboratory of Resources of the South Pacific was disbanded. Ultimately, Dr. Nossov formed a private business that brought the data to Australia, where the data entry was completed.

PROJECT OBJECTIVES

• Computerize all fishery data from Soviet research cruises in Australian waters, 1963-1975: species composition by trawl, length frequency, maturity stage, and stomach content data

* Validate and update species identifications for species in instances where the identification is in doubt, based on cruise material

• Archive the data in the AFZIS database

• Analyze the data for patterns of distribution and relative abundance for major species

METHODOLOGY

Soviet trawl methods

The Soviet data were from demersal and pelagic trawls using similar vessels with similar gears. The demersal trawls typically had a wing spread of 8 m (the wingspread of their largest net was 12 m). The headline was approximately 6 m. The total length of the net was approximately 40 m. The mesh in the codend was 40 mm (stretched mesh); often 4 layers were used. The ground gear consisted of approximately 5 large steel bobbins, 60-70 cm in diameter, and separated by steel spacers. The combined sweep and bridle length was about 40 m. The pelagic trawls were typically 24 m across the mouth, about 15 high and towed at 6 knots. Vessel power was approximately 5,000 HP.

Data entry

All data from the Soviet surveys in Australian waters were computerized through entry into an Oracle database: the position, depth and other characteristics of the trawl sample; the catch composition by species based on numbers or weights or simply their presence (total catch weight was always noted but the entry for different species was not consistent); and the length frequency, maturity stage (spawning condition) and stomach fullness for key species, as available. The data fields in the database are provided in Appendix 3. 40% of the data were entered at the TINRO laboratory (Table 1). Validation procedures at the TINRO laboratory cannot be ascertained and the original data books for those data were not available to check the initial data entry. Data entry at CSIRO was checked for accuracy and was initially poor but was re-done, until it was virtually error-free.

Taxonomic validation

Some 1,700 nominal species of fish are named in the original catch composition lists but many of these are problematic. The identifications were made at a time when the Australian marine demersal fish fauna was relatively poorly known and documented in only a preliminary literature. Hence, many of the species names in the data set are out of date, valid species are recorded under a variety of synonyms or mis-spelled names, and other species do not occur in the region. Before analysing the TINRO data set it was necessary to evaluate the reliability of the original species identifications and upgrade them to the extent possible.

Voucher specimens collected during the TINRO cruises are held at the Russian Academy of Sciences Institute of Oceanography and State University Zoological Museum in Moscow, and the Russian Academy of Sciences Zoological Institute in St. Petersburg. These specimens were re-identified by two fish taxonomists from CSIRO Marine Research, Drs. Peter Last and Alan Williams, and the information used as a guide to estimate the reliability of identifications for the data set and enable computer coding of species (using the revised CAAB list) for analysis.

Priority was given to the dominant species based on their frequency of occurrence in the data set. However, specimens of some key species and several important families were not represented in the collections. For each species examined, we recorded the original and upgraded identifications, the vessel, region and year of capture, and scored the reliability of the original identification based on a set of criteria (Table 1a). In total, 106 and 113 lots were examined in Moscow and St. Petersburg collections, respectively.

Using this assessment, a reliability score on a scale of 0 –10 was then estimated for the identification of each species in the data set based on the criteria in Table 1b. As part of this process, junior synonyms were upgraded, mis-spelled names corrected, and spurious and pelagic species flagged. CAAB codes were also added.

Species identifications were upgraded on a region by region basis using the eight disjunct biogeographic provinces identified for Australian fishes (IMCRA 1998) (see Figure 1). This was seen as an important step in the methodology because the reliability levels for re-identified species were not uniform geographically (more reliable for temperate than tropical faunas). Some species were accurately identified within one bioregion but misidentified in others. Similarly, the accuracy of identifications within regions depended on the identification sources available at the time, and this varied between regions, and temporally by cruise.

Based on the material re-identified, the reliability of original identifications appeared to be most dependent on region. Reliability was higher in temperate than tropical Australia, with the GAB/SW Australian region most reliable and the NW Shelf least reliable (Table 2). Among cruises, those of the *Akedemia Berg* and *Dimitriy Mendelev* were most reliable, but these were in temperate regions. There was no clear trend in reliability with year, probably because region and/or vessel were more important.

Species were ranked by their frequency of occurrence in each bioregion and tackled in priority order. Typically, there were many infrequently occurring species in each region, which formed a long ‘tail’ at the end of each regional list. Because these species were usually more difficult to identify and are likely to have low utility in analyses, they were not considered in detail. They were assigned reliability scores using the total list of species, i.e. without reference to their region of occurrence. Because of this, their inclusion in analyses should be with due care. This may be by combining them at supraspecific levels but this may not be applicable. Similarly, use of infrequently occurring species in estimating total species numbers by region should be done with great caution. The report of Drs. Last and Williams is included as Apendix 4 and the list of species and reliability scores as Appendix 5.

Table 1 Criteria used to assign reliability scores to (a) the original TINRO identifications of fishes in collections, (b) species from each bioregion

(a) Reliability of original identifications

1. correct to current species or contemporary name

2. correct to most closely available literature taxon

3. sibling species confusion within available literature

4. family correct but wrong group within family

5. mis-identified at family level

– evaluation not possible

(b) Criteria used to assign reliability levels to species lists

Level 1 (Score 8-10)

- species name correct or synonomy clear
- no sibling species or, if siblings, these resolved through literature
- species highly distinctive
- good regional references available at time of original identification

Level 2 (Score 5-7)

- species distinct but not in references available at the time
- species name/synonomy with high likelihood of being correct
- no unresolved sibling species or cryptic siblings
- remote possibility of confusion with species of similar form
- species not in geographical area but unconfused

Level 3 (Score 2-4)

- presently unresolved synonyms
- taxa identified only to "sp" but with chance of resolution
- highly problematic taxonomic group
- species name/synonomy with high likelihood of being incorrect
- species not in geographical area and possibly confused

Level 4 (Score 1)

- taxa referred to as multiple species
- siblings existing and unresolved at time
- species not in geographical area and probably confused
- no regional reference available for group

Level 5 (Score 0)

- true pelagic species or non-fish taxon

Table 2. Reliability scores (see Table 1a for criteria) for original fish identifications by bioregion.

	Reliability scores					
Bioregion	1	2	3	4	5	—
NE Australia	1		4	2	4	
E Australia	1		4		1	
SE Australia/ Tasmania	14	3	7	1		
GAB, SW Australia	46	3	12	2		
W Australia (SW Australia)	1		3	1		
NW Shelf (W Australia)	1	6	22	5		1
N Australia (NW Shelf)	10	2	4	2		2

Methods of analysis

The data set was analyzed to examine total catch rates by depth and by region, using the bioregionalization of Lyne (1996). Analysis of variance was used to examine for significant differences. The catch weight data were log-transformed to normalize the distribution and achieve homogeneity of the variance. The Bonferroni least-significant difference (LSD) and Scheffe multiple range tests were used to test that specific depths or regions were significantly different, due to the problems of multiple comparison.

Because of the uneven taxonomic quality of the data, analyses of species diversity were not carried out.

Data from species with reliability levels of 1 or 2 (Table 1b) were used in multivariate analyses to examine community structure, in particular to examine zonation by depth and bioregion. Because the catch data contained a mix of weights, numbers and simple presence, a distance metric, variously known as the Sorensen's, Lance-Williams or binary Bray-Curtis measure (Legendre and Legendre 1983, p. 200) was used. It is based on presence-absence and is not affected by the joint absence of a species from a sample.

The data set was too large for hierarchical cluster analysis, so non-hierarchical cluster analysis was used, based on the K-MEANS routines in SPSS. Because it is necessary to initially specify the number of clusters in non-hierarchical cluster analysis, we sought the 'optimal' number of distinct clusters. This was based on examination of a range of solutions, the 'optimal' number being the largest number that appeared to correspond to distinct regions, depths, seasons and that contained distinct assemblages of species.

We also assessed the robustness of the clustering through use of the 'jack-knife' method, whereby discriminant functions were defined based upon 90% of stations and the resulting functions were used to classify the remaining 10% of stations into clusters. This was carried out ten times to classify successive 10% portions of the data set, i.e. until all stations had been classified and the proportion classified into the original clusters could be examined.

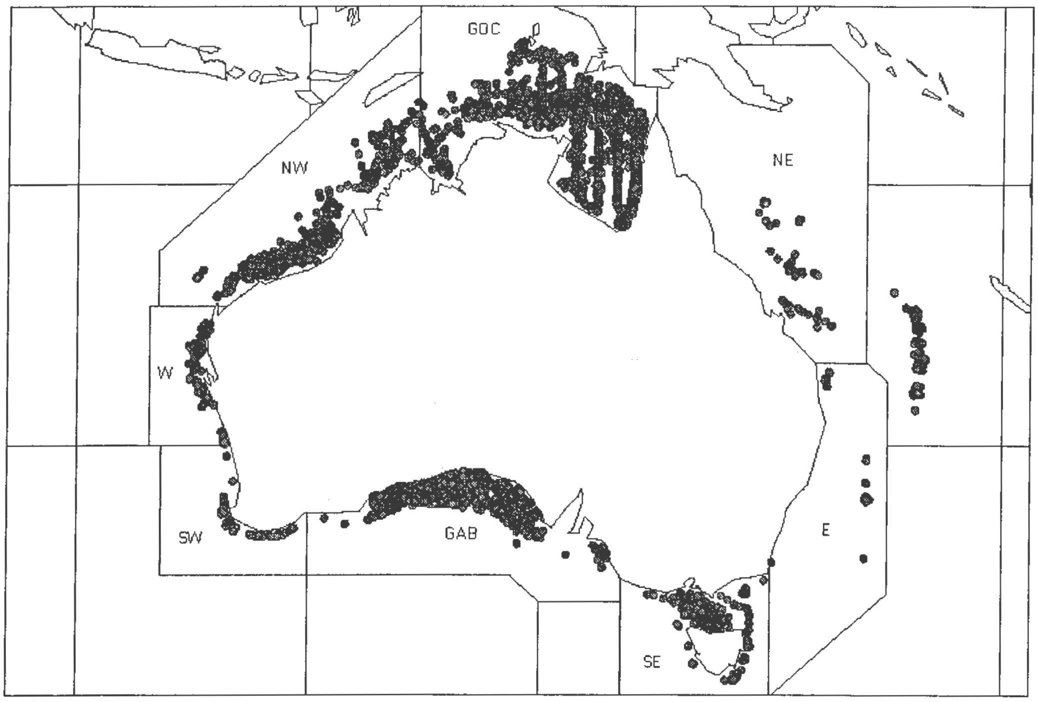


Figure 1. Chart showing the position of Soviet demersal survey trawls carried out in Australian waters, 1964-78.

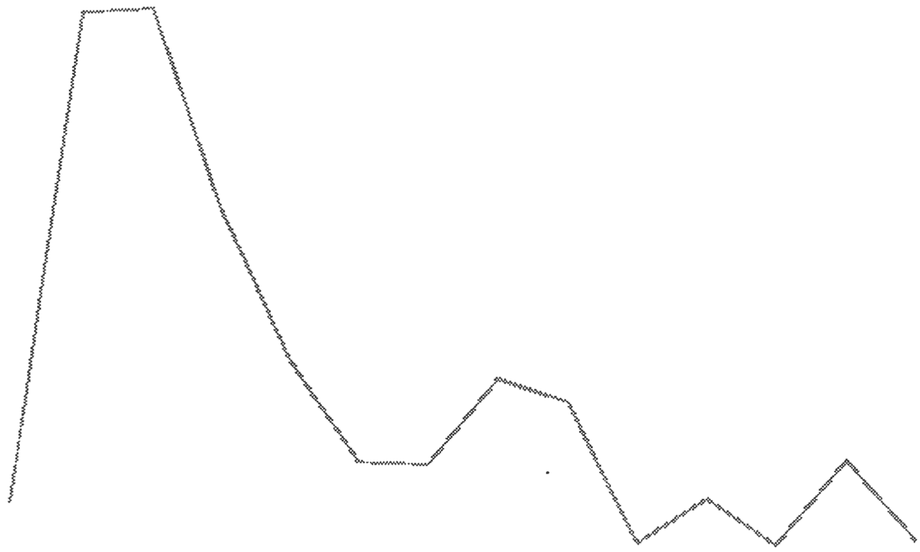


Figure 2. Soviet survey effort (number of trawls per year) in Australian waters, 1964-1978.

RESULTS/DISCUSSION

Survey effort

In all, there were 5376 demersal, 83 pelagic and 204 modified pelagic trawls entered into the database over the period of the Soviet surveys, covering the period from 1965 to 1978. The data from two of the first cruises, which were carried out in 1963 and containing 385 trawls, were not recovered. The distribution of demersal trawls by bioregion is shown in Figure 1. Because of the preponderance of demersal trawls, they are the only data treated in subsequent analyses. The research effort was greatest in the early years of the survey period (1965-68) and dropped off sharply near the end, particularly after 1973 (Figure 2). The distribution of demersal survey effort around Australia is shown in Figure 1. Survey effort was minimal off the central west coast and from northeastern (Cape York) to southeastern Australia (Spencer Gulf), with the exception of Bass Strait. The Great Australian Bight and the area from the Northwest Shelf to the Gulf of Carpentaria were well sampled.

Sampling effort varied by depth in the different bioregions (Figure 3). In the best-sampled areas, sampling concentrated on the continental shelf to 200 m in the Great Australian Bight and southwestern Australia, but concentrated in the upper 100 m in Bass Strait and northern and northwestern Australia.

Figure 3. The number of trawls by region and depth zone. (See Figure 1 for location of regions.)

Catch rate

Comparison of catch rates across regions or depths was hindered by the uneven sampling distribution across regions and depths. It was therefore not possible to analyze the influence of either of these variables in isolation. We also found a highly significant interaction between depth and region when the log-transformed catch data were examined with ANOVA ($F= 9.70$, $df = 28$, $p < 0.001$). As seen from plots of mean catch rate by depth across regions and by region across depths (Figures 4a, b), catches off the shelf (> 250 m) were considerably lower than on the shelf everywhere except off southeast Australia where they were notably higher.

Figure 4a. Mean catch rates by depths within regions.

Figure 4b. Mean catch rates by regions within depths.

Due to the significant interaction effect between region and depth, it was necessary to examine for regional differences by depth and for depth differences by region. Examination of depth differences indicated that the primary depth differences in catch rate were at the deepest and shallowest depth (Table 3). Off southeast Australia, the highest catch rates were at depths greater than 250 m, but this depth range had significantly lower catch rates in areas from the Great Australian Bight and around Western Australia to the Northwest Shelf. Other depth effects were not consistent across regions. The shallowest depth range ((< 100 m) had significantly higher catch rates in the Gulf of Carpentaria but significantly lower catch rates than other depths (except depths > 250 m) in the Great Australian Bight . In general there appeared to be a trend toward the catch rates being highest at shallow depths in tropical regions and at deeper depths on the slope off southern Australia. There were no significant depth differences off eastern Australia, probably due to limited sampling there.

There were significant regional differences at all depths. At all depths except on the slope (> 250 m), the highest catch rates were generally in the Great Australian Bight and the Southwest Cape regions and lowest in the tropics off the Northwest Shelf and in the Gulf of Carpentaria. On the slope, the catch rates off southeastern Australia were significantly higher than elsewhere in Australian waters.

Table 3. Mean catch rates by depth and region (as shown in Figure 1). Number of trawls in parentheses beneath individual catch rates. Standard deviation in parentheses beneath means.

Region	0m - 50m	50m - 100m	100m - 150m	150m - 200m	200m - 250m	> 250m	Region Mean
NE	50.48	1266.54	433.33	750.00		50.86	584.34
	(3)	(8)	(3)	(2)		(7)	(1432.47)
E		538.33		366.67	52.00		403.09
		(6)		(3)	(2)		(607.34)
SE	1002.22	827.16	379.64	145.97	440.00	1059.51	807.95
	(17)	(267)	(18)	(11)	(2)	(29)	(1643.22)
GAB	303.58	499.57	971.00	706.02	697.82	117.15	671.48
	(197)	(719)	(602)	(925)	(59)	(32)	(2585.08)
SW	10.00	1332.19	166.73	116.67	60.00	114.49	718.01
	(1)	(17)	(11)	(3)	(1)	(2)	(1713.22)
W		308.48	536.38	259.71	321.12	58.70	351.91
		(36)	(19)	(7)	(4)	(4)	(628.33)
NW	597.70	352.53	159.83	103.90	196.00	65.94	337.14
	(60)	(503)	(81)	17	(5)	(16)	(562.03)
GOC	494.34	253.43	95.64	38.40	48.22	456.09	346.69
	(408)	(482)	(51)	(5)	(5)	(5)	(557.57)
Depth Mean	458.54	454.68	793.02	679.75	573.35	406.63	560.36
	(805)	(947)	(3260)	(2842)	(1013)	(722)	(2003.18)

Community structure

The demersal trawl species composition data were analyzed with nonhierarchical cluster analysis based upon initial specification of between 5 and 19 clusters. The 11-cluster solution provided the maximum number of clusters that could be distinguished in terms of depth, region, seasonal and species characteristics. These clusters tended to be stable from one solution to another, such that although a few clusters split or became grouped together at higher or lower solutions, most remained unchanged. At higher-cluster solutions, single-station clusters appeared, a clear indication that the optimal number had been surpassed.

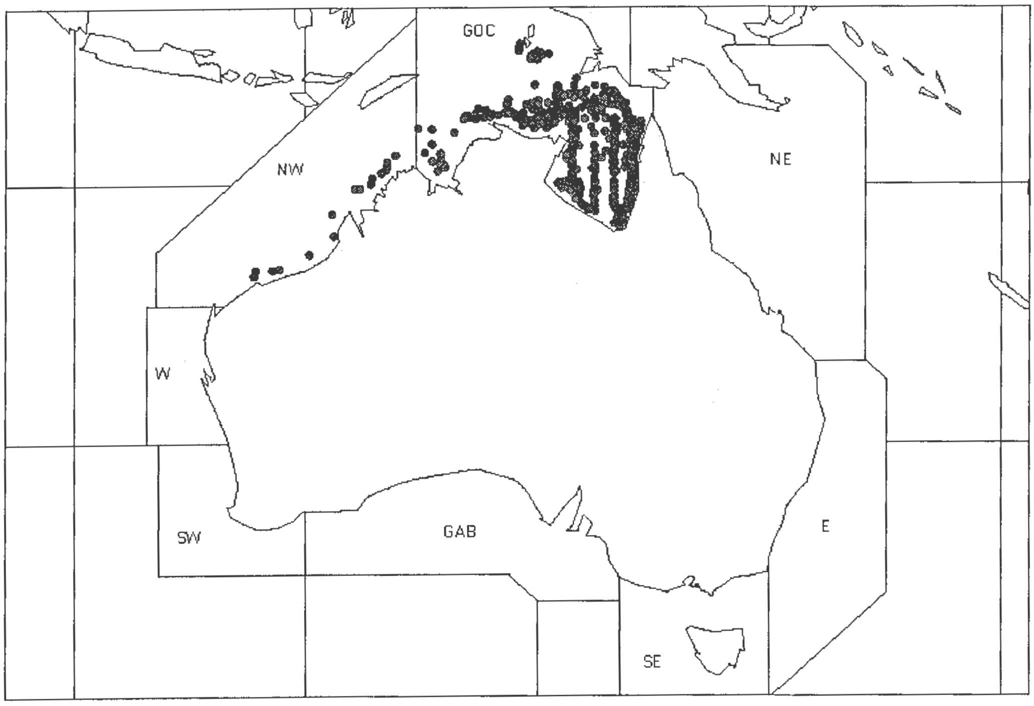
Of the 11 clusters in this solution, 10 appeared to be valid clusters. However, one large cluster (Cluster 4 with ~1060 trawls: 23% of total), which appeared in all solutions, contained a high proportion of stations from all regions and appeared to consist of those stations not clearly linked with any others. These stations contained fewer species on average (12 cf a weighted mean of 18 species per cluster for the other clusters).and, moreover, did not contain any characteristic species. All other clusters contained at least one species that occurred, typically, in >80% of stations, but no species in this cluster occurred in more than 20% of cases. Even when the stations in this cluster were analyzed separately, further significant clusters did not form. The presence of this large cluster seems to indicate a degree of heterogeneity in the samples that cannot be resolved by this clustering method.

Results of the 11-cluster groupings are shown in Figure 5 and Tables 4-7 by bioregions (as defined by Lyne et al. 1996), depth, season and year. The dominant species characterizing each cluster are shown in Table 8.

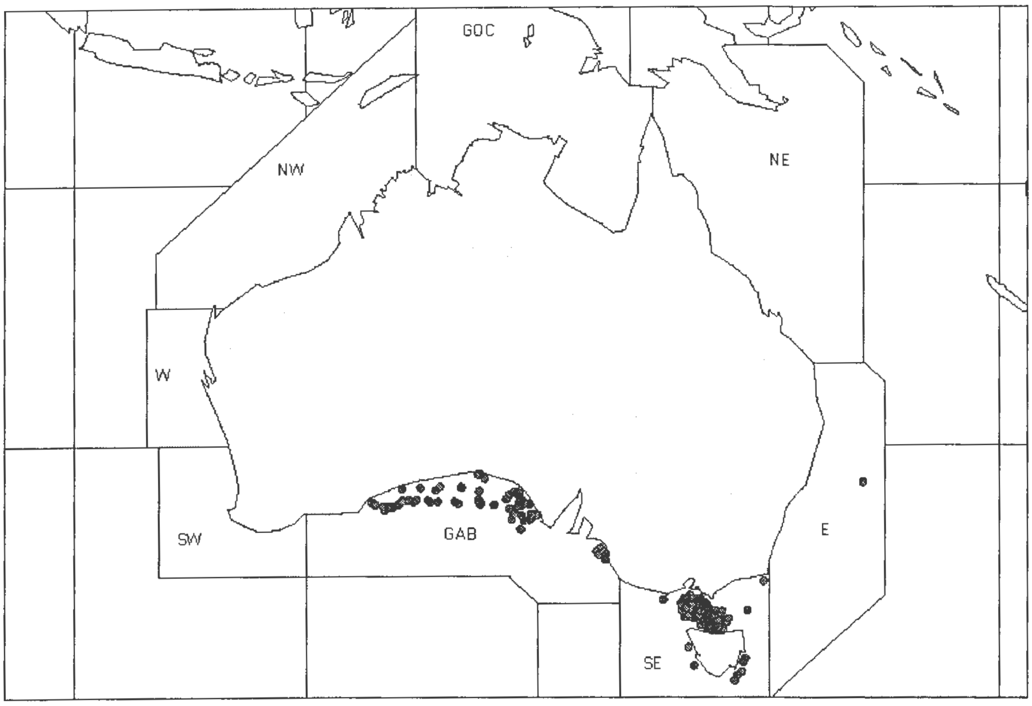
None of the poorly sampled bioregions (i.e. those containing < 100 stations), such as off eastern, western or southwestern Australia appeared to have a distinct assemblage of species (Table 4). It is notable, however, that stations from these areas did not cluster with the clusters from adjacent bioregions, tending rather to fall in with the large indistinct cluster (Cluster 4). This indicates a degree of distinctness between adjacent bioregions, albeit poorly defined.

In the 11-cluster solution, three clusters (Clusters 2, 8 & 11) were distinctly associated with Tasmania. Cluster 11, a small cluster consisting of only 25 stations was most distinct, being based on slope stations at depths > 250 m and being characterized by a suite of characteristically upper slope species: blue grenadier (*Macruronus novaezelandiae*), toothed whiptail (*Lepidorhynchus denticulatus*), ocean perch (*Helicolenus percoides*), mirror dory (*Zenopsis nebulosus*), gemfish (*Rexea solandri*). The other two Tasmanian clusters were based on sampling on the shelf, primarily in Bass Strait, although about a third of the stations in Cluster 2 were from the Great Australian Bight. The stations in Cluster 8 were sampled only in autumn and predominantly from only two years, 1965 and 1969, whereas the assemblage in Cluster 2 was obtained in all seasons and over a number of years. There was considerable overlap in species composition of these two clusters, but Cluster 2 was dominated by only a single species, couta (*Thyrsites atun*), which was present in 94% of stations; no other species was present in even a third of the stations in this cluster. Couta was also a dominant species in Cluster 8, but overall there were ten species that occurred in >50% of stations in this cluster, including silver dory (*Cyttus australis*), white-spotted dogfish (*Squalus acanthias*), red gurnard (*Chelidonichthys kumu*), morwong (*Nemadactylus macropterus*), velvet leatherjacket (*Meuschenia scaber*), ocean perch and common gurnard perch (*Neosebastes scorpaenoides*).

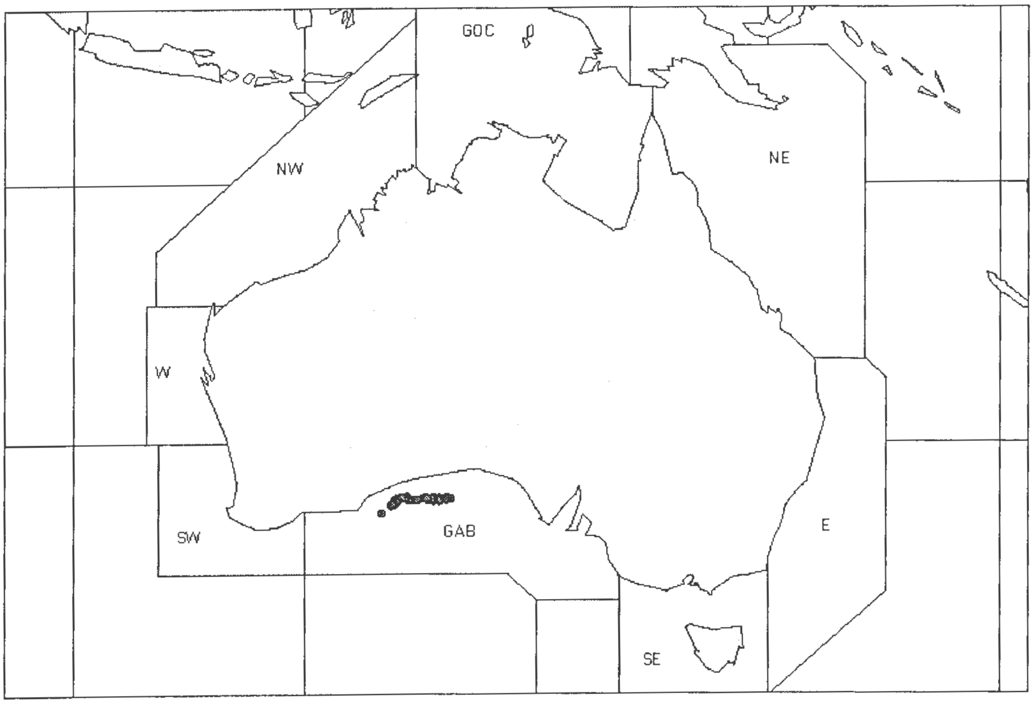
Figure 5. Results of cluster analysis. Maps showing trawl positions by cluster.



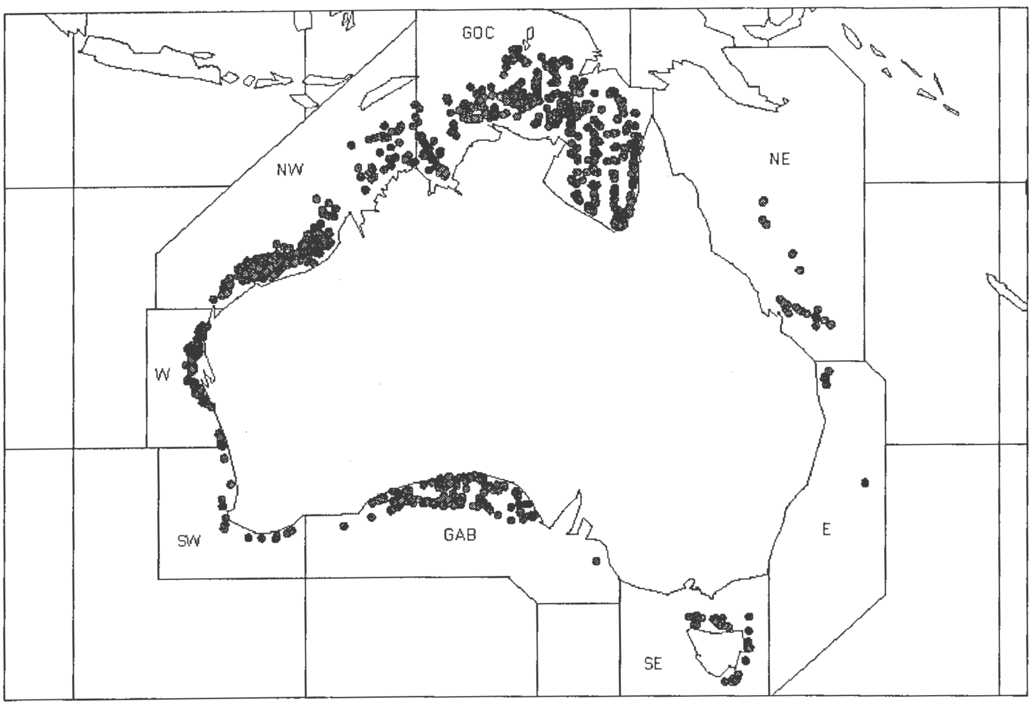
Cluster 1

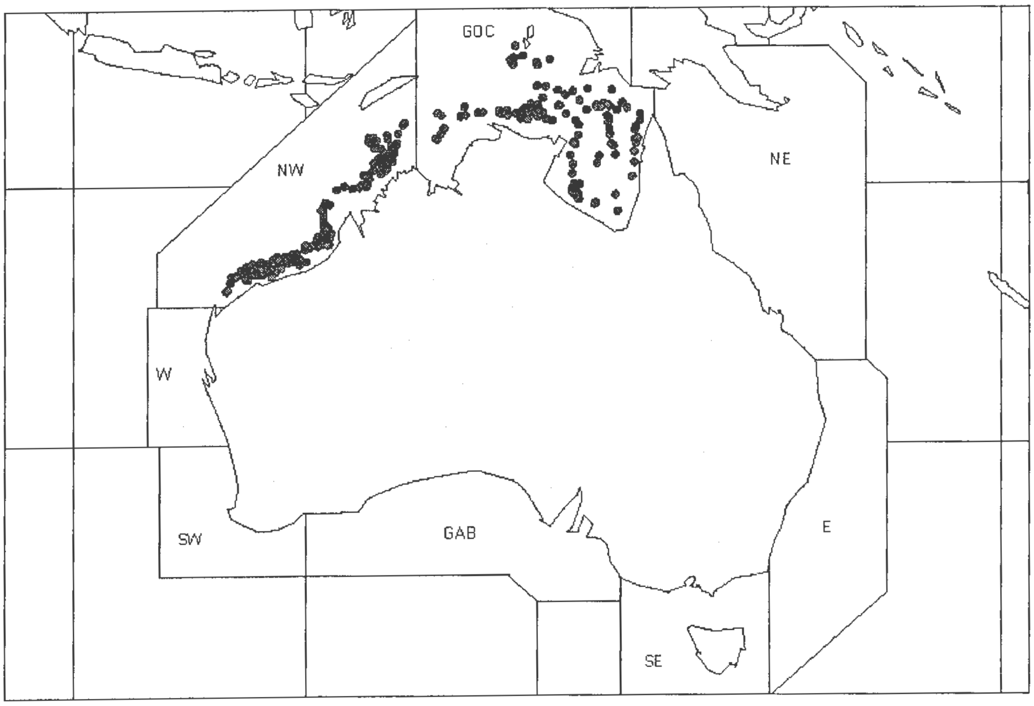


Cluster 2

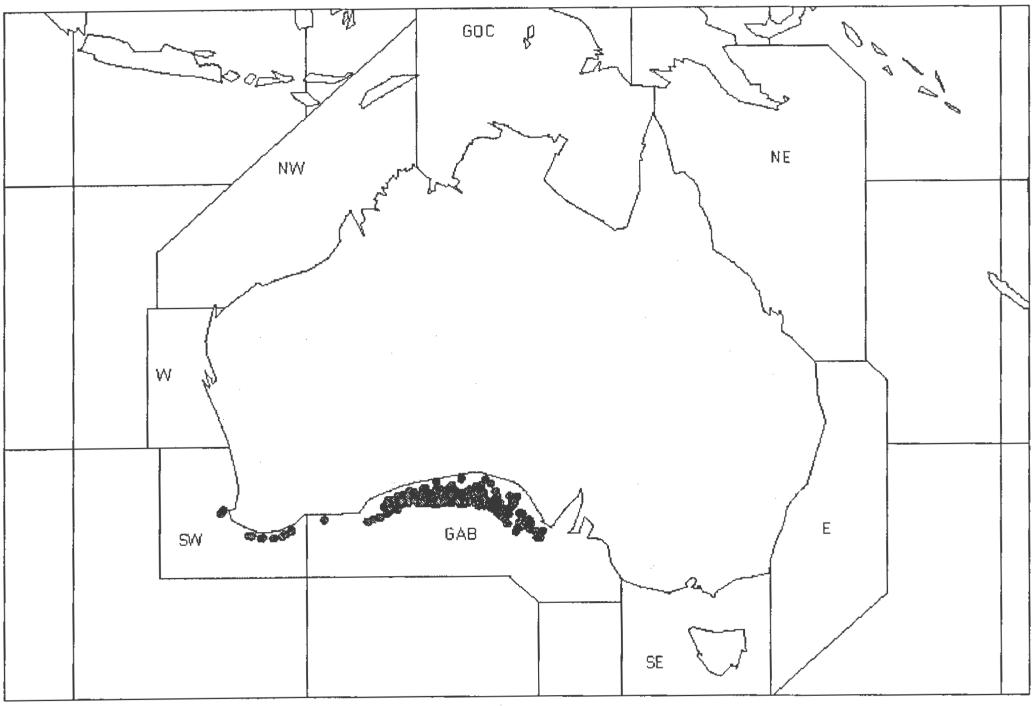


Cluster 3

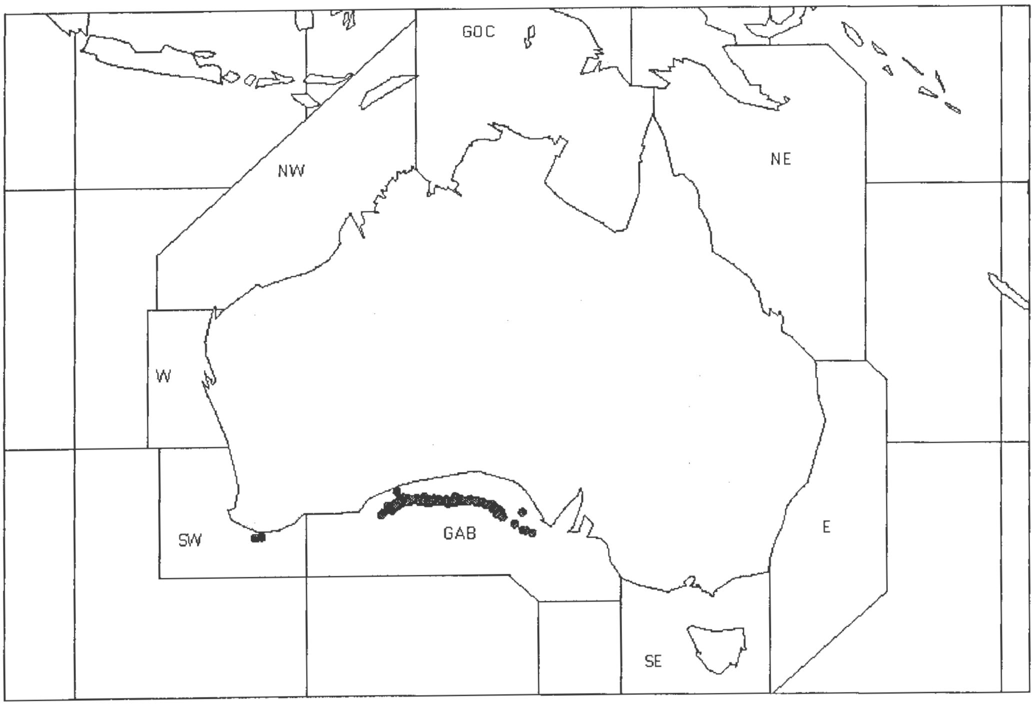




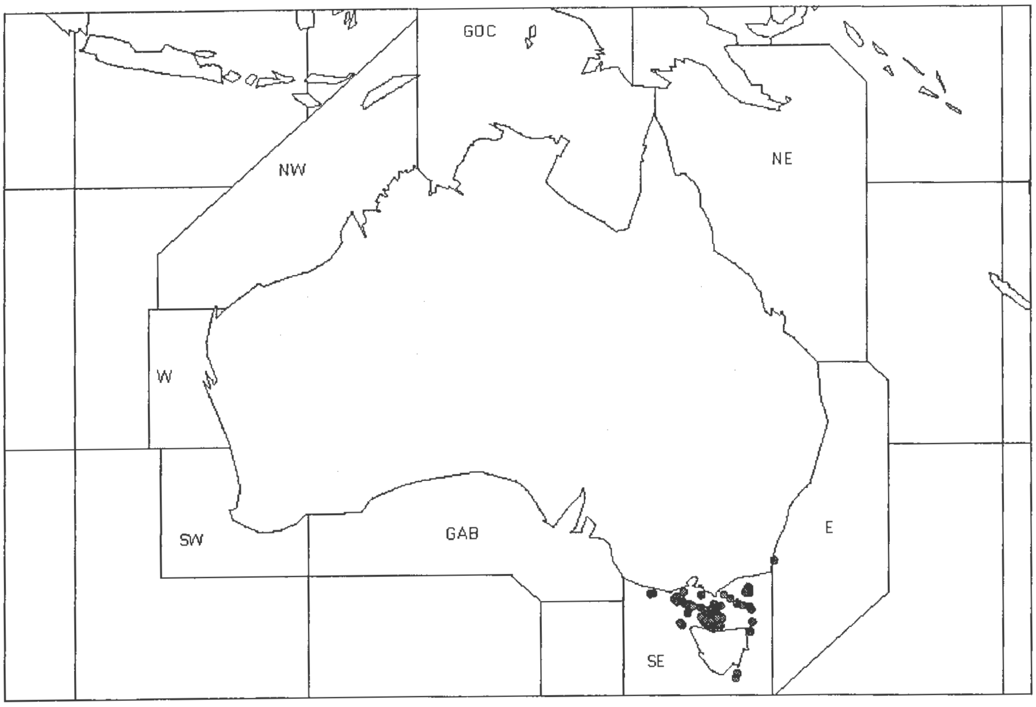
Cluster 5



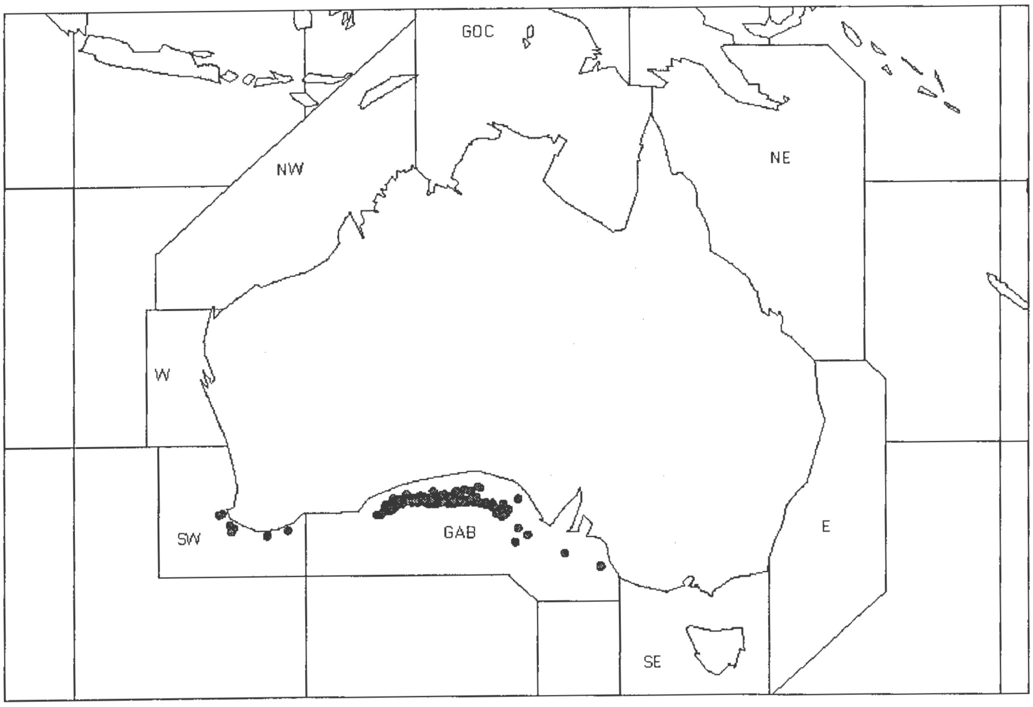
Cluster 6



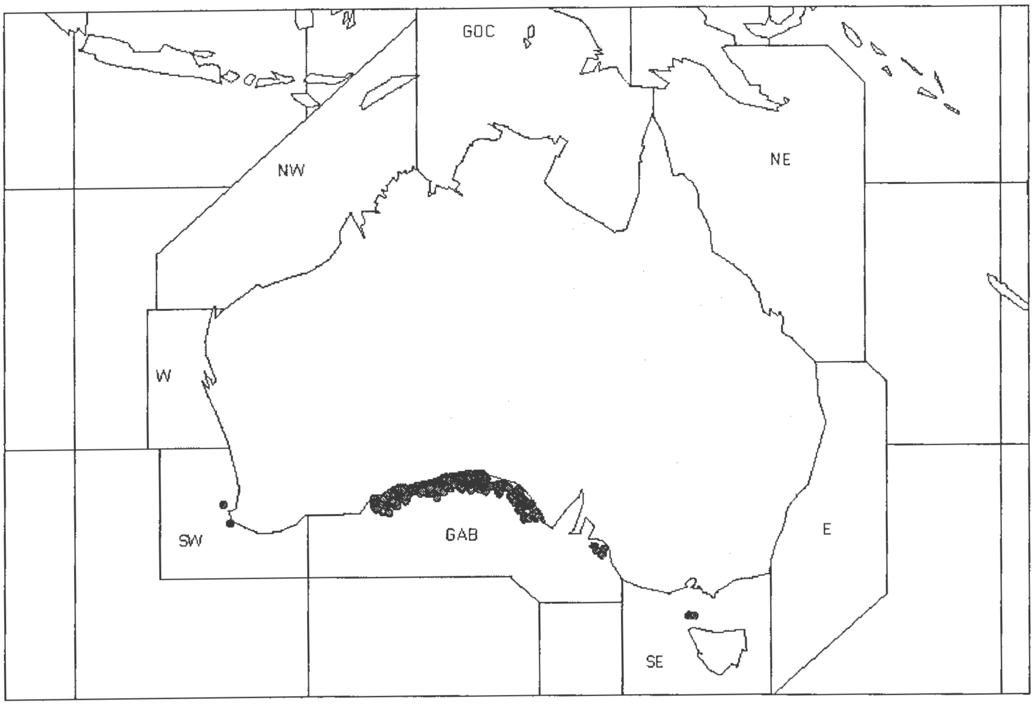
Cluster 7



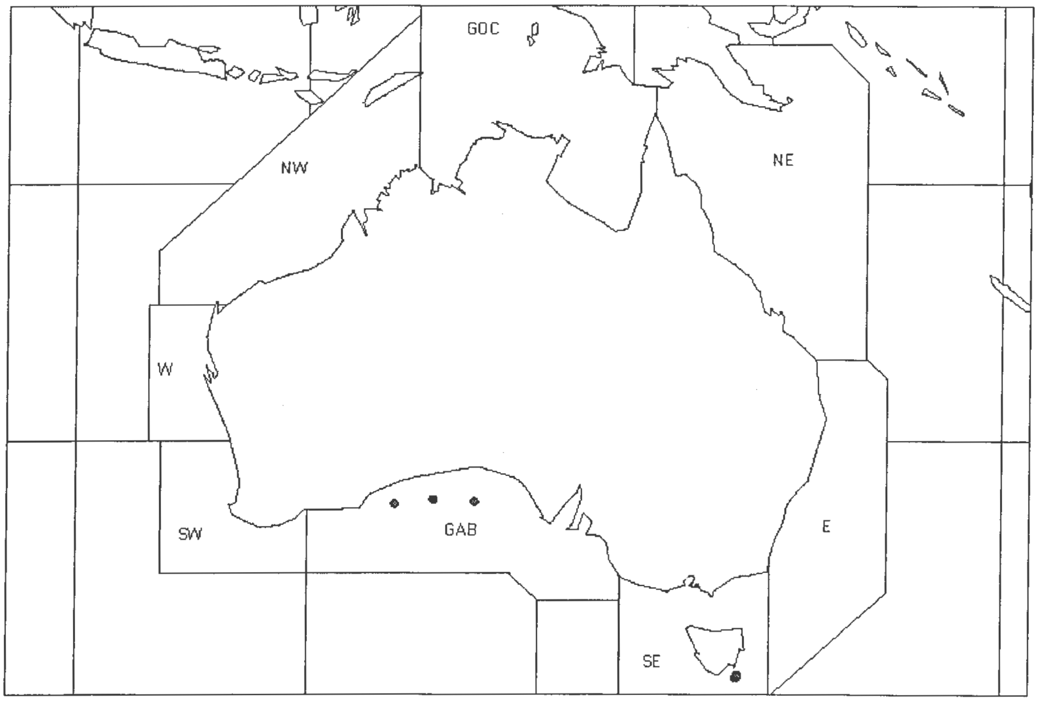
Cluster 8



Cluster 9



Cluster 10



Cluster 11

Table 4. Distribution of trawls from cluster analysis by region (as in Figure 1).

Cluster	NE	E	SE	GAB	SW	W	NW	GOC	Row Total	% of Total
1							18	458	476	10.18%
2		1	191	90					282	6.03%
3				102					102	2.18%
4	25	8	52	166	17	65	348	380	1061	22.69%
5							296	115	411	8.79%
6				408	11				419	8.96%
7				558	2				560	11.97%
8		1	85						86	1.84%
9				618	7				625	13.36%
10			2	629	2				633	13.53%
11			19	3					22	0.47%
Column Total	25	10	349	2574	39	65	662	953	4677	100.00 %
% of Total	0.53%	0.21%	7.46%	55.04%	0.83%	1.39%	14.15%	20.38%	100.00%	

Table 5. Distribution of trawls from cluster analysis by depth.

Cluster	0m - 50m	50m - 100m	100m - 150m	150m - 200m	200m - 250m	>250 m	Row Total	% of Total
1	249	224	3				476	10.18%
2	21	228	18	13	1	1	282	6.03%
3			38	61	2	1	102	2.18%
4	194	592	146	61	14	54	1061	22.69%
5	54	333	24				411	8.79%
6	5	108	211	93	2		419	8.96%
7	2	4	140	397	14	3	560	11.97%
8	3	67	11	3	1	1	86	1.84%
9	3	33	189	349	41	10	625	13.36%
10	166	462	5				633	13.53%
11						22	22	0.47%
Column Total	697	2051	785	977	75	92	4677	100.00%
% of Total	14.90%	43.85%	16.78%	20.89%	1.60%	1.97%	100.00%	

Table 6. Distribution of trawls from cluster analysis by seasons.

Cluster	Mar-May	Jun-Aug	Sep-Nov	Dec-Feb	Row Total	% of Total
1	72	52	194	158	476	10.18%
2	92	36	87	67	282	6.03%
3	102				102	2.18%
4	333	295	217	216	1061	22.69%
5	73	246	57	35	411	8.79%
6	198	137	49	35	419	8.96%
7	375	51	2	132	560	11.97%
8	86				86	1.84%
9	290	25	126	184	625	13.36%
10	52	239	294	48	633	13.53%
11	1			21	22	0.47%
Column Total	1674	1081	1026	896	4677	100.00%
% of Total	35.79%	23.11%	21.94%	19.16%	100.00%	

Table 7. Distribution of trawls from cluster analysis by year.

Cluster	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	Row Tota l	% of Total
1			39	197	142	60		31	7						476	10.2%
2	2	102	28	61	26	4	17	30	6		1		3	2	282	6.0%
3	102														102	2.2%
4	6	82	189	189	136	90	50	73	133	41	22	9	41		1061	22.7%
5			223	50	56	22	3	37	17	2		1			411	8.8%
6	1	51	167	55	24	2	41	24	54						419	9.0%
7		291	219	18	6	2	20	2	2						560	12.0%
8	31		2		51		2								86	1.8%
9	1	352	140	94	5	1	13	5	14						625	13.4%
10		86	133	74	2	20	52	174	91	1					633	13.5%
11		2					1						19		22	0.5%
Column Total	143	966	1140	738	448	201	199	376	324	44	23	10	63	2	4677	100.0%
% of Total	3.1%	20.7 %	24.4 %	15.8 %	9.6%	4.3%	4.3%	8.0%	6.9%	0.9%	0.5%	0.2%	1.3%	0.0%	100. 0%	

Table 8. The dominant ten species in each cluster from the 11-cluster analysis, showing the percentage of stations in which each species occurs.

Cluster	CAABCODE	NAME	Frequency	Percent
1	337015	Selaroides leptolepis	393	82.56%
1	118013	Saurida gracilis	276	57.98%
1	441012	Rastrelliger kanagurta	237	49.79%
1	278002	Fistularia petimba	231	48.53%
1	465011	Abalistes stellaris	223	46.85%
1	321003	Terapon theraps	221	46.43%
1	346003	Lutjanus vitta	216	45.38%
1	311057	Epinephelus tauvina	165	34.66%
1	440004	Trichiurus lepturus	158	33.19%
1	467033	Arothron hispidus	153	32.14%
2	439001	Thyrsites atun	264	93.62%
2	264002	Cyttus australis	89	31.56%
2	465005	Meuschenia scaber	71	25.18%
2	023001	Pristiophorus nudipinnis	68	24.11%
2	469001	Diodon nicthemerus	61	21.63%
2	020006	Squalus megalops	52	18.44%
2	287001	Helicolenus percoides	52	18.44%
2	287006	Neosebastes thetidis	52	18.44%
2	337003	Trachurus novaezelandiae	51	18.09%
2	007001	Heterodontus portusjacksoni	48	17.02%
3	258004	Centroberyx gerrardi	101	99.02%
3	377003	Nemadactylus macropterus	97	95.10%
3	465002	Acanthaluteres vittiger	94	92.16%
3	465006	Nelusetta ayraudi	94	92.16%
3	369002	Oplegnathus woodwardi	91	89.22%
3	020008	Squalus acanthias	90	88.24%
3	367005	Zanclistius elevatus	88	86.27%

3	023001	Pristiophorus nudipinnis	87	85.29%
3	288001	Chelidonichthys kumu	83	81.37%
3	024002	Squatina tergocellata	77	75.49%
4	465011	Abalistes stellaris	213	20.08%
4	278002	Fistularia petimba	181	17.06%
4	118013	Saurida gracilis	172	16.21%
4	440004	Trichiurus lepturus	172	16.21%
4	346007	Lutjanus malabaricus	169	15.93%
4	346004	Lutjanus sebae	162	15.27%
4	337015	Selaroides leptolepis	153	14.42%
4	351001	Lethrinus nebulosus	100	9.43%
4	321003	Terapon theraps	99	9.33%
4	346003	Lutjanus vitta	98	9.24%
5	346003	Lutjanus vitta	340	82.73%
5	351001	Lethrinus nebulosus	301	73.24%
5	346004	Lutjanus sebae	274	66.67%
5	118013	Saurida gracilis	245	59.61%
5	346007	Lutjanus malabaricus	243	59.12%
5	278002	Fistularia petimba	192	46.72%
5	353006	Argyrops spinifer	171	41.61%
5	465027	Pseudobalistes fuscus	169	41.12%
5	337027	Caranx ignobilis	142	34.55%
5	346034	Lutjanus fulviflamma	142	34.55%
6	465006	Nelusetta ayraudi	378	90.21%
6	258004	Centroberyx gerrardi	317	75.66%
6	369002	Oplegnathus woodwardi	299	71.36%
6	288006	Pterygotrigla polyommata	280	66.83%
6	367001	Paristiopterus gallipavo	277	66.11%
6	367005	Zanclistius elevatus	250	59.67%
6	038008	Urolophus expansus	243	58.00%
6	258005	Centroberyx lineatus	181	43.20%
6	377003	Nemadactylus macropterus	152	36.28%
6	024002	Squatina tergocellata	147	35.08%
7	377003	Nemadactylus macropterus	524	93.57%
7	258004	Centroberyx gerrardi	518	92.50%
7	020006	Squalus megalops	488	87.14%
7	369002	Oplegnathus woodwardi	487	86.96%
7	465006	Nelusetta ayraudi	472	84.29%
7	367005	Zanclistius elevatus	467	83.39%
7	345002	Plagiogeneion macrolepis	391	69.82%
7	288006	Pterygotrigla polyommata	366	65.36%
7	024002	Squatina tergocellata	363	64.82%
7	023001	Pristiophorus nudipinnis	291	51.96%
8	264002	Cyttus australis	70	81.40%
8	439001	Thyrsites atun	64	74.42%
8	020008	Squalus acanthias	58	67.44%
8	288001	Chelidonichthys kumu	57	66.28%
8	377003	Nemadactylus macropterus	57	66.28%
8	023001	Pristiophorus nudipinnis	49	56.98%
8	465005	Meuschenia scaber	49	56.98%
8	287001	Helicolenus percoides	48	55.81%
8	287005	Neosebastes scorpaenoides	47	54.65%

8	469002	Allomycterus pilatus	47	54.65%
9	258004	Centroberyx gerrardi	535	82.06%
9	377003	Nemadactylus macropterus	512	78.53%
9	369002	Oplegnathus woodwardi	492	75.46%
9	345002	Plagiogeneion macrolepis	287	44.02%
9	367005	Zanclistius elevatus	252	38.65%
9	465006	Nelusetta ayraudi	251	38.50%
9	337003	Trachurus novaezelandiae	240	36.81%
9	465031	Balistoides conspicillum	216	33.13%
9	377004	Nemadactylus valenciennesi	179	27.45%
9	020006	Squalus megalops	165	25.31%
10	085002	Sardinops neopilchardus	460	72.67%
10	439001	Thyrsites atun	383	60.51%
10	469002	Allomycterus pilatus	360	56.87%
10	039001	Myliobatis australis	358	56.56%
10	367005	Zanclistius elevatus	285	45.02%
10	465006	Nelusetta ayraudi	255	40.28%
10	258005	Centroberyx lineatus	237	37.44%
10	038008	Urolophus expansus	216	34.12%
10	355029	Upeneichthys vlamingii	210	33.18%
10	349001	Parequula melbournensis	196	30.96%
11	227001	Macruronus novaezelandiae	21	95.45%
11	232004	Lepidorhynchus denticulatus	21	95.45%
11	287001	Helicolenus percoides	20	90.91%
11	264003	Zenopsis nebulosus	19	86.36%
11	439002	Rexea solandri	17	77.27%
11	042003	Hydrolagus lemures	16	72.73%
11	228002	Genypterus blacodes	14	63.64%
11	017008	Galeorhinus galeus	12	54.55%
11	013004	Parascyllium variolatum	11	50.00%
11	445001	Hyperoglyphe antarctica	11	50.00%

The Great Australian Bight (GAB) was best sampled, containing 2574 demersal trawls (55% of the total), and it contained the most clusters: 5 (Clusters 3, 6, 7, 9 & 10: Table 4). Several, but not all, these clusters showed distinct depth zonation. Cluster 10 was composed of stations within the inner shelf (depths < 100m) and was dominated by pilchard (*Sardinops neopilchardus*) and also couta, the eagle ray (*Myliobatus australis*) and Australian burrfish (*Allomycterus pilatus*). The remaining clusters from the GAB were comprised of stations from the mid to outer shelf and exhibted considerable overlap in species composition. Cluster 3 was based on a single cruise in autumn 1965 and from stations predominantly on the outer shelf (150-200 m). Fully nine species were obtained on over 80% of these stations. The dominant species in this assemblage was red snapper (*Centroberyx gerrardi*), but the characteristic species of this assemblage, not found among the ten most common species of the other GAB assemblages, were toothbrush leatherjacket (*Acanthaluteres vittiger*), white-spotted dogfish and red gurnard. Clusters 7 and 9 were also comprised of stations from the outer shelf (predominantly 150-200 m) and exhibited no clear differences in terms of the season or years in which the stations were sampled. There was also considerable overlap in their species composition. The clearest difference between the assemblages in these clusters was in their diversity or evenness, such that Cluster 7 contained 10 species found in >50% of stations, whereas Cluster 9 contained only three such species, the red snapper, jackass morwong (*Nemadactylus macropterus*) and the knifejaw (*Oplegnathus woodwardi*). The remaining GAB cluster was centred at mid-shelf depths (100-150 m) and was dominated by Chinaman leatherjacket (*Nelusetta ayraudi*). Again there was considerable overlap in species composition with other GAB clusters from mid- to outer-shelf depths, including red snapper; however, several species characteristic of this cluster were not found among the 10 most common species of other GAB clusters: yellowspotted boarfish (*Parisiopterus gallipavo*) and the wide stingaree (*Urolophus expansus*).

Two tropical bioregions were clearly defined by the cluster analysis: the Northwest Shelf (Cluster 5)and the Gulf of Carpentaria (Cluster 1). Both regions were well sampled: over 400 trawls distributed over all four seasons and a number of years. Both were predominantly sampled at depths less than 100 m. Dominant species in the Northwest Shelf were a suite of five lutjanids and the spangled emperor (*Lethrinus nebulosus*). The dominant species in the Gulf of Carpentaria assemblage was yellow-striped trevally (*Selaroides leptolepis*). Several species were common in both areas, such as the lizardfish, *Saurida gracilis*, rough flutemouth (*Fistularia petimba*) and one-band sea-perch (*Lutjanus vitta*).

Discussion

Completion of this project required that severe obstacles be overcome, beyond those normally experienced by a scientific project. Our counterparts at the TINRO laboratory in Vladivostok, apparently unfamiliar with the normal standards for management of a scientific project, reneged on promises of cooperation in signed MOUs, disbanded our partner in the enterprise, the Laboratory of the Resources of the South Pacific and fired its scientists. As a result the data entry, envisaged to require 6 months, required several years to complete and eventually was carried out at CSIRO rather than at TINRO. However in the end, all objectives of the project were ultimately completed within budget.

There are two primary uses of the Soviet historical survey data. The first is as a benchmark of the species composition, catch rate and length composition of fish stocks in the Australian EEZ in the 1960s and 1970s. To some degree, this usefulness is compromised by the uneven sampling by depth and region of the surveys; by uncertainty regarding the taxonomy, particularly of less common species; by the particular fishing characteristics of the Soviet fishing gear, which is unlikely to be duplicated in future surveys; and by the inconsistent recording of catch weight and numbers for the various taxa. As a result, only major changes in these variables are likely to be detected by reference to this historical data set. The data are archived with FRDC and at CSIRO and are available for further analysis. Despite these caveats, however, the Soviet data set stands as the most comprehensive data set available for Australian continental shelf fishes. Length-frequency and reproductive stage data, which were not analyzed as part of this project, are also available.

The second major use is to define the primary biogeographic zones, or bioregions, within the Australian EEZ. Due to limited sampling off eastern, western and southwestern Australia, the data set is of limited use for those regions. However, as the largest Australian fishery data set, collected with consistent gear and sampling methodology, it can be used to clarify some biogeographic issues and provide distributional information on a wide range of commercial species.

Several biogeographic schemes for the Australian marine fauna have been proposed in the past. The first major biogeographic scheme was developed by Whitley (1937): a Solanderian region off the Great Barrier Reef; a Dampierian region from Torres Strait to approximately Geraldton; a Flindersian region from southwest Australia and across the Great Australian Bight; the Maugean region around Tasmania; and a Peronian region along the east coast. Wilson and Allen (1987) reduced this to two major faunal regions, the tropical in the north and temperate in the south, with transition zones extending along the eastern and western coasts. On the other hand, a more recent bioregionalization scheme expands Whitley's scheme to 9 provinces along the continental shelf (IMCRA Technical Group 1998): splitting the Dampierian region into the Gulf of Carpentaria and Northwest Shelf provinces, and splitting the Flindersian province into three: a Western Australia, Southwest Australian and Great Australian Bight regions.

Some of the differences between authors may be semantic, i.e. the level of distinctness required to define a distinct bioregion, such as the number of distinct or endemic species or the range limits of faunal elements. Our community analysis separates regions, rather, on the basis of differences in the overall assemblages.

Our analysis supports several aspects of the IMCRA bioregionalization: in particular, the separation of the Gulf of Carpentaria and Northwest Shelf and the distinctness of the GAB from areas around Tasmania to the east and the area off southwestern and western Australian to the west. It also indicates that the Tasmanian area is distinct from areas to the north along the east coast of Australia. However, the areas off eastern and northeastern, western and southwestern Australia did not form distinct clusters, possibly due to insufficient sampling.

In the most comprehensively sampled areas, such as the GAB, the cluster analysis also indicates significant structure with depth in the fish assemblages. There have been relatively few fishery surveys around Australia sufficiently comprehensive and extensive to delineate the depth structure of Australian fish communities; exceptions are the surveys of the continental slope off western and southeastern Australia (May and Blaber 1989, Koslow et al. 1994, Williams et al. submitted). These findings highlight the need to explore and determine the structure of Australian fish communities in areas still not adequately explored, such as the continental slope around northern and eastern Australia and depths beyond about 1200 m in all parts of the Australian EEZ.

BENEFITS

The primary beneficiaries of the project are the scientists, managers and industry engaged in demersal fin fisheries around Australia, particularly in the Great Australian Bight, the Northwest Shelf and Gulf of Carpentaria. The study provides the most comprehensive data set extant for the relative abundance and distribution for commercially and ecologically important species, as well as overall community structure, for these areas. Data are also available for analysis of historical length frequencies and reproductive stages.

FURTHER DEVELOPMENT

The primary steps to be taken to further develop results of this project include:

• comparison of catch rates, length frequencies and maturity stage (spawning condition) data for key commercial species, particularly in well-sampled areas (e.g. the GAB, NW Shelf and Gulf of Carpentaria), between the Soviet and more recent data.

• comparison of species composition for bioregions between the Soviet and more recent trawl surveys, particularly to examine changes in the relative abundance of heavily exploited species (to examine the potential impacts of fishing and of species near the edge of their range (to examine the potential influences of climate change)

• design of future surveys of well-sampled regions should take account of the particular areas, gear, and protocols used by the Soviet surveys to obtain data that can be used to assess changes in length frequency, maturity stage, abundance and species composition

• complete the task of exploring and delineating the fish communities of Australia's EEZ; in particular through systematic surveys of slope fauna off eastern and northern Australia and of the waters deepeer than about 1200 m around all of Australia's EEZ.

CONCLUSION

The project was successful overall. The large Soviet historical fishery survey data set was computerized and transferred to Australia. This database comprises data from some 35 cruises carried out by Soviet research vessels to survey fisheries in Australian waters between 1964 and 1976, prior to declaration of the 200-mile EEZ. There were ~ 4000 demersal trawls carried out during this period with comparable gear. Most trawls were on the continental shelf, primarily in the Great Australian Bight, North West Shelf and Gulf of Carpentaria. There was very little coverage off eastern Australia and relatively little of western Australia. Overall, however, the Soviet fishery surveys represent the most comprehensive and important data set for Australian fisheries.

The taxonomy of the Australian fish fauna was poorly known at the time of the surveys, so species identifications were validated so far as was possible. Unfortunately much of the original fish collections had been lost. The reliability of all identifications was therefore evaluated on a regional basis, and only taxa with a high degree of certainty were used in subsequent analyses. It would not be appropriate to analyze the Soviet data for species diversity.

Overall catch rates were analyzed by region and depth. Catch rates were generally highest in the Great Australian Bight and Southwest Cape regions and lowest in the tropics. Catches on the shelf were higher on the shelf than over the slope, except off southeastern Australia, where relatively high catches were obtained over the upper slope. The Soviets did not explore depths greater than about 500 m.

The Soviet cruises consistently recorded catch composition by species, but neither weight nor numbers of individual species were consistently recorded, so most analyses were based on presence-absence data. The length frequency, maturity stages and diet of important commercial species were also recorded.

Nonhierarchical cluster analysis was used to examine community structure around the Australian continent. Ten distinct and robust clusters of stations were found, most of which corresponded to distinct geographic regions and depths and to a lesser degree with season. Results of the cluster analysis were largely consistent with a prior bioregionalization of the Australian marine fish fauna. Of particular note, confirmation was found for considering the Northwest Shelf and Gulf of Carpentaria to have distinct assemblages. Distinct assemblages were found characteristic of the fish fauna around Tasmania and the Great Australian Bight. On the other hand, distinct fauna communities were not observed off the west or east coasts of Australia, but this may be due to insufficient sampling.

Despite the limitations of the data, the data set provides a valuable historical benchmark for Australian fisheries, which can now be used to examine changes in Australian fish populations arising from anthropogenic impacts an d environmental variability. Analysis of the data contributes significantly to the classification of Australia's marine bioregions, the basis for planning a national system of representative marine protected areas. Information on the depth structure of Australian fish communities contributes to the management of species within regions. Large-scale distributional information contributes to understanding of stock and sub-species structure, which often conforms to the broad bioregionalization.

Literature cited

May, J.L. and S.J.M. Blaber. 1989. Benthic and pelagic fish biomass of the upper continental slope off Eastern Tasmania. Marine Biology 101: 11-25.

Interim Marine and Coastal Regionalisation for Australia Technical Group. 1998. Interim marine and coastal regionalization for Australia: an ecosystem-based classification for marine and coastal environments. Version 3.3. Environment Australia. Canberra. 103 pp.

Koslow, J.A., C.M. Bulman, and J.M. Lyle. 1994. The mid-slope demersal fish community off southeastern Australia. Deep-Sea Research 41: 113-141.

Whitley, G. 1937. Marine zoogeographical regions of Australia Australian Naturalist 8: 166-167.

Williams, A., J.A. Koslow, and P.R. Last. Submitted. The continental slope demersal fish fauna off the west coast of Australia (latitudes 20-35°S): biomass, diversity and community structure. Marine Ecology Progress Series.

Wilson, B.R. and G.R. Allen. 1987. Major components and distribution of marine fauna. In: Fauna of Australia, vol. 1A. Dyne, G.R. and D.W. Walton (eds). Australian Government Publishing Service. Canberra. Pp. 43-68.

APPENDIX 1

Intellectual property

No intellectual property resulted from this project. The database of Soviet historical fishesry data is available from the authors or from the FRDC.

APPENDIX 2

Staff

Dr. Tony Koslow, principal investigator (CSIRO Division of Marine Research)

Dr. Alan Williams, fisheries biologist (CSIRO Division of Marine Research)

Dr. Peter Last, fish taxonomist (CSIRO Division of Marine Research)

Mr. Aubrey Harris, database specialist (CSIRO Division of Marine Research)

Mr. Paul Sabourenkov, analyst, database specialist, translator (University of Tasmania)

Dr. Ernst Nossov, Russian fishery biologist (TINRO)

Ms. Yevgenia Nossov, data entry (TINRO)

APPENDIX 3

Data fields in the database

Summary of tables in the database

Table name	Set	Description
Caabcode	Auxiliary	Describes association between species names and corresponding CAAB codes.
Cisspp	Original	Contains the original names of species as recorded from the station logs.
Cruises	Additional	Describes cruises by vessel and dates in operation as recorded from the station logs. An upgraded version of the Vessels table.
Depthzon	Analysis	Describes depth zones as used in by-depth analysis.
Fishan	Original	Describes the number and type of measurements made for a particular species and trawl.
Fishbiol	Original	Describes fish biology measurements for a particular species and trawl.
Fishlen	Original	Describes fish length measurements for a particular species and trawl.
Scinames	Auxiliary	Contains CAAB database as used in this project.
Trdere	Analysis	Describes association between trawls, depth zones and bioregions.
Trwlby	Original	Contains trawl bycatch comments.
Trwlca	Original	Contains trawl catch information.
Trwllog	Original	Describes stations.
Trwlregn	Analysis	Describes association between trawls and bioregions, this table was produces by MapInfo.
Upgrspcl	Additional	An upgraded version of Cisspp table, associates original Russian species codes with corresponding CAAB codes and bioregions.
Vessels	Original	Describes cruises in the original comment format as recorded from the station logs.

Cisspp table fields.

Field Name	Field Type	Length / Dec	Description
NAME	char	30	Species name as recorded in trawl logs
SPCODE	char	8	Original Russian species code, key field
RECORDER	char	10	Name of the operator entering data for this specie
COMMENTS	char	60	General comments
SYNON	char	30	Other species name synonymous with the current one

Fisan table fields.

Field Name	Field Type	Length / Dec	Description
TRAWLID	char	12	Uniquely identifies the trawl these measureents were taken in, part of key
NAME	char	30	Species name as recorded in trawl logs, part of key
TAGS	int	5	Number of tags
LENGTHS	int	5	Number of length measurements taken
BIOLS	int	5	Number of biological measurements taken

Fishbiol table fields.

Field Name	Field Type	Length / Dec	Description
SPCODE	char	8	Russian species code, uniquely identifies a specie, key field
FISHID	char	5	Sequential number of an individual fish being measured
LT	float	6.2	Length (in cm)
WT	int	6	Wight (in g)
GONADWT	int	6	Gonad weight (in g)
SEX	char	1	Sex
STAGE	char	3	Maturity stage
FULLNESS	int	3	Stomach fullness
AGE	char	3	Age
COMMENTS	char	60	General comments
SEQUENCE	int	8	Used during data entry
TRAWLID	char	12	Uniquely identifies a trawl, key field
FAT	int	3	Unit unknown

Fishlen table fields.

Field Name	Field Type	Length / Dec	Description
SPCODE	char	8	Russian species code, uniquely identifies a specie, key field
LT	float	6.2	Length (in cm)
NOS	int	5	Number of individuals of this length
SEQUENCE	int	8	Used during data entry
TRAWLID	char	12	Uniquely identifies a trawl, key field
GROUPED	int	1	Purpose unknown

Trwlby table fields.

Field Name	Field Type	Length / Dec	Description
COMMENTS	char	240	Description of bycatch
TRAWLID	char	12	Uniquely identifies a trawl, key field
CHECKTRANS	int	1	Purpose unknown

Trwlca table fields.

Field Name	Field Type	Length / Dec	Description
SPCODE	char	8	Russian species code, uniquely identifies a specie, key field
NAME	char	30	Specie name, as in table Cisspp, superfluous as SPCODE is used to identify a specie
NO	float	8.2	Number of individuals of this specie caught
WT	float	8.2	Total weight caught of this specie (in kg)
TRAWLID	char	12	Uniquely identifies a trawl, key field
SEQUENCE	int	6	Used during data entry
COMMENTS	char	100	General comments

Trwllog table fields

Field Name	Field Type	Length / Dec	Description
VESSEL	char	25	Vessel name
THEDATE	date	8	The date of the trawl
STATION	int	4	Station number
SEA	char	30	Name of the sea
WIND	char	8	Direction of wind
SEASTATE	char	8	State of the sea
TEMP	float	4.1	Air temperature (in C)
TOTCATCH	float	8.2	Weight in total catch (in kg)
NORMTOT	float	16.2	Calculated field - normalised total catch weight by hour
BEGIN	int	2	Trawl start time - hours
BEGIN1	int	2	Trawl start time - minutes
END	int	2	Trawl end time - hours
END1	int	2	Trawl end time - minutes
DEPTH	int	5	Depth (in m) at the beginning of the trawl
BOTTOMTYPE	char	30	Seabed type at the beginning of the trawl
BOTTOMTEMP	float	4.1	Water temperature (in C) at trawl depth at the beginning of the trawl.
LAT	int	3	Latitude at the beginning of the trawl - degrees
LATMIN	int	2	Latitude at the beginning of the trawl - minutes
LATSEC	int	2	Latitude at the beginning of the trawl - seconds
LON	int	3	Longitude at the beginning of the trawl - degrees
LONMIN	int	2	Longitude at the beginning of the trawl - minutes
LONSEC	int	2	Longitude at the beginning of the trawl - seconds
TRAWLTYPE	char	15	Description of trawl type used
WIRELENGTH	int	5	Wire length
SPEED	int	5	Vessel speed during trawl (in knots)
DUR	int	3	Duration of the trawl (in minutes)
MESHSIZE	char	6	Trawl mesh size
EDEPTH	int	5	Depth (in m) at the end of the trawl
EBOTTOMTYP	char	30	Seabed type at the end of the trawl
EBOTTOMTEM	float	5.2	Water temperature (in C) at trawl depth at the end of the trawl
ELAT	int	3	Latitude at the end of the trawl - degrees
ELATMIN	int	2	Latitude at the end of the trawl - minutes
ELATSEC	int	2	Latitude at the end of the trawl - seconds
ELON	int	3	Longitude at the end of the trawl - degrees
ELONMIN	int	2	Longitude at the end of the trawl - minutes
ELONSEC	int	2	Longitude at the end of the trawl - seconds
COURSE	char	30	Description of the vessel course and method used to calculate it
COMMENTS	char	240	General comments
DEPTH1	int	5	Depth during trawl - entry one.
DEPTH2	int	5	Depth during trawl - entry two.
DEPTH3	int	5	Depth during trawl - entry three.
DEPTH4	int	5	Depth during trawl - entry four.
DEPTH5	int	5	Depth during trawl - entry five.
DEPTH6	int	5	Depth during trawl - entry six.
DAYID	char	12	Purpose unknown
SEQUENCE	int	7	Used during data entry
CHECKCATCH	char	1	Boolean (Y/N) field, purpose unknown
CHECKLENGT	char	1	Boolean (Y/N) field, purpose unknown
CHECKBIOL	char	1	Boolean (Y/N) field, purpose unknown
TRAWLID	char	12	Unique identifier for this trawl, key field

TRAWLSIZE	int	5	Purpose unknown
TRAWL	char	10	Purpose unknown

Field use (in number of records and percent) in Trwllog table.

Field Name	# records used	% records used
VESSEL	5629	100%
THEDATE	5629	100%
STATION	5629	100%
SEA	5483	97%
WIND	5441	97%
SEASTATE	4495	80%
TEMP	5181	92%
TOTCATCH	5182	92%
NORMTOT	5514	98%
BEGIN	5538	98%
BEGIN1	5538	98%
END	5538	98%
END1	5538	98%
DEPTH	5618	100%
BOTTOMTYPE	2649	47%
BOTTOMTEMP	432	8%
LAT	5629	100%
LATMIN	5629	100%
LATSEC	5629	100%
LON	5629	100%
LONMIN	5629	100%
LONSEC	5629	100%
TRAWLTYPE	0	0%
WIRELENGTH	5578	99%
SPEED	5584	99%
DUR	5628	100%
MESHSIZE	44	1%
EDEPTH	5568	99%
EBOTTOMTYP	2634	47%
EBOTTOMTEM	3848	68%
ELAT	5618	100%
ELATMIN	5618	100%
ELATSEC	5618	100%
ELON	5618	100%
ELONMIN	5618	100%
ELONSEC	5618	100%
COURSE	5604	100%
COMMENTS	598	11%
DEPTH1	2632	47%
DEPTH2	2485	44%
DEPTH3	2494	44%
DEPTH4	2180	39%
DEPTH5	2123	38%
DEPTH6	2141	38%
DAYID	0	0%
SEQUENCE	5629	100%
CHECKCATCH	15	0%
CHECKLENGTH	14	0%

CHECKBIOL	15	0%
TRAWLID	5629	100%
TRAWLSIZE	5401	96%
TRAWL	5621	100%

Vessels table fields

Field Name	Field Type	Length / Dec	Description
VESSEL	char	25	Name of the vessel
DETAILS	char	240	Description of the vessel and dates in operation, in text format as recorded in trawl logs

APPENDIX 4

Report on taxonomic validation

Objective: To assess the reliability of fish identifications in the TINRO fishery data from Australian waters for the purpose of upgrading the catch data

Dates: 14/4/97—7/5/97

Project: FRDC funded to Tony Koslow (CSIRO DMR)

Background

Fisheries data collected by the Soviet Far Seas Fisheries Laboratory (TINRO) during exploratory fishing in Australian waters in the 1960s and 1970s is now in a database at the CSIRO Division of Marine Research in Hobart. Some 1,700 nominal species of fish are named in the original catch composition lists but many of these are problematic. The identifications were made at a time when the Australian marine demersal fish fauna was relatively poorly known and documented in only a preliminary literature. Hence, many of the species names in the data set are out of date, valid species are recorded under a variety of synonyms or mis-spelled names, and other species do not occur in the region. Before analysing the TINRO data set it was necessary to evaluate the reliability of the original species identifications and upgrade them to the extent possible.

Voucher specimens collected during the TINRO cruises are held at the Russian Academy of Sciences Institute of Oceanography and State University Zoological Museum in Moscow, and the Russian Academy of Sciences Zoological Institute in St. Petersburg. These specimens were re-identified and the information used as a guide to estimate the reliability of identifications for the data set and enable computer coding of species (using the revised CAAB list) for analysis.

Method

We spent most time at the State University Zoological Museum (Moscow) and the Zoological Institute (St. Petersburg) as their collections held most of the TINRO material and were easiest to access. At both institutions the collections were well organised on the shelves and well documented. We examined only registered material with original identifications. (An unregistered collection was housed at Zoological Institute but its location was unknown and was apparently dominated by pelagic species.)

We took with us several key taxonomic guides to the Australian fish fauna which were not available at the Russian museums. The most important of these were the 5-volume CSIRO field guide to North West Shelf fishes which provides the best available coverage of Australia's tropical shelf fauna. Also, the regional references available at the time the TINRO cruises were undertaken: Fishes of Southern Australian- Scott et al. (1966), and the Handbook of Australian fishes– Munro (1966-).

Priority was given to the dominant species based on their frequency of occurrence in the data set. However, specimens of some key species and several important families were not represented in the collections. For each species examined, we recorded the original and upgraded identifications, the vessel, region and year of capture, and scored the reliability of the original identification based on a set of criteria (Table 1a). In total, 106 and 113 lots were examined in Moscow and St. Petersburg collections, respectively (Table 2).

Using this assessment, a reliability score on a scale of 0 –10 was then estimated for the identification of each species in the data set based on the criteria in Table 1b. As part of this process, junior synonyms were upgraded, mis-spelled names corrected, and spurious and pelagic species flagged; CAAB codes were also added.

We upgraded the species identifications on a region by region basis using the eight disjunct biogeographic provinces identified for Australian fishes by IMCRA (1998) (Fig. 1). This was seen as an important step in the methodology because the reliability levels for re-identified species were not uniform geographically (more reliable for temperate than tropical faunas). Some species were accurately identified within one bioregion but misidentified in others. Similarly, the accuracy of identifications within regions depended on the identification sources available at the time, and this varied between regions, and temporally by cruise. Hence, we were able to gain an estimate of accuracy levels within regions using this method.

Based on the material we re-identified, the reliability of original identifications appeared to be most dependent on region. Reliability was higher in temperate than tropical Australia, with the GAB/SW Australian region most reliable and the NW Shelf least reliable (Table 3). Among cruises, those of the *Akedemia Berg* and *Dimitriy Mendelev* were most reliable— but these were in temperate regions (Table 4). There was no clear trend in reliability with year, probably because region and/or vessel was more important (Table 5).

Species were ranked by their frequency of occurrence in each bioregion and tackled in priority order. Typically, there were many infrequently-occurring species in each region which formed a long 'tail' at the end of each regional list. Because these species were usually more difficult to identify, and are likely to have low utility in analyses, they were not considered in detail. They were assigned reliability scores using the total list of species, ie without reference to their region of occurrence. Because of this, their inclusion in analyses should be with due care. This may be by combining them at supraspecific levels but this may not be applicable. Similarly, use of infrequently-occurring species in estimating total species numbers by region should be done with great caution .

Data reduction and data use

Compartmentalisation of the data into regional faunas provides a natural structure for subsequent analyses. Within bioregions, the reliability scores and species codes in the upgraded species lists can be further used to reduce the entire data set to subsets of species at given levels of reliability, and/ or taxonomic levels. Data reduction can also be by using only a certain number of species, or by using only those species making up a certain percentage of total occurrences. We suggest the following ways to reduce the data and to explore for appropriate cut-off points:

use only species with Level 1 and 2 reliability (based on Table 1b)

use a subset of species from each bioregion (based on a percentage of total number of species by using standard methodology but including species stepwise based on diminishing highest reliability's, ie from 10–0).

examine trends in reliability in each region by plotting the mean reliability of sequential blocks of ten species in the ranked species lists to assess cut-off points

consider the reliability's of different vessels and years where multiple cruises occur in a single region

consider the utility of supraspecific groupings which can be easily derived using CAAB codes. For example, grouping species to the level of genus will eliminate sister-species uncertainties; higher-level groupings may relate to ecomorphotypes or ecological characteristics

Although many highly-ranked species are identified at a relatively high level of reliability, some are not. It might be possible to improve their reliability scores by examining their distribution with respect to depth, latitude or year of identification. For example, previously unrecognised sister-species might be separated in different trawl samples on the basis of non-overlapping distributions by depth. Species in this category are flagged in the database with a symbol (D-depth, R-region, C-chronology). However, we suggest this is done only in particularly important instances due to the time and effort that will be involved.

Work required to complete the upgrade of the data set

Because neither of the computers taken on the trip worked reliably on the Russian power supply the entry of reliability scores and CAAB codes was done on paper copies. These data need to be punched and the data set checked prior to analysis. The following steps are required:

1. Add CAAB codes to data set

2. Add reliability scores to data set

3. Generate upgraded lists of species by bioregion ranked by number of occurrences and check for missing CAAB codes or reliability scores

4. Explore data reduction options

Concluding remarks

The museum survey and reliability assessment was successful as we were able to examine a large number of specimens (more than expected), and gained an excellent insight into the quality of the various subsets of the data. Major errors in identification in some bioregions vindicated this assessment, but, conversely, meaningful data sets are now known to exist for other bioregions. Without this insight, uninformed analysis of the general data set would have been misleading.

Alan Williams & Peter Last, 13 May 1997

APPENDIX 5

List of species and reliabiliity