Science Policy

NIWA

Taihoro Nukurangi

Estimate of the absolute abundance of black and smooth oreo in OEO 3A and 4 on the Chatham Rise

I.J. Doonan, R.F. Coombs, P.J. McMillan, A. Dunn

Final Research Report for Ministry of Fisheries Research Project OEO9701

National Institute of Water and Atmospheric Research

October 1998

Final Research Report

Report Title	Estimate of the absolute abundance of black and smooth oreo in OEO 3A and 4 on the Chatham Rise
Author	I.J. Doonan, R.F. Coombs, P.J. McMillan, A. Dunn
1. Date	22 September 1998
2. Contractor	NIWA
3. Project title	Estimate of the absolute abundance of black and smooth oreo in OEO 3A and 4 on the Chatham Rise
4. Project code	OEO9701
5. Project leader	Roger Coombs

6. Duration of project

Start Date:	1 October 1997
Completion date:	30 September 1998

Attached are two reports that comprise the Final Research Report for this project. The first report covers the bulk of the work whilst the second deals with the 'objective' approach to target identification.

7. Executive Summary

Part 1: Acoustic estimation of oreo biomass in areas OEO 3A and 4

I.J. Doonan, R.F. Coombs, and P. J. McMillan

A successful acoustic survey of area OEO 3A was carried out between 10 November and 19 December 1997 on *Tangaroa* (voyage TAN9713). An adaptive, 2-phase stratified random approach was used. A partial survey of area OEO 4 was also carried out.

In both areas acoustic data were collected concurrently on both towed and hull mounted transducers. The OEO 3A survey covered 74 transects on the 'flat' and 34 on hills. A total of 51 trawls was carried out for target identification and to estimate target strength and species composition. The OEO 4 survey included 8 transects on the 'flat' and 14 on hills. A total of 11 trawls was carried out.

In situ and swimbladder target strength data were collected and these have yielded new estimates of target strength for both black and smooth oreos.

For OEO 3A the estimated biomass of recruited smooth oreos on the flat is 15 400 t with a *c.v.* of 46% and 95% confidence interval of 6–33 000 t. For recruited black oreos the estimated biomass is 18 800 t with a *c.v.* of 44% and 95% confidence interval of 6–39 000 t. The *c.v.* in both cases includes a subjective allowance for target strength variance. Estimated biomass of smooth oreos on hills is 2 300 t with a *c.v.* of 34% and 95% confidence interval of 730–3 900 t.

For OEO 4 the estimated biomass of recruited smooth oreos on the flat (stratum 12) is 172 400 t with a *c.v.* of 77%. For recruited black oreos the estimated biomass is 2670 t with a *c.v.* of 68%.

Part 2: Oreo mark identification

A. Dunn

Data collected by *Tangaroa* on an acoustic survey of black and smooth oreos in 1997 from oreo management area 3A on the Chatham Rise (TAN9713) are analysed using classification and predictive statistical methods in order to formulate decision rules on the species composition of acoustic marks. The acoustic marks used in this analysis are those collected during the survey where the species composition was estimated from an associated trawl.

We discuss a range of supervised statistical methodologies including discriminant analysis methods (including linear and flexible discriminant analyses) and classification trees, and apply these to the available data. The discriminant analysis results suggest that a decision rule based on gross parameters of acoustic marks may be useful in predicting species composition. The classification tree is less convincing. However, all methods are hampered by a lack of data.

Cross validation techniques cannot be carried out due to the small number of data points and hence a comparison of the predictive properties of these methods is unavailable. In addition, some research on an appropriate cost or risk function is required to allow such a comparison.

8. Objectives

Programme Objective:

To estimate the absolute abundance of black oreo and smooth oreo in OEO3A and 4 on the Chatham Rise.

Objective for 1997/1998:

To estimate the absolute abundance, with a target coefficient of variation of the estimate of 20–30%, of black oreo and smooth oreo in OEO3A and 4 on the Chatham Rise.

Part 1: Acoustic estimation of oreo biomass in areas OEO 3A and 4

I.J. Doonan, R.F. Coombs and P.J. McMillan

1. INTRODUCTION

The south and east Chatham Rise (OEO 3A and OEO 4) is the main oreo fishing area in the New Zealand EEZ, with reported landings of 14 520 t in 1995/96 compared to the EEZ wide catch of 23 572 t (Annala & Sullivan 1997). There is also a substantial orange roughy fishery in the area with reported landings of 1 400 t. Oreos from under-sea hills have made up an increasing proportion of the total oreo catch in recent years.

In the past, oreo relative biomass has been estimated using catch per unit fishing effort (CPUE) and trawl survey methods. Because of problems with these, particularly the difficulty of deriving absolute abundance from them, alternative ways of estimating absolute biomass were assessed by the Deepwater Fishery Assessment Working Group who considered that acoustic techniques offered the possibility of fishery independent, absolute abundance estimates covering both areas of flat and undulating sea-bed ('flat') and under-sea hills (Annala & Sullivan 1997). A first evaluation of the approach was made using the Simrad EK500 in October-November 1995 (voyage TAN9511). A trial survey on the flat, intended primarily for target strength, target identification and acoustic equipment development was carried out in April 1997 (voyage TAN9705). The first full survey, which is described here, was carried out between 10 November and 19 December 1997 on *Tangaroa* (voyage TAN9713).

2. METHODS

2.1 Biomass estimation

The following description deals with the estimation of smooth oreo biomass. The same procedure is applicable to the estimation of black oreo biomass.

The acoustic data were classified into types of 'marks' as described in the section on target identification. For a hill or a stratum, i, the biomass of smooth oreo in mark-type, m, is given by:

$$B_{i,m} = \frac{\text{abscf}_{i,m}}{\overline{\overline{\sigma}}_{bs,m}} \times p_{sso,m} \times \text{area}_i \times \overline{w_m}$$

where area_i is the area of the hill or stratum, $\text{abscf}_{i,m}$ is the mean back scattering in units of fish m^{-2}, $\overline{\overline{\sigma}}_{bs,m}$ is the mean tilt-averaged acoustic cross-section for the species mix, $p_{sso,m}$ is the proportion of smooth oreo, and $\overline{w_m}$ is the mean weight of a smooth oreo.

$$\overline{\overline{\sigma}}_{bs,m} = \sum_{j}^{\text{species}} p_{jm} \overline{\overline{\sigma}}_{bs,jm}$$

where j indexes species, p_{jm} is the proportion in numbers of species j in the mix, and $\overline{\overline{\sigma}}_{bs,jm}$ is the mean tilt-averaged cross-section for species j (which depends on its length distribution in mark-type m).

Mean cross-section, $\overline{\overline{\sigma}}_{bs,jm}$, is given by $\sum_l f_{sso,m,l} 10^{\frac{\langle\text{TS}\rangle_{sso}(l)}{10}}$ for smooth oreo and by $\sum_l f_{j,m,l} 10^{\frac{\langle\text{TS}\rangle_j(Ljm)}{10}}$ for other species, where $f_{sso,m,l}$ is the fraction of smooth oreo in mark-type m with length l, $\langle\text{TS}\rangle_j(l)$ is the tilt-averaged or *in situ* target strength-to-length function for species j, L_{jm} is the mean length of species j in mark-type m, and $\langle\text{TS}\rangle_j(l) = a_j + b_j \times \log_{10} l$.

The tilt-averaged acoustic cross-section is given by

$$\overline{\sigma}_{bs} = \int \sigma_{bs}(\theta) g(\theta) d\theta$$

where θ is the tilt angle (in the pitch plane only), $\sigma_{bs}(\theta)$ is the acoustic cross-section as a function of θ and $g(\theta)$ is the probability of a fish being at an angle θ. Tilt-averaged target strength, $\langle\text{TS}\rangle$, is given by $10\log_{10}\overline{\sigma}_{bs}$.

The lengths, mean weights, species composition and proportion of smooth oreo in the population were obtained by trawling during the survey.

For several strata (*strata*) and mark-types (*marks*) the total biomass is given by

$$\frac{1}{f_{area}} \sum_{i}^{strata} \sum_{m}^{marks} B_{i,m}$$

where f_{area} is the fraction of recruited fish in the survey area. $B_{i,m}$ refers to recruits so two $p_{sso,m}$ must be estimated: one for recruits and another for pre-recruits.

f_{area} has been calculated from the ratio of commercial catches within the survey area to those for the management area as a whole.

2.2 Survey design

2.2.1 Flat

The survey on flat areas was a two-phase stratified random design (Francis 1984, Jolly and Hampton 1990). In each stratum, randomly positioned acoustic transects were defined in the north-south direction. Recruited fish were assumed to be in schools and randomly chosen schools in each stratum were sampled by trawl to obtain the length frequencies and proportions of smooth oreo, black oreo and other species.

In allocating trawl tows and acoustic transects to strata, for the first phase, three sources of variation were considered:

- sampling error in the acoustic data
- sampling error in the proportions of both oreo species in the species mix
- experimental error in the determination of the target strength of both oreos.

Estimates of all of these were made using existing trawl survey data, acoustic data collected with the Simrad EK500 and orange roughy target strength data (McClatchie *et al.* in press). Transects and trawl tows were then allocated to give an expected biomass coefficient of variation (*c.v.*) of <30%.

Second phase transects were allocated from an analysis of the first phase results.

2.2.2 Hills

The hill survey was a mixed design. One hill (Neil's Pinny) was selected explicitly and two hills (Neil's Condom and Hill A) were selected at random from a complex of eight hills near 176° E. The list of hills was compiled from positions provided by the Fishing Industry, catch statistics and hills encountered during this and previous surveys.

The approach to surveying hills was to use systematically allocated transects either as evenly spaced parallel tracks, a grid or a 'star-burst' pattern. For the estimates presented here, these transects have been treated as simple random samples.

2.3 Acoustic system

The acoustic data were collected with NIWA's Computerised Research Echo Sounder Technology (*CREST*). *CREST* is computer based, using the concept of a 'software echo sounder'. It supports multi-channels, each channel consisting of at least a receiver and usually also a transmitter. The receiver has a broad-band, wide dynamic range pre-amplifier and serial analog-to-digital converters (ADCs) which feed a digital signal processor (DSP56002). The ADCs have a conversion rate of 100 kHz and the data from these are complex (quadrature) demodulated, filtered and decimated. The filter was a 100 tap, linear-phase finite impulse response digital filter. For the biomass survey work this had a bandwidth of 1.5 kHz and the data were decimated to 4 kHz. For target strength work the bandwidth was 3.5 kHz and the decimated frequency 10 kHz. Following decimation, for surveys a 20 Log R time-varied gain was applied and for target strength 40 Log R. In both cases the results were shifted to give 16 bit resolution in both the real and imaginary terms and the complex data were stored for later processing.

The transmitter is a switching type with a nominal power output of 2 kW rms. For all the work described here the operating frequency was 38 kHz. For surveys the transmitted pulse length was 1 ms (38 cycles) and the effective pulse length 0.78 ms. For target strength work it was 0.32 ms (12 cycles). Time between transmits was 4 s in all cases.

Three *CREST* systems were used for the survey:

1. A single channel system connected to a hull-mounted Simrad model 38–7 transducer.
2. A single channel system connected to a towed Edo model 6978 single beam transducer via 1 km of Rochester type A301301 tow cable.
3. A four channel towed system, with underwater electronics, connected to a Simrad type ES38DD split beam transducer.

In systems 1 and 2 the receiver and transmitter were mounted on the vessel. Thus for system 2 the tow cable forms part of the calibrated circuit. In system 3 the receiver and transmitter were mounted in the towed body. The same type of flat-nosed, torpedo-shaped, 3 m long 'heavy weight' towed body was used for both towed systems (2 and 3).

The digital data from the receiver are sent, in all cases, to a control computer where they are combined with position and transect information and stored. In the underwater system (3 above) the data are transmitted via the tow cable to the control computer on the towing vessel. In this system all four transducer quadrants (beams) are energised simultaneously from a single transmitter but on receive the system operates as four semi-independent echosounders. Data are processed independently on the four channels but operation is tightly synchronised by the transmit key and by using a common clock for all the ADCs. In subsequent analysis of survey data the four channels are summed to form a single beam. However, for target strength, the beams are treated separately to reject multiple echoes and calculate the position of the echoes in the beam.

Calibration data for biomass estimation for all three systems are shown in Table 1(a) and for target strength estimation using the split-beam system in Table 1(b).

Table 1: Calibration data for the three systems used (a) for the biomass survey and (b) for target strength estimation. G_r is the gain of the system at the reference range *r*.

(a)

System	**1**	**2**	**3**
Transducer model	Simrad 38-7	Edo 6978	Simrad ES38DD
Transducer serial no.	23421	102	28326
Nominal 3dB beamwidth (°)	7.3	6.5	6.9
Effective beam angle (sr)	0.0087	0.00864	0.0081
Operating frequency (kHz)	38	38	38
Transmit interval (s)	4	4	4
Nominal pulse length (ms)	1	1	1
Effective pulse length (ms)	0.78	0.78	0.78
Filter bandwidth (kHz)	1.5	1.5	1.5
Initial sample rate (kHz)	100	100	100
Decimated sample rate (kHz)	4	4	4
TVG	20 Log R	20 Log R	20 Log R
SL+SRT (dB)	51.0	40.2	60.7
Reference range, *r* (m)	28	28	28
$20 \log_{10} G_r$	120.58	120.76	111.56

(b)

System	**3**
Transducer model	Simrad ES38DD
Transducer serial no.	28326
Nominal 3dB beamwidth (°)	6.9
Effective beam angle (sr)	0.0081
Operating frequency (kHz)	38
Transmit interval (s)	4
Pulse length (ms)	0.32
Filter bandwidth (kHz)	4
Initial sample rate (kHz)	100
Decimated sample rate (kHz)	10
TVG	40 Log R
SL+SRT (dB)	60.7
Reference range, *r* (m)	28
$20 \log_{10} G_r$	111.56

2.4 Target strength

Target strength for both black and smooth oreos was estimated from swimbladder casts and *in situ*. Swimbladder casts were made using the technique described in McClatchie *et al.* (1996) and target strength was estimated by modelling (McClatchie *et al.* 1996).

2.4.1 In situ data collection

To collect *in situ* data, marks that were expected to be black or smooth oreos were located and the *CREST* system deployed at about 2 knots, 30–70 m above the schools, usually for about an hour. The school was then trawled to identify the species and estimate size composition.

2.4.2 In situ data processing

The recorded complex data preserves both amplitude and phase information and allows both target position and amplitude to be calculated. To estimate target strength it is first necessary to filter out all echoes that do not originate from a single fish. To do this we have checked the following echo characteristics and associated conditions:

- The width of the echo was between 78% and 180% of the transmit pulse width at half the maximum echo amplitude (the 6 dB amplitude points).
- The standard deviation of the electrical phase of the echo, between the 6 dB amplitude points, was < 0.196 rad on both the individual and combined beams.
- The width of the four individual echoes at the 6 dB amplitude points varied by < 63% of the transmit pulse width.
- The echo peak was more than 0.375 m in range from other echoes.
- The mean and standard deviation of the difference between the echo amplitude on beam 1 and the same echo on beams 2, 3 and 4 were < 1.5 and < 3 dB respectively for all three comparisons.

These characteristics are based on those listed by Soule *et al.* 1995, Soule *et al.* 1997 and Soule *pers. com.* They were used to filter the data to reject all echoes from more than one fish. The values of these characteristics were set by recording data from two spheres at constant angles in the beam but at a range of different distances, in the large tank (8 x 4 x 4 m) at Greta Point. The experimental approach was similar to that of Soule *et al.* 1997.

After filtering, the positions of the echoes qualifying as single echoes were calculated (Ehrenberg, 1979) and the amplitudes corrected accordingly.

2.4.3 Associated species

Target strength for associated species was based on Foote's (1987) equation, adjusted to incorporate swimbladder data for rattails.

Population target strengths were estimated by applying the length-target strength relationships to the length frequency data collected during trawl sampling.

2.5 Target identification

2.5.1 Mark classification

One of the reasons for believing that acoustic surveys were viable was the observation during the years of trawl surveys that oreos of both species formed schools and aggregations. The schools appeared as distinctive 'marks' on echograms. The trial survey carried out in 1995 was based on EK500 echograms. The experience gained with this was used to design this survey. The various types of marks seen were classified in the first instance using EK500 echograms associated with trawl catches. The following characteristics were used to classify marks:

- Density of backscatter in terms of a colour code (in ascending order of backscatter intensity: dots, blocks of colour of light blue, blue, purple and green)
- Thickness of the mark
- Whether the mark was on the flat or the top edge of a drop-off or mound
- Whether the mark was a plume or a layer
- The depth of the mark.

During the voyage some acoustic data were also collected during trawls using the hull *CREST* system. These data were used to assist in assigning marks to the above categories. *CREST* towed system data collected during acoustic transects were also used to further specify the characteristics.

Trawls were grouped into the following catch categories according to the percentage (by weight) of the main species in the catch:

- Black oreos.
- Smooth oreos.
- Mixtures of oreos ('OREO').
- Black oreo and others.
- Others.

'Others' are species other than black and smooth oreo.

Although we have made use of all the types of echogram 'image' data available, the approach still relies on human pattern recognition and includes a subjective element, particularly since there is not usually likely to be exact correspondence between the echogram marks and the trawl catch.

2.5.2 Underwater camera

During the voyage a Benthos underwater camera was used to take still photographs in fish marks as an additional aid to target identification. The camera was dropped into suitable marks and set to take photographs at 1 minute intervals.

2.5.3 'Objective' mark identification analysis

A separate report (attached) has been prepared for this section.

2.6 Estimating variance and bias

Sources of variance are:

- the sampling error in mean backscatter
- the proportion of oreos in schools as opposed to the background
- sampling error in catches (affects the estimates of p_{sso} and p_{boe})
- variance in the estimate of oreo target strengths, $\langle TS \rangle_{sso}$ and $\langle TS \rangle_{boe}$
- error in the target strengths of other species in the mix (high)

These sources of variation were combined using simple bootstrapping as follows:

- For acoustic sampling, acoustic transects were re-sampled from those within a stratum
- For trawl sampling, the school stations were re-sampled from those within each mark-type
- For target strength of oreos, $\langle TS \rangle_{sso}$ and $\langle TS \rangle_{boe}$, the TS-length data were re-sampled and the TS-length regression re-estimated.
- For target strength of other species, the TS for each species was re-sampled in 3 steps. First, the intercept in the TS-at-length relationship was randomly shifted to the constant for the individual relationship of one of the component species in Foote's data (only the intercept was included because the slope was constant at 20). This intercept remained the same for each species in a particular bootstrap run. For each species, a difference to this mean was randomly chosen from those for Foote's component species. The TS was selected from the distribution of TS-at-length (assuming this distribution to be normal with a mean equal to the above chosen constant and difference to the mean, and a standard deviation equal to the residual standard error, 1.47 dB), from the TS-length regression .

Potential sources of bias are:

- classification of marks
- differences in relative catchability of other species compared to oreos
- the species composition and species distribution in the background layer
- the proportion of oreos in the shadow zone
- the validity of the target strength-length relationship used for estimating the target strength of associated species.
- signal loss from transducer motion
- signal loss from bubbles (for the hull transducer)
- uncertainty about absorption of sound in water (particularly for the hull transducer)
- fish movements (oreos moving to and from schools on both hills and flat to the background population).

The effects of these were investigated by sensitivity analysis except for the shadow zone for which a special analysis was done. This is described in the following section.

2.7 Shadow zone

The shape of the sound beam projected by a transducer is quite complex with a main lobe and several side lobes (*see* Urick 1983). In echo-integration an equivalent, simple conical beam is assumed. The response of the transducer is uniform inside this cone and zero outside. For the Simrad ES38DD transducer used in the survey, the half-angle of the idealised beam is 2.91° (i.e., a solid angle of 0.0081 steradians). The other transducer used (Edo 6978) has a similar idealised beam (Table 1).

If an idealised sound beam is projected downwards at right angles to a flat bottom, the spherical wave-front of the beam first reaches the bottom at its centre. It returns an echo from the bottom that obscures any echoes from fish that are to one side of the centre of the beam. The part of the beam volume where fish echoes are obscured is called the shadow zone or dead zone.

The presence of the shadow zone results in fish close to the bottom being under-sampled by the acoustic beam. For an ideal beam and a uniform density of fish, a correction for this under-sampling can be calculated (Ona and Mitson 1996):

$$\text{correction} = \frac{B_{0-m} H_{eq}}{m}$$

where $B_{0\text{-}m}$ is the backscatter in the layer from the bottom up to m metres and H_{eq} is the equivalent height of a conical layer in the idealised beam that has the same volume as the shadow zone. H_{eq} is calculated as the ratio of the volumes, V_{eq}/V_{ref}, where V_{ref} is the volume of the 1 metre layer in the conical beam that bounds the upper extent of the shadow zone, and V_{eq} is the volume of the shadow zone (Ona and Mitson 1996). For an idealised beam with a half-angle A at a range R, these volumes are:

$$V_{ref} = \frac{2\pi}{3}[R^3 - (R-1)^3][1-\cos(A)]$$

and

$$V_{eq} = \frac{2\pi R^3}{3}\left[\frac{\tan(A)}{2} - 1 + \cos(A)\right]$$

For the Simrad ES38DD transducer on a flat bottom, H_{eq} is 0.35 m at 500 m range and 0.63 m at 900 m range. The formula for H_{eq} on a sloping bottom is not available so an approximation is used here: $H_{eq} = E/2$, where E is the greatest vertical distance from the outer edge of the idealised beam to the bottom for the wave-front that just touches the bottom. Its formula is given in Cordue (1996). Using this approximation, H_{eq} for the ES38DD transducer on a flat bottom is 0.35 at 500 m range and 0.62 at 900 m range (i.e., the approximation is good for flat bottom).

Kloser (1996) used an operational definition of H_{eq} which can be calculated from the bottom echo and this appears to be equivalent to E. No justification was given. Using $H_{eq} = E$ would double the shadow zone correction to the backscatter.

We apply an adaptation of Ona and Mitson 's (1996) correction by fitting a linear regression line to the backscatter of the lowest m one-metre layers and use the regression line to predict the backscatter in each meter layer in the shadow zone.

To spread the samples amongst the different sizes of marks and the bottoms they occurred on, samples were selected from categories that were based on the slope of the bottom (subjectively), their height, and their intensity (again, subjectively) (Table 2).

About a quarter of the marks in the OREO mark-type were sampled and estimates of the backscatter in their shadow zones made.

Table 2: OREO mark-type, categories, number of marks in each and sample sizes for shadow zone analysis

	Category definition			
Code	**Slope**	**Height, or intensity of mark**	**Number**	**Sample size**
PF	Flat	>50m	0	0
PS	Sloping	>50m	3	1
F50	Flat	<50m	2	1
S50	Sloping	<50m	7	2
DF	Flat	Diffuse	23	4
DS	Sloping	Diffuse	4	2
OF		Mark is well off the bottom	8	2

2.8 Hydrography

Drops were made with a Guildline conductivity-temperature-depth (CTD) probe to collect temperature and salinity profile data, allowing calculation of sound absorption and sound speed.

3. RESULTS

3.1 Acoustic survey

The survey was carried out on *Tangaroa* between 10 November and 19 December 1997 (voyage TAN9713). Data were recorded simultaneously with both towed and hull mounted transducers throughout most of the survey. However, there was a short period when only the hull system was used and for about half the time, poor weather meant that the hull data were of limited quality.

EK500 echogram data were collected for all but 4 trawls. For the first half of the voyage the *CREST* hull system was not synchronised with the two ship's sounders (EK500 and Kaijo Denki KMC 2000) used for fishing operations. This was remedied for the second half and *CREST* data were collected for 18 trawls.

System 2 (Edo transducer) was used as the primary system for surveying flat ground and as a backup to system 3 (underwater) for hills. For flat surveys the transducer was towed at about 300 m deep at about 4 $m.s^{-1}$. The primary system for hill surveys was system 3 with system 2 as backup. For hill surveys we aimed to fly the towed body about 100 m above the top of the hill and the speed needed to achieve this was typically less than 1.5 $m.s^{-1}$.

The acoustic transects surveyed in area OEO 3A are listed in Table 3 and the flat transects are plotted in Figure 1.

In the hill surveys, both Neil's Condom (NC) and Neil's Pinny (NP) were visited twice (two snapshots) and surveyed 3 times. In both cases one survey used E-W parallel transects and one N-S parallel transects. The third survey on NC was a star pattern and on NP a set of NW-SE parallel transects. Hill A (HA) was visited only once and a star pattern was used.

The acoustic transects carried out in OEO 4 are listed in Table 4 and plotted in Figure 2.

Table 3: Summary of acoustic transects surveyed in area OEO 3A by stratum and snapshot. On the hills, NC is Neil's Condom, NP Neil's Pinny and HA Hill A

(a) **Flat**

Snapshot	Stratum	Transects	
		Hull	Towed
1	1	1–5	1–5
1	2	1–20	1–20
2	2	1–7	3–7
3	2	2	2
1	3	1–12	1–12
2	3	1–8	6–8
3	3	2–5	2–5
1	4	1–5	1–5
1	5	1–3	1–3
1	15	1–5	1
2	15	1–4	1–4

(b) **Hills**

Snapshot	Stratum	Transects	
		Hull	Towed
1	NC	1–12	1–12
2	NC	1–4	1–4
1	NP	1–7	1–7
2	NP	1–7	1–7
1	HA	1–4	1–4

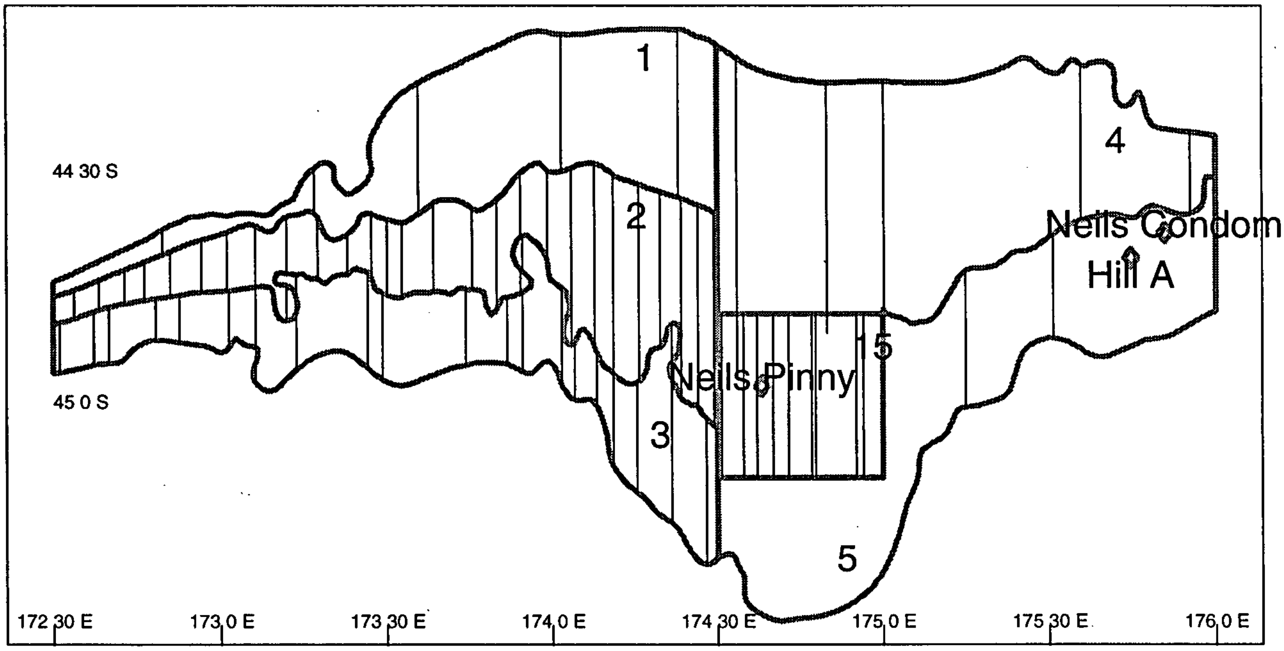

Figure 1: Survey areas for Chatham Rise acoustic survey showing strata and transects on the flat and the positions of the hills surveyed in area OEO 3A.

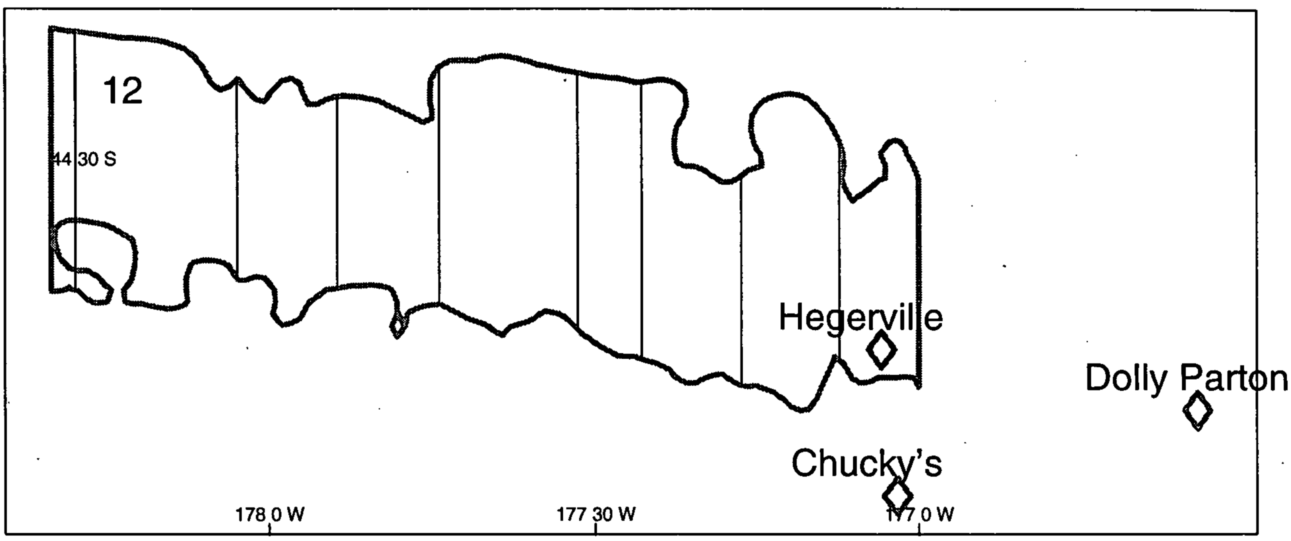

Figure 2: Survey area for Chatham Rise acoustic survey showing strata and transects on the flat and the positions of the hills surveyed in area OEO 4.

3.2 Trawling

A total of 51 trawls were carried out of which 6 were randomly chosen 'background' trawls, 8 were target trawls associated with target strength work and the rest targeted trawls for species identification. The positions of the trawls are plotted in Figure 3.

Table 4: Number of acoustic transects surveyed in area OEO 4

	Stratum	Transects	
		Hull	Towed
Flat	12	8	7
Hills	Dolly Parton	5	1
	Chucky's	4	1
	Hegerville	5	2

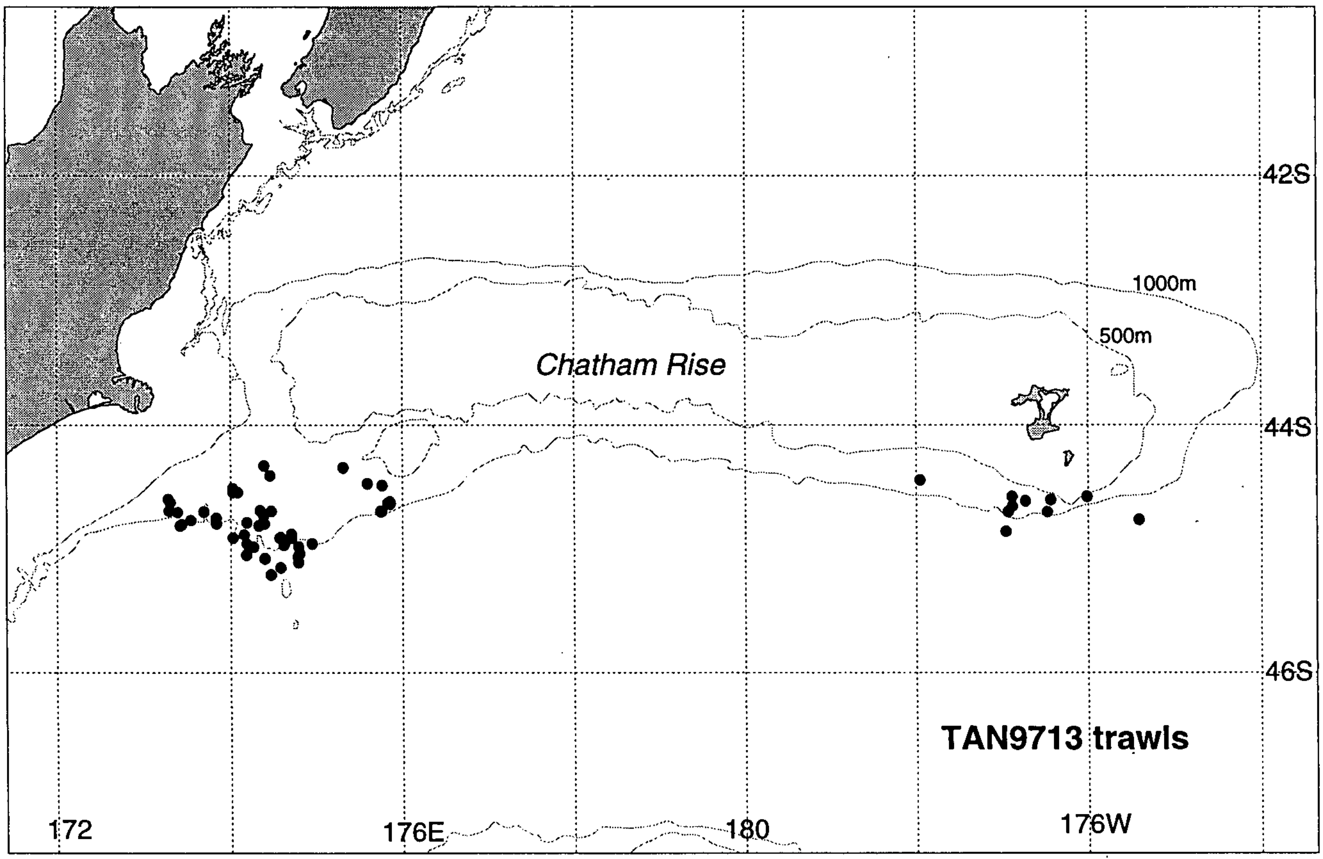

Figure 3: Trawl stations carried out during the Chatham Rise acoustic survey in areas OEO 3A and OEO 4, hills and flat. Area OEO 3A trawls form the western group of points.

3.3 Target strength

3.3.1 Swimbladder modelling

Swimbladder casts for both species were made during 2 voyages in the 1996/97 fishing year (TAN9705 and TAN9708). Black oreo swimbladders proved susceptible to damage and it was difficult to make adequate casts from them. As a result only a small sample was collected. In the absence of any tilt angle data for oreos, a typical distribution (mean 0° and standard deviation 15°, McClatchie *et al.* 1996) was used to estimate target strength. The estimated relationships are shown in Figure 7.

3.3.2 In situ estimation

Target strength runs were made over 10 fish schools but for 4 of these the associated catches were small (< 0.5 t) with a mixture of species. The remaining 6 were predominantly oreos and are summarised in Table 5.

Table 5: Smooth and black oreo schools used for target strength estimation. The letters identifying the schools match with those in Table 6 and Figures 4–6

School	Total catch (kg)	Smooth %	Black %	Depth (m)
A	2 084	84	11	990
B	1 976	94	4	1 000
C	3 295	98	0	1 240
D	749	76	20	1 030
E	665	0	75	700
F	1 526	1	93	850

Flying the towed body close (30–70 m) to black oreos produced no discernable avoidance behaviour. In the case of smooth oreos there was a slight reaction at the closer range. Target strengths have been estimated from the schools in Table 5 by matching modes in the target strength and total length frequency distributions. The modes were matched subjectively, by eye and allocated on the basis that black oreo target strength is higher than smooth (from the swimbladder modelling results). Where possible all modes have been matched up but in the two mixed schools the higher target strength mode in A and the lower mode in B are mixtures of black oreo and smooth oreo and have not been included.

Some of the target strength distributions were hard to interpret, particularly that for school C. Although the catch was a very clean one, the target strength data suggest that there were large numbers of very small scatterers not being sampled by the trawling. Figure 4 shows the distribution with relaxed (achieved by omitting some of the criteria listed in section 2.4.2) and more strict multiple target rejection. The more strict criteria (Figure 4, 2) have removed most of the mode at about -75 dB and changed the balance of the other modes. The balance changes even further with the even stricter criteria used for the analysis (Figure 5 C). Matching the obvious modes here gives a higher target strength than for the other schools.

The distributions for smooth oreos are shown in Figure 5 and black oreos in Figure 6. School D has been used for both smooth oreos and large black oreos. The target strength and length modes are listed in Table 6 and plotted in Figure 7 together with swimbladder modelling data.

Table 6: Target strengths and lengths of smooth and black oreos extracted from modes in the target strength and length distributions in Figures 5 and 6. The school identifying letters match those in Figures 4–6

Species	School	Target strength (db)	Total length (cm)
Smooth	A	-51.2	35.5
	A	-46.2	39.5
	B	-49.3	37.5
	C	-42.7	39.5
	D	-49.1	38.5
Black	D	-35.0	36.5
	E	-46.6	28.5
	E	-35.1	34.5
	F	-45.6	28.5
	F	-40.2	32.5

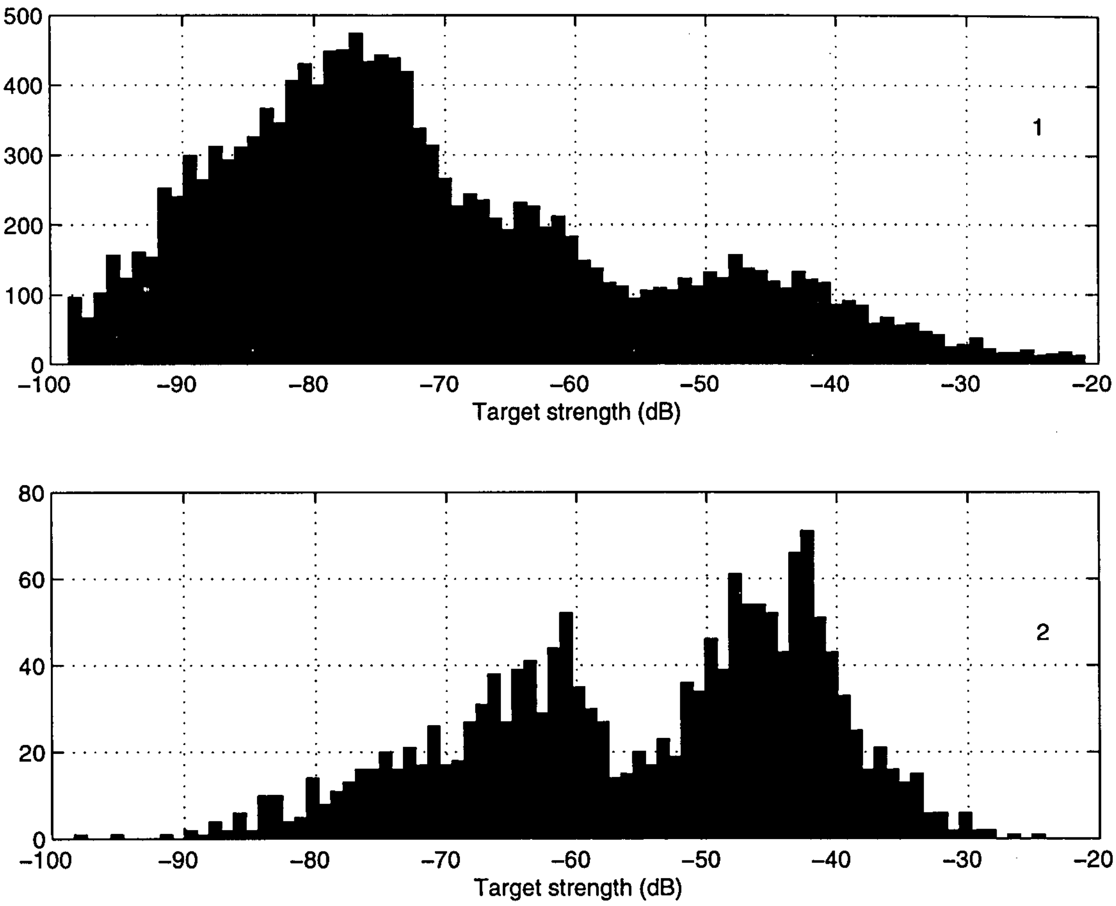

Figure 4: Target strength distributions for school C, with relaxed (1) and more strict (2) criteria for rejecting multiple echoes.

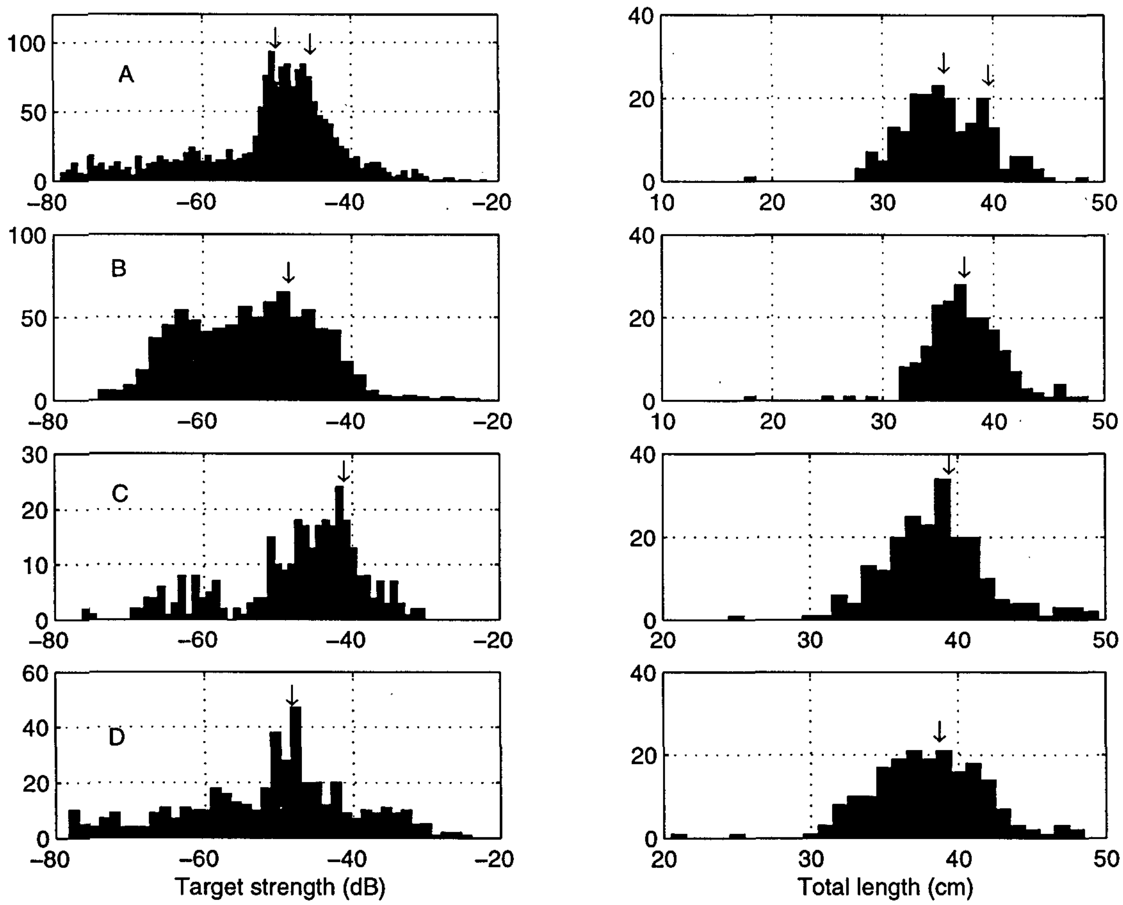

Figure 5: Target strength and length frequency distributions for smooth oreos for 4 fish schools. The y axis is in numbers in all cases. The positions of the modes used are marked with arrows.

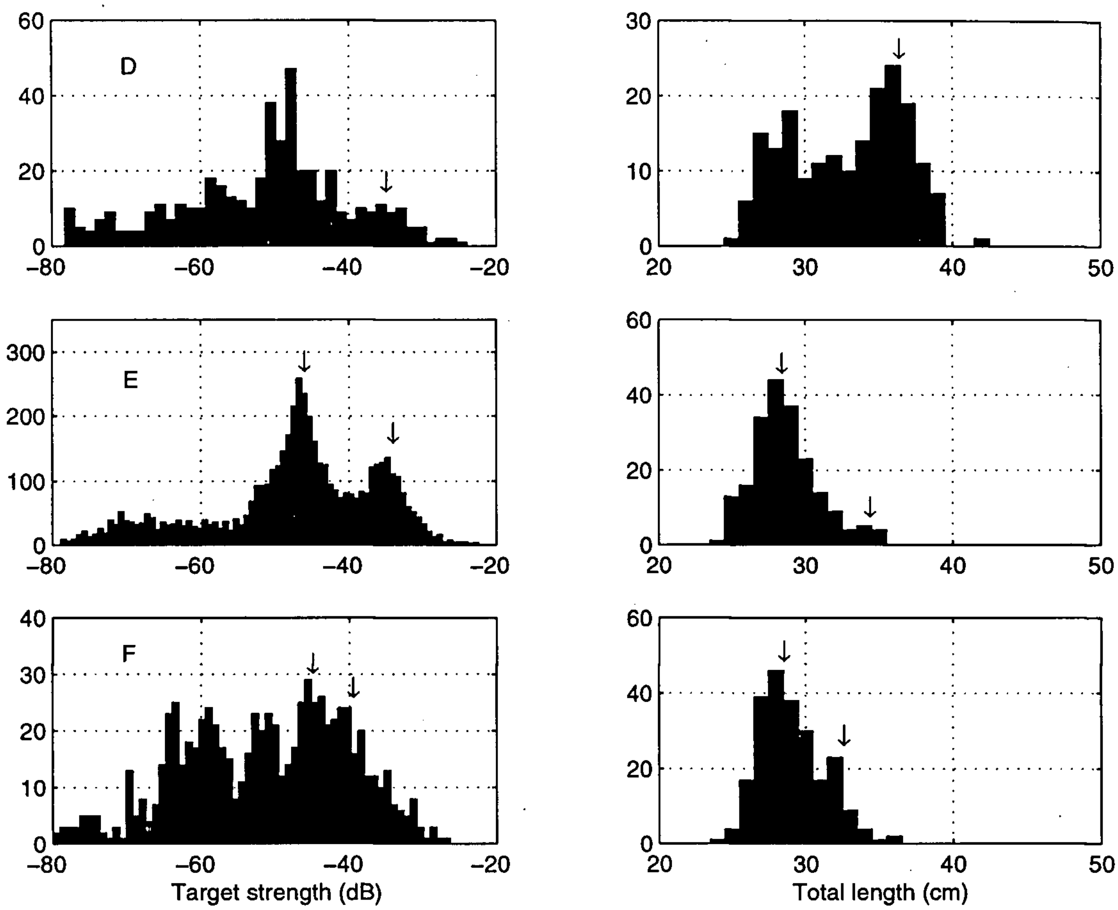

Figure 6: Target strength and length frequency distributions for black oreos for 3 fish schools. The y axis is in numbers in all cases. The positions of the modes used are marked with arrows.

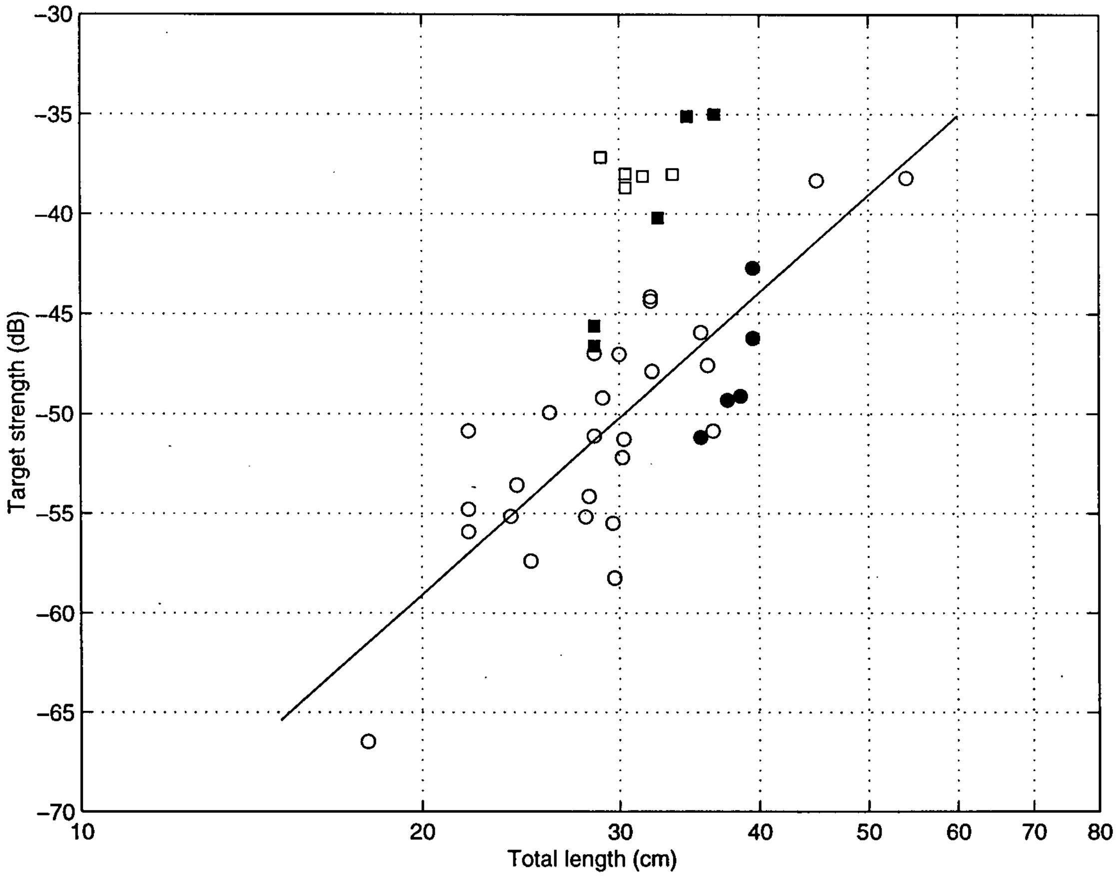

Figure 7: Target strength – length frequency relationships. Black oreo *in situ*, ■, black oreo swimbladder model, □, smooth oreo *in situ*, ●, smooth oreo swimbladder model, ○. The line is fitted to the smooth oreo swimbladder model points.

3.3.3 Target strength relationships used

The following target strength relationships have been used in the biomass estimates. For smooth oreos, the modelling results covered a much wider size range than the *in situ* data and the slope was therefore estimated from these. However, the *in situ* points are more likely to represent the true relationship and they were used to estimate the intercept. The estimated relationship is:

$$\langle \mathrm{TS} \rangle_{sso}(l) = -127.30 + 50.36 \times \log_{10} l$$

For a 35 cm fish, the *c.v.* is 26%

For black oreos only a small number of observations are available and a working relationship has been derived by assuming the slope is the same as for smooth oreo and estimating the intercept from the *in situ* data. The estimated relationship is:

$$\langle \mathrm{TS} \rangle_{boe}(l) = -116.24 + 50.36 \times \log_{10} l$$

For a 35 cm fish, the *c.v.* is 56%.

3.3.4 Associated species

Swimbladder casts were collected from a range of species associated with oreos during the 1996/97 fishing year (TAN9705 and TAN9708). Target strength-length relationships have been estimated from these data for the rattail species *Coryphaenoides serrulatus* and *C. subserrulatus.*

Because target strength is related to the cross-sectional area of fish, a square relationship of the form:

$$\mathrm{TS}(l) = c + 20 \times \log_{10} l$$

(where c is a constant) is commonly assumed (e.g. Foote 1987). This implies that fish body and swimbladder proportions remain constant as the fish grows. The above relationships show that this assumtion is not true for oreos and may not be for any deepwater fish. However, since we have only limited information available, for associated species, we have followed Foote's (1987) model.

For species with swimbladders (*sb*) the slope has been taken to be 20 and the intercept estimated from the mean of intercepts for the rattail and smooth oreo relationships:

$$\langle \mathrm{TS} \rangle_{sb}(l) = -79.4 + 20 \times \log_{10} l$$

For species without swimbladders (*nsb*) Foote's (1987) relationship has been used directly:

$$\langle \mathrm{TS} \rangle_{nsb}(l) = -77.0 + 20 \times \log_{10} l$$

Foote's (1987) data have been used for bootstrapping to estimate the overall mixed species variance.

3.4 Target identification

3.4.1 Mark classification

EK500 echogram data were collected for 47 of the 51 trawls in OEO 3A. Of these, 41 were on the flat and 6 on the hills. Two of the trawls on the flat were 'misses' and the data were not used. Because of problems with synchronising with the other ship's sounders, CREST hull data was not collected until the second half of the voyage and only 14 of the flat, and 4 of the hill trawls were recorded.

The trawl data were initially compared with the associated EK500 acoustic data to find the best match between the trawl groups and mark categories. The categories were then refined using the CREST hull data which contained more detail, particularly when the weather was poor.

The key mark attributes were depth and backscatter density. The denser and larger the mark, the higher the percentage of oreos in the associated catch. Catch rates below 180 $kg.km^{-1}$ showed a wide scatter in proportion of oreos (0–70%). The scatter reduced and the lower limit increased to 40% for catch rates from 180–1000 $kg.km^{-1}$ and at rates higher than 1000 $kg.km^{-1}$, the proportion was 80–100%. Thus, large marks over 30–40 m thick, or plumes over 40 m high, were usually more than 80% oreos. These large marks were also denser, with a higher colour code. In the shallower depths large marks were mainly black oreo, but at greater depths were a mixture of black oreo and smooth oreo. There was usually a background layer or layers of lighter marks, up to 20 m thick, over most of the bottom. They were composed of smaller black oreo and others in shallow and mid-range depths, but contained almost no oreo in deeper waters.

The analysis yielded four mark types as shown in Table 7.

Table 7: Echogram colour and depth thresholds that define the 4 mark types, the number of trawls in each group (n), the average proportion of smooth (SSO) and black (BOE) oreos in the catch

Mark-type	Echogram colour	Depth (m)	n	SSO (%)	BOE (%)
OREO	blue or more	860–1 000	7	63	33
	+ purple or more	+ >1 000			
BOE	blue or more	<860	11	2	74
BOE-back	Up to light blue	<1 000	11	5	38
Back-deep	Up to light blue	>1 000	8	2	1

The different types of marks have the following characteristics:

OREO marks

A mixture of smooth and black oreos, mainly recruited sized fish, with very few non-oreo species. Of the 7 trawls, 4 caught more than 70% smooth oreos, 1 caught black oreo and 2 had a mixture of both. The catch rates were all high. Marks were thicker than 10 m and were usually on the top edge of drop-offs or mounds and were frequently in plumes. Two trawls were excluded because they had heavy plumes yet caught less than 200 kg.

BOE marks

Mainly black oreos (about 5% of which was recruited size fish) with some non-oreo species. These marks were thick (20 m or more) layers on the bottom (or sometimes 50 m off the bottom), on gentle slopes. There was one plume. The catch rates were moderate and high.

BOE-back marks

This was the shallow background layer group and was a mixture of small black oreo and non-oreo species. The layers were 20 m or less thick, usually on flatter bottom. The catch rates were low. A transition zone between 900 and 1000 m existed where some of the background catches were of the back-deep type. The depth threshold was set at 1000 m because catches deeper than this had almost no oreos (the survey lower depth boundary was 1200 m). Deep background marks will not be contaminated with oreos from the transition zone.

Back-deep marks

This was the background layers in deeper waters. It was almost devoid of oreos and contained a higher percentage of rattails and Baxter's dogfish than shallower waters. Catch rates were all low. The marks were layers, 20 m or less thick, on flat bottoms.

The above categories represent an initial compromise between assigning many groups (so that boundaries between them represent variability in the samples) and fewer groups (which will have more variability in the catches within a group). The number of trawls in each group must be large enough so that bootstrapping for the biomass variance will be valid. Each of the four groups chosen here contains about 8 trawls which is about the minimum needed for bootstrapping. The scheme captures the principal acoustic structures of the trawl catch data, but may need to be revised. With more analysis and more data it may prove possible to separate black oreo and smooth oreo.

OEO 4

The species mix differed in OEO 4 from OEO 3A with a different range of rattail species and a significant presence of small orange roughy. Black oreo made up only 1% of the catch so only the 'SSO' mark-type was retained from the above OEO 3A analysis and a new 'Background' type was established. Since only limited trawling was carried out, the criteria for assigning marks to these categories were largely based on the OEO 3A results.

3.4.2 Underwater camera

Two camera drops were carried out to collect target identification and target orientation data. These yielded pictures of lantern fish, a medusa, deep-water prawns and other creatures but no oreos.

3.5 Hydrography

Temperature and salinity profiles were collected at three sites: 44° 42.68' S 173° 24.03' E (drop 1); 44° 56.02' S 174° 37.25' E (drop 2); 44° 38.27' S 177° 32.28' W (drop 3). The absorption was very similar for all three drops (Table 8).

Table 8: Absorption coefficients (dB.km^{-1}) calculated using the methods of both Fisher & Simmons (1977) and Francois & Garrison (1982) for the three CTD drops carried out. – no data

Depth	Fisher & Simmons			Francois & Garrison		
	Drop 1	Drop 2	Drop 3	Drop 1	Drop 2	Drop 3
25	7.9	7.9	7.5	10.1	10.1	9.6
75	7.9	8.0	7.8	10.0	10.1	9.9
125	7.9	8.0	7.9	10.0	10.1	10.0
175	7.9	8.0	8.0	9.9	10.0	10.0
225	7.9	8.0	8.0	9.9	10.0	10.0
275	7.9	8.1	8.0	9.8	10.0	9.9
325	7.9	8.1	8.0	9.8	9.9	9.9
375	7.9	8.0	8.0	9.8	9.9	9.8
425	8.0	8.0	8.0	9.8	9.8	9.8
475	8.0	8.0	8.0	9.7	9.7	9.7
525	7.9	8.0	7.9	9.6	9.7	9.7
575	7.9	8.0	7.9	9.6	9.6	9.6
625	7.9	8.0	7.9	9.5	9.6	9.5
675	7.9	8.0	7.9	9.5	9.5	9.5
725	7.9	8.0	7.9	9.4	9.5	9.4
775	7.9	8.0	7.9	9.4	9.5	9.4
825	7.8	8.0	7.9	9.3	9.4	9.4
875	7.9	–	7.9	9.3	–	9.3
900	7.9	–	7.9	9.3	–	9.3
950	–	–	7.9	–	–	9.2
1 000	–	–	7.9	–	–	9.2

3.6 Biomass estimates

3.6.1 OEO 3A

3.6.1.1 Flat

The biomass estimates for recruited smooth oreos from the flat survey of area OEO 3A by stratum and mark-type are listed in Table 9 and for recruited black oreos in Table 10. Estimates for pre-recruits are in Table 11.

Table 9: Biomass (t) of recruited smooth oreo in OEO 3A by stratum and mark-type. The grand total is 15 400 t

Stratum	1	15	2	3	4	5
OREO	0	1 845	7 127	2 419	1 697	0
BOE	413	34	33	0	509	0
BOE-back	234	57	231	0	257	148
back-deep	0	71	0	322	0	7
Total	647	2 006	7 391	2 741	2 462	155

Estimated biomass of recruited smooth oreos is 15 400 t with a *c.v.* of 46% and 95% confidence interval of 6–33 000 t. The *c.v.* includes a subjective allowance for the experimental error in the estimate of smooth oreo target strength. If this is excluded the *c.v.* reduces to 32%. For recruited black oreos the estimated biomass is 18 800 t with a *c.v.* of 44% and 95% confidence interval of 6–39 000 t. This also includes an allowance for target strength measurements errors and without this the *c.v.* is 28%.

Table 10: Biomass (t) of recruited black oreo in OEO 3A by stratum and mark-type. The grand total is 18 800 t

Stratum	1	15	2	3	4	5
OREO	0	493	1 903	646	453	0
BOE	4 378	359	353	0	5 396	4
BOE-back	883	214	869	0	967	557
back-deep	0	231	0	1 055	0	22
Total	5 261	1 297	3 125	1 701	6 816	582

Table 11: Biomass ('000 t) in OEO 3A for recruit smooth oreo (SSO), pre-recruit smooth oreo (SSO*pr*), recruit black oreo (BOE) and pre-recruit black oreo (BOE*pr*)

Stratum	1	15	2	3	4	5	Total
BOE	5	1	3	2	7	1	19
BOE*pr*	66	6	10	1	81	2	166
SSO	1	2	7	3	2	0	15
SSO*pr*	1	1	2	1	1	0	6

3.6.1.2 Hills

The data from all OEO 3A hills were treated as though they were from a simple random design. This is a reasonable assumption for the parallel transects because drift and position resolution introduced a random element and transect spacing was not uniform. For the star patterns, the top of the hill (centre of the star) is over-sampled relative to the flanks. The areal backscatter for the different sets of transects for the 3 hills is shown in Table 12. Backscatter is consistent between the different runs on the same hill.

To estimate the biomass on each hill the areal backscatter was averaged over all transects for the hill and multiplied by the area of the hill. The results for both oreo species are shown in Table 13. The variances were calculated assuming that transects were independent estimates of abundance which is not strictly correct for the parallel transects and incorrect for the star patterns. However, they give a rough indication.

The estimates are possibly biased by the fact that not all the hills now known to exist in OEO 3A were included in the survey.

Table 12: Areal backscatter (numbers of fish x 10^{-6} per m^2) for the different transects carried out on the three OEO 3A hills

Hill	Direction	Date	Number of transects	Areal backscatter
NC	E-W	23/11	5	4.2
NC	N-S	23/11	5	4.5
NC	Star	9/12	4	3.7
NP	E-W	24/11	3	12.3
NP	N-S	24/11	3	3.6
NP	NW-SE	8/12	6	7.0
HA	Star	9/12	4	2.9

Table 13: Recruited biomass and biomass estimate confidence intervals for smooth oreos on the hills surveyed and total for all hills in area OEO 3A. Recruit biomass estimates only are shown for black oreos

Hill	Area km^2	Smooth Biomass (t)	*c.v.* %	95% C.I.	Black Biomass (t)
NC	4.50	330	52	50–590	1.5
NP	5.66	414	58	99–806	2.9
HA	3.49	214	29	100–300	1.0
Total	–	2 300	34	730–3 900	–

3.6.1.3 Smooth oreo total biomass

The estimated total recruited biomass of smooth oreos in area OEO 3A including both flat and hill areas is 13 700 t with a *c.v.* of 39% (29% if the subjective estimate of variance on target strength is excluded).

3.6.2 OEO 4

3.6.2.1 Flat

Only one flat stratum, stratum 12, was covered in area OEO 4. The number of target identification tows carried out was only 7 which is too few for a full target identification analysis. Only two mark-types, 'SSO' and 'Background' were used as described earlier. Two tows were assigned to Background and five to SSO. Data from the 1995 trawl survey (TAN9511) were used to augment the catch data. Trawls were assigned to the two mark-types on the basis of smooth oreo total catch rate such that catch rates >400 $kg.m^{-1}$ were classified as SSO. This level is consistent with the few trawls carried out in the survey and with the background trawls in OEO 3A.

The estimated biomass of recruited smooth oreos in stratum 12, based on these mark-types, was 172 000 t with a *c.v.* of 77%. The *c.v.* is possibly of limited value since most of the catch data came from a trawl survey which was assigned on the basis of catch rate. In addition, the population *c.v.* (i.e., for a single transect) for acoustic transects in stratum 12 is 187%. This is unexpectedly high and is due to an increasing trend in the number and extent of SSO marks from one end of the stratum to the other.

Black oreos formed only a small part of the population and are essentially part of the background species. The estimate of biomass of recruited black oreos in stratum 12 is 2670 with a *c.v.* of 68%.

The estimates for both species by mark-type are shown in Table 14 for both recruited and pre-recruit fish. Overall estimates for both species and orange roughy are shown in Table 15.

Table 14: Biomass in OEO 4, stratum 12, by mark-type for recruit and pre-recruit smooth oreo (SSO), and recruit black oreo (BOE)

Mark-type	Recruit SSO	Pre-recruit SSO	Recruit BOE
SSO	166 000	78 500	2 600
Background	6 200	10 600	70

Table 15: Biomass in OEO 4, stratum 12, for recruit and pre-recruit smooth oreo (SSO), black oreo (BOE) and orange roughy (ORH)

	Recruits	Pre-recruits
SSO	171 900	89 100
BOE	12 700	3 100
ORH	5 900	2 900

3.6.2.2 OEO 4 Hills

Hills surveyed in OEO 4 included Hegerville, Dolly Parton and Chucky's. Abundance estimates for the two hills with usable data are given in Table 16 and biomass estimate *c.v.*s are in Table 17.

Details of the sampling on each hill are as follows:

Hegerville

Hegerville has an estimated surface area of 17.8 km^2 and was surveyed with parallel transects. An additional transect, perpendicular to the others, was not included in the biomass calculation. Two tows were carried out to estimate species composition. Eight tows from a previous trawl survey (TAN9511) were used to provide more information.

Dolly Parton

Dolly Parton has an estimated surface area of 5.1 km^2 and was surveyed using a star-burst design. Only one tow was carried out. There are no additional data available. With only one tow, the variance of the species composition cannot be determined and has been set to zero. As a consequence, the total *c.v.* of the biomass is underestimated.

Chucky's

Chucky's is a very deep hill and problems with acoustic noise meant that no usable data were obtained. One trawl produced 24 t of recruit sized smooth oreo and 1 t of pre-recruit sized smooth oreo.

Table 16: Biomass estimates of recruit sized fish (t)

	SSO	BOE	ORH	*c.v.* (% SSO)
Hegerville	2 100	20	108	35
Dolly Parton	184	0	0	22

Table 17: Biomass *c.v.* (%) for each source of variation

	Species composition	Acoustic backscatter	Target strength	
			Mix	SSO
Hegerville	16	21	18	14
Dolly Parton	0	8	6	20

3.7 Sensitivity analysis (OEO 3A)

3.7.1 Target strength

For a uniform, monospecific population, a 2 dB decrease in the target strength of the species would produce a 58% increase the backscatter while a 2 dB increase would produce a 37% decrease. However, for mixtures of species the changes are less straightforward and Table 18 shows the effect of target strength changes for a mix of smooth and black oreos on the estimated biomass of each species. The biomass of smooth oreos is insensitive to the smooth oreo target strength but is sensitive to changes in black oreo target strength. The effect is non-linear and there is a proportionally smaller change in smooth oreo biomass if both target strengths are increased together compared to the change if only the black oreo value is increased. Conversely, there is a proportionally bigger change in smooth oreo biomass if both target strengths are decreased together compared to the change if only the black oreo value is decreased.

The biomass of black oreos is insensitive to changes in smooth oreo target strength but sensitive to changes in black oreo target strength.

Table 18: Sensitivity of estimated biomass of the smooth (SSO) and black (BOE) oreos to changes in target strength

Change in TS (dB)		Change in biomass (%)	
SSO	BOE	SSO	BOE
2	0	-11	-3
-2	0	+9	+2
0	2	-30	-33
0	-2	+38	+47
-2	-2	+55	+50
2	-2	+18	+42
-2	2	-26	-32
2	2	-36	-34

The sensitivity of the estimates to the 'other' species target strength was investigated by reverting to Foote's (1987) relationship, i.e., $\langle \mathrm{TS} \rangle_{sb}(l) = -67.5 + 20 \times \log_{10} l$. This reduced the estimate for smooth oreo by 24% and black oreo by 43%.

3.7.2 Background layer

The effect of the height of the background layer was investigated by integrating up to 25 m instead of 50 m. All marks in the non-background categories were included up to their full height. The lower layer height reduced the smooth oreo estimate by14% and the black oreo estimate by 34%.

3.7.3 Proportion of smooth oreo in the commercial catch

Commercial catches were selected where the total oreo catch was 5 t or more, and the catch was in stratum 2 or 3, for fishing years 1991–92 to 1996–97. The proportion of smooth oreos in the catch was 48% by weight (the number of tows was 1554 and unspecified oreo catches were discarded.) For each fishing year, the proportions of smooth oreo were (in order, starting with 1991–92): 54, 42, 44, 66, 49, and 35%. The data showed a strong depth cline in the proportion of smooth oreo in the catch (more smooth oreo at greater depth) and a step in the proportion at about 900 m. When the data were restricted to depths of 900m or more, the proportion of smooth oreo rose to 58%. By comparison, the proportion of smooth oreo in the catches of tows assigned to the OREO mark-type was 67%.

3.7.4 Oreo layers

Various alternative classifications and proportions in the oreo (OREO, BOE and BOE-back) mark categories were investigated.

- The threshold density between OREO, BOE and background categories was increased by 20% (based on the received voltage level), increasing the backscattering allocated to the background. This reduced the smooth oreo biomass by 9% but increased the black oreo biomass by 13%.
- Excluding mid-water BOE mark-types increased smooth oreo biomass by 3% but reduced black oreo biomass by 11%
- Reducing the proportion of smooth oreo in the OREO mark-type from 59% to 50% (to reflect the proportion in commercial catches) reduced the smooth oreo biomass estimate by 30% and increased the black oreo biomass estimate by 7%.

3.7.5 Transducer motion

Transducer motion due to movement of the towed body in bad weather may result in biomass being underestimated. Transducer motion was not analysed so the extent of the bias from this source is unknown but is not thought to be substantial.

3.7.6 Absorption coefficient

The absorption coefficient for smooth oreo only was adjusted from 8 $dB.km^{-1}$, the Fisher & Simmon (1977) value used in the base case, to 9.4 $dB.km^{-1}$, an alternative value estimated from a relationship derived by Francois & Garrison (1982). The biomass estimate increased by 60%.

3.7.7 Shadowed zone

The median of the extra backscatter in the shadowed zone (Table 19) was 7.4% (the mean was 7.6%). The marks were selected approximately in proportion to their numbers in each category so the median (or mean) can be applied to the oreo biomass. Strictly, the sample sizes from each mark class should be proportional to their contribution to the biomass. An approximate correction to the biomass would therefore be to increase it by 7–8%.

Table 19: Estimated shadowed zone height (equivalent height of a conical layer) and the extra backscatter in the shadowed zone (as a percentage of the backscatter in the mark)

Code	Transect	Stratum	Shadowed zone height (m)	Extra back-scatter (%)
PS	46	2	3.7	10
OF	112	2	1.1	1
OF	134	15	1.6	4
F50	102	15	1.1	2
S50	26	2	1.9	11
S50	73	4	1.7	10
DF	12	2	1.2	24
DF	96	3	1.7	7
DF	52	15	1.5	3
DF	35	2	1.7	0
DS	25	2	0.9	11

4. REFERENCES

Annala, J.H & Sullivan, K.J. (Comps) 1997: Report from the Fishery Assessment Plenary, April-May 1997: stock assessments and yield estimates. 381 p. (Unpublished report held in NIWA library, Wellington.)

Cordue, P. L. 1996: On the feasibility of using acoustic techniques to estimate orange roughy and oreo biomass on underwater hills. NZ Fisheries Assessment Research Document 96/22. 21 p.

Ehrenberg, J. E. 1979: A comparative analysis of in situ methods for directly measuring the acoustic target strength of individual fish. *IEEE Journal of Oceanographic Engineering,* Vol. OE-4, No. 4, October 1979.

Fisher, F. H. & Simmons, V. P. 1977: Sound absorption in seawater. *Journal of the Acoustical Society of America 62*: 558–564.

Foote, K.G. 1987. Fish target strengths for use in echo integrator surveys. *Journal of Acoustic Society of America 82:*981–987.

Francis, R.I.C.C. 1984: An adaptive strategy for stratified random trawl surveys. *New Zealand Journal of Marine and Freshwater Research 18:* 59–71.

Francois, R. E. & G. R. Garrison (1982). Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption. *Journal of the Acoustical Society of America 72*: 1879–1890.

Jolly, G. M. & Hampton, I. 1990: A stratified random transect design for acoustic surveys of fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences 47*: 1282–1291.

Kloser, R. J. 1996: Improved precision of acoustic surveys of benthopelagic fish by means of a deep-towed transducer. *ICES Journal of Marine Science 53*: 407–413.

McClatchie, S., Alsop, J., Ye, Z. & Coombs, R.F. 1996: Consequence of swimbladder model choice and fish orientation to target strength of three New Zealand fish species. *ICES Journal of Marine Science 53:* 847–862.

McClatchie, S.; Macaulay, G., Ye, Z.; Coombs, R.F., Grimes, P., & Hart, A. in press: Target strength of the deep-water fish, orange roughy (*Hoplostethus atlanticus*). Part 1: Experiments. *Journal of the Acoustical Society of America.*

McMillan, P.J., Doonan, I.J., Coburn, R.P. & Hart, A.C. 1996: Is the south Chatham Rise trawl survey series providing an index of smooth oreo abundance in OEO 4 N.Z. Fisheries Assessment Research Document 96/16. 18 p.

Ona, E & Mitson, R. B. 1996: Acoustic sampling and signal processing near the seabed: the deadzone revisited. *ICES Journal of Marine Science 53*: 677–690.

Soule, M., Barange, M., & Hampton, I. 1995: Evidence of bias in estimates of target strength obtained with a split-beam echo-sounder. *ICES Journal of Marine Science 52*: 139–144.

Soule, M., Barange, M., Solli, H. & Hampton, I. 1997: Performance of a new phase algorithm for discriminating single and overlapping echoes in a split-beam echosounder. *ICES Journal of Marine Science 54*: 934–938.

Urick, R. J. 1983: Principles of underwater sound. McGraw-Hill.

Part 2: Oreo mark identification

Alistair Dunn

1. INTRODUCTION

The project objective (Objective 1) is "To estimate the absolute abundance, with a target coefficient of variation (*c.v.*) of the estimate of 20–30%, of black oreo and smooth oreo in OEO 3A and 4 on the Chatham Rise". This report partially fulfils the target "Target 4. Improving mark identification". In particular, this report discusses the statistical analysis of predictive models as a means of improving the target mark identification for black and smooth oreos.

2. METHODS

The data analysed in this report consist of 37 observations of acoustic marks and associated fishing tow data collected by *Tangaroa* during the acoustic survey of black and smooth oreos of OEO 3A in October 1997 (TAN9713). The problem was to formulate decision rules on the species composition of these acoustic marks that may assist in the future identification of acoustic marks found on Chatham Rise and other oreo surveys — decision rules that would be based on predictive statistical models.

Identifying the composition of acoustic marks is problematic. In some fisheries, an estimate of composition is made by observers who use biological knowledge, anecdotal evidence, and past experience to calculate — within broad parameters — the species composition of these marks. On oreo surveys, a simplified decision algorithm is used; where each mark on an echogram is classified using simple rules based on the depth and the intensity of the image. In general, trawls are carried out on selected marks during the survey to inform and confirm the decision process. However, the identification process may often amount to little more than an "educated guess". This is not a fault of the observers, but rather a lack of analyses and empirical data on which to derive predictive models. In general, there are few successful objective analyses that have been conducted on either acoustic mark identification or the subsequent conclusions drawn by observers.

One key aspect of mark identification is the problem of *ground truthing*. This can be interpreted as a problem in identifying the actual (or true) species composition of an acoustic mark once it has been located. Without this information, it is difficult to deduce from *unsupervised* statistical methods (i.e., methods that look at relationships *within* a set of data) the association between resulting classification and actual species composition. When true classification information is available, *supervised* methods (i.e., methods that attempt to describe the classifications based on some explanatory data) can be used to generate predictive models. However, the process typically requires a large body of data, especially when the number of possible explanatory variables is also large. The usual approach to developing such models is to fit, then evaluate each model via cross-validation. This would usually allow a comparison of the relative predictive power of a range of methodologies that may be employed, as well as an estimation of the misclassification rate for each model. However, for the oreo data, only small numbers of data points are available. This restricts the overall complexity of individual models and any subsequent validation that can be employed. With only a small amount of data, the methods employed here constitute only a preliminary investigation.

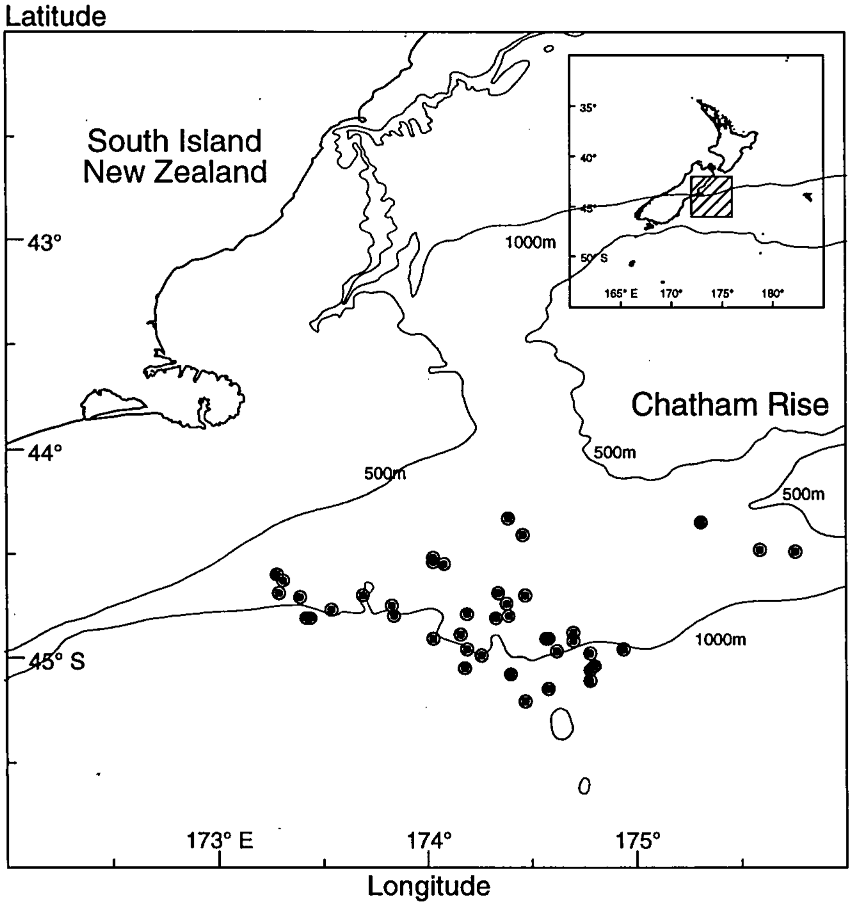

Figure 1: The Chatham Rise, with locations of the oreo acoustic marks and associated fishing tows (circles)

We discuss a range of supervised statistical methodologies, initially using traditional methods for which the distributional theory is well established within the statistical literature and later extend to more modern methods about which much less is known. Supervised methods are generally those that attempt to define the conditions that divide a group of observations into predefined groups (Ripley 1996). The supervised methodologies that are discussed are discriminant analysis methods (including linear and flexible discriminant analyses) and classification trees.

2.1 Description of the data

The acoustic marks used in this analysis are those recorded during TAN9713 where species composition had been estimated by trawling. Once an acoustic mark had been located, an appropriate net was deployed, and attempts made to tow it through the centre of the mark. While this is known to be a source of many potential biases, we assume that the resulting estimated species composition provides the true relative species composition for that mark.

The variables consist of the estimated species composition resulting from a single tow together with a summary of the characteristics and locations of the acoustic mark. These are listed in Table 4 in the Appendix. Table 1 gives the variables collected and associated definitions. Species composition is determined from the tow through the mark, and is the proportion (by weight) of smooth oreos, black oreos and other species resulting. This is used as the ground truth for the classification and predictive models.

Many of the methods that can be employed in classification analysis require that each observation belong to a discrete class (or category). Continuous methods are available, but are often more restrictive in terms of the assumptions required. We investigate only discrete methods. The species composition of the catch is hence converted into a single categorical variable.

The analysis was restricted to classifications of smooth oreos, black oreos, and other species. Adult (or recruited sized fish) and juveniles (non-recruited sized fish) were combined. The categories determined from the species composition in the trawl used the following algorithm:

0 Any species making up more than 67% of the total catch is coded as that species, i.e., Smooth (S), Black (B), & Other (O)
1 Of the remaining observations, the combination of any two species making up more than 67% of the catch is coded as a combined species, i.e., Smooth/Black (sb), Smooth/Other (so), & Black/Other (bo).

Acoustic marks can range from single, through various forms of clustered echoes, and can be represented graphically as an echogram (*see* Figure 3 for an example echogram). A large proportion of estimated biomass of oreos tends to be found in highly aggregated, dense clusters which produce large acoustic marks. The data, therefore, comprise of a sample of dense, clustered marks that have been sampled by trawling. The variable type records the approximate shape of the acoustic mark, with the values defined by the nomenclature of Cordue (1996). Density records the approximate relative (log) density of the mark, on a 0–5 scale. This value is based on the colour coding used to represent relative density in echograms, which is also a log scale. Height records the estimated height or vertical length of the mark in metres. This value is estimated from visual inspection of the echogram. Similarly, depth records the depth of the sea floor (in metres) beneath the centre of the mark, and is estimated in the same way. Location records the location of the mark in relation to the physical features on the sea floor. Oreos are believed to be found in close proximity to small hills and drop-offs (or cliffs), as well as on flatter surfaces. See Figure 2 for a graphical representation of the variables and associated definition

The small number of observations is of some concern. In order to simplify analysis, the descriptor variables have been reduced to simple classification values. More complex and detailed descriptions of the acoustic marks are available, but would introduce a complexity into the analysis that is unjustifiable in terms of the amount of data available.

Table 1: The descriptions and possible values of variables characterising the acoustic marks

Variable	Description	Values
Species composition		
	% smooth oreos within the catch	
	% black oreos within the catch	
	% other species within the catch	
Mark type	A descriptor of the type of mark	Plume: a vertical spike, pole, or column Layer: a flat long structure
Mark density	An estimate of the maximum mark density	Log scale, from 0–5, based on the colour of pixels making up the mark
Mark height	The height of the mark in metres	Measured by the sonar device (range 3–160m)
Mark depth	The depth of the sea floor beneath the mark	Measured by the sonar device (range 600–1200m)
Mark location	The location of the mark with reference to sea floor features	Top: at the top of a small hill or drop-off Bottom: at the bottom of a small hill or drop-off Flat: On a flat section of sea floor
Group	The classification of the catch into a categorical variable	S: smooth oreos B: black oreos O: Other species sb: smooth oreos and black oreos so: smooth oreos and other species bo: black oreos and other species

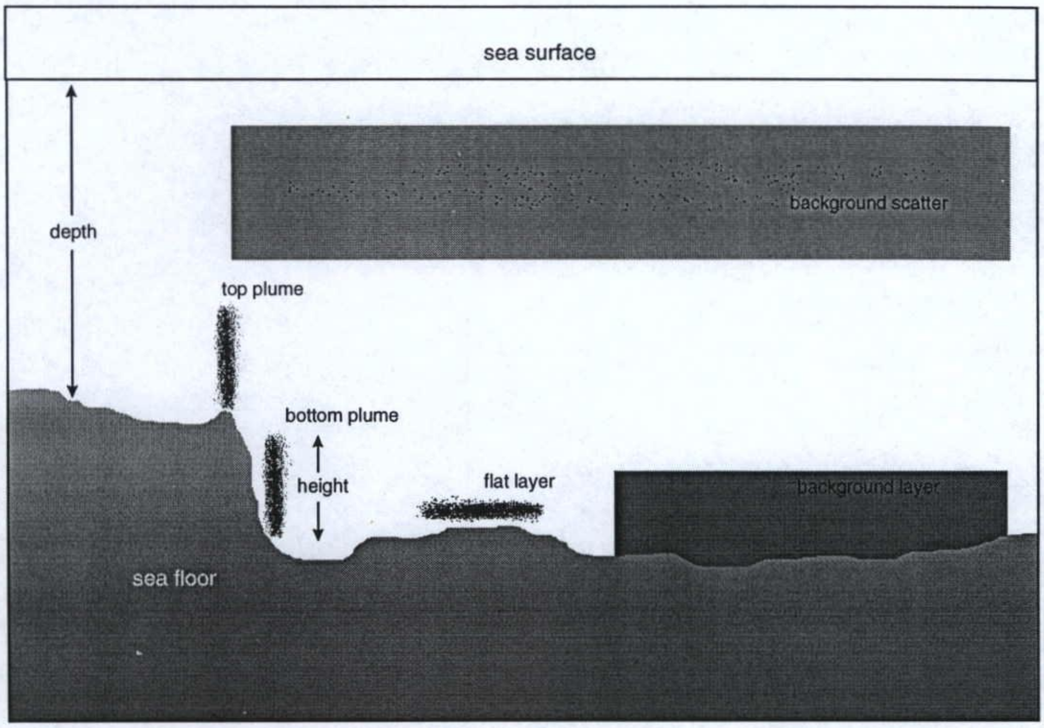

Figure 2: Graphical representation of the variables used in classifying acoustic marks; mark location, type, height and depth

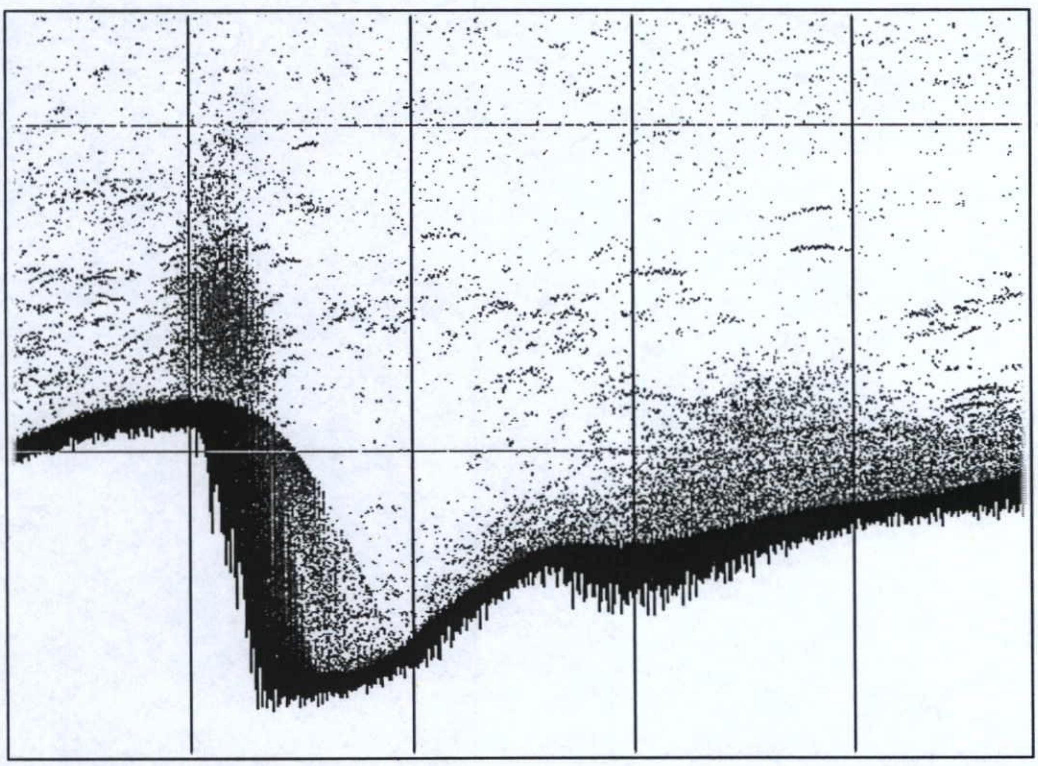

Figure 3: An example echogram, showing a plume located at the top of a hill, with a large degree of background scatter and shadowing (note that a large degree of density information has been lost in rendering the echogram onto a grey scale)

3. RESULTS

3.1 Linear discriminant analysis

Linear discriminant analysis (Fisher 1936) seeks a linear combination of the variables which has a maximal ratio of the separation of the class (or group) means to the within-class variance (Venables & Ripley 1994). The linear discriminant analysis was fitted using the function `lda` (Venables & Ripley 1994) in S-Plus (Mathsoft 1997) using all five explanatory variables. The resulting estimated linear discriminants are presented below (*see* Table 2).

Table 2: Coefficients of the estimated linear discriminants

Variable	LD1	LD2	LD3	LD4
Density	-0.68	-0.64	0.83	-0.05
Height	0.38	1.44	1.03	1.22
Type	-0.80	-0.92	-0.87	-1.73
Location	-1.03	-0.66	-0.45	0.61
Depth	0.85	-1.25	0.89	0.22

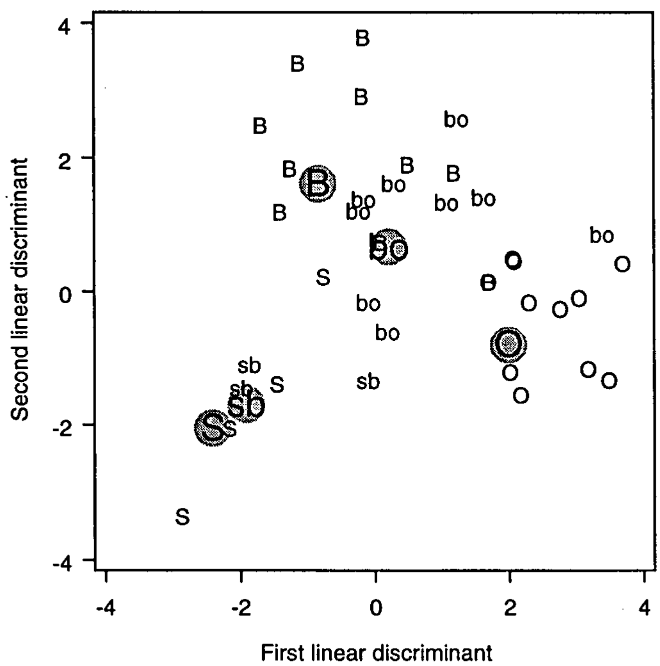

Figure 4: Equal-scaled plot of the first and second linear discriminants, with group means circled

The first and second linear discriminants explained approximately 96% of the between group variance in the data. The classifications resulting from the first two linear discriminants are plotted below in Figure 4, with the misclassification rate shown in Table 3. There is reasonable separation between each of the groups. The linear discriminant analysis has some difficulty distinguishing between those groups which encompass more than one type of species and groups that are only one species. The plot of the linear discriminants suggests that there is a real signal within the data, with the placement of the three single species groups outside that found for the composite groups — a result that we might have desired in a classification algorithm. A misclassification rate can be estimated from the analysis by using the linear discriminants to "predict" the categories of the original input data. For this data, linear discriminant analysis misclassified 9 of the original observations (thought this is not strictly an accurate computation as the data from which the predictions are drawn are not independent of the model).

Table 3: The number of misclassifications and rate for each method

Method	Number misclassified	Per cent misclassified
Linear discriminant analysis	9	24.3%
Polynomial discriminant analysis (with degree=1)	4	10.8%
Polynomial discriminant analysis (with degree=2)	5	13.5%
MARS flexible discriminant analysis	14	37.8%
BRUTO flexible discriminant analysis	11	29.7%
Classification tree	10	27.0%

3.2 Flexible discriminant analysis

Flexible discriminant analysis (Hastie *et al.* 1994) is an extension of Fisher's linear discriminant analysis to non-parameteric methods. Fisher's linear discriminant analysis uses linear boundaries, and requires the assumption of normality. Hastie *et al.* (1994) proposed flexible discriminant analysis as an alternative, where a range of non-parameteric regression algorithms with non-linear boundaries can be employed to fit a range of models. An additional advantage is the ability to use categorical data within the regression fit, rather than the sometimes arbitrary numerical mapping required for Fisher's method.

We employ three approaches to flexible discriminant analysis here; polynomial regression; the MARS (multivariate adaptive regression splines) procedure (Freidman 1991), and BRUTO (adaptive additive modelling) procedure. See Hastie *et al.* (1994) for details on the derivation and implementation of each of these methods. Each of the methods were fitted using the S-Plus implementation `fda` (Hastie *et al.* 1994).

(i) Polynomial regression (with degree=1)

Polynomial regression fits using the same method as for MARS and BRUTO, but uses a polynomial regression rather than the more modern regression techniques. If the highest degree polynomial is set to one, then this can be considered to be equivalent to Fisher's linear discriminant analysis above. However, unlike Fisher's method, this allows the use of categorical variables within the regression.
Figure 5 shows the plot of the first two (linear) discriminants. As would be expected, the resulting discriminants are similar to that for Fisher's linear discriminants above, with the first two discriminants explaining 92% of the between group variance. Again, the single species are well separated, with combined species groups sitting between their respective single species.

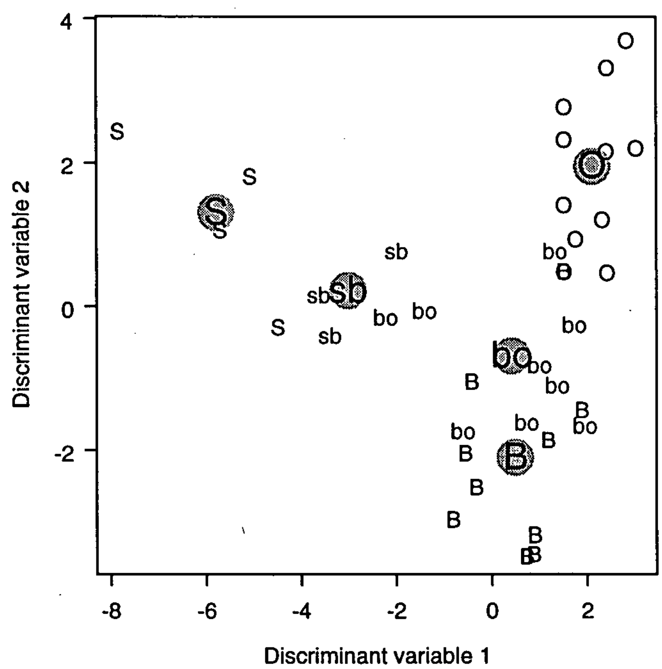

Figure 5: The first two discriminants for polynomial (degree=1) flexible discriminant analysis, with group means circled

(ii) Polynomial regression (with degree=2)

Polynomial regression with degree 2 was fitted. As Figure 6 shows, the resulting fit is similar to that for the polynomial regression with degree=1, though a slightly higher misclassification rate (13.5% versus 10.8%) and the first two discriminants explaining a slightly lower 90% of the between group variance.

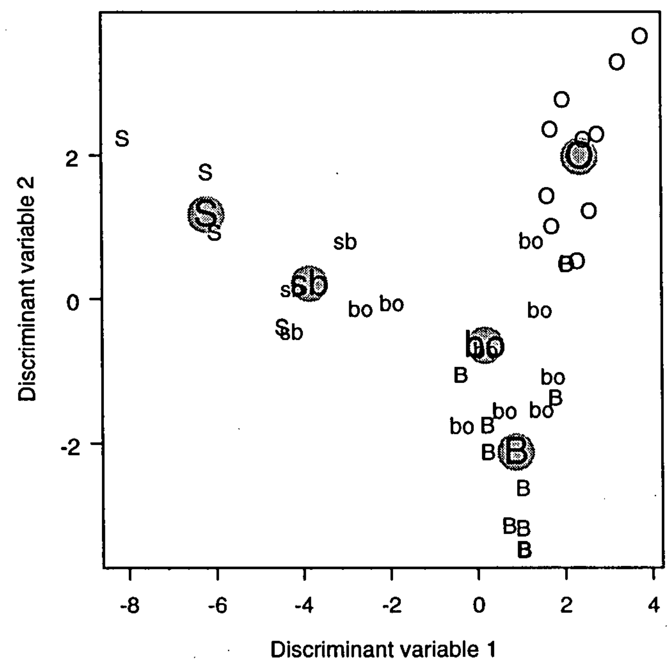

Figure 6: The first two discriminants for polynomial (degree=2) flexible discriminant analysis, with group means circled

(iii) MARS flexible discriminant analysis

Figure 7 shows the plots of the first two discriminants using the MARS algorithm. The model fit is poor in comparison to that of Fisher's method with a higher misclassification rate. However, again clear discrimination between the single species groups is evident. Less obvious is the discrimination between combined groups of species. The model shows some evidence of over-fitting, presumably due to the extra parameters required for the non-parameteric regression in conjunction with the small number of data points available for analysis. Misclassification is high (38%), though the model estimated that 94% of between group variance is explained.

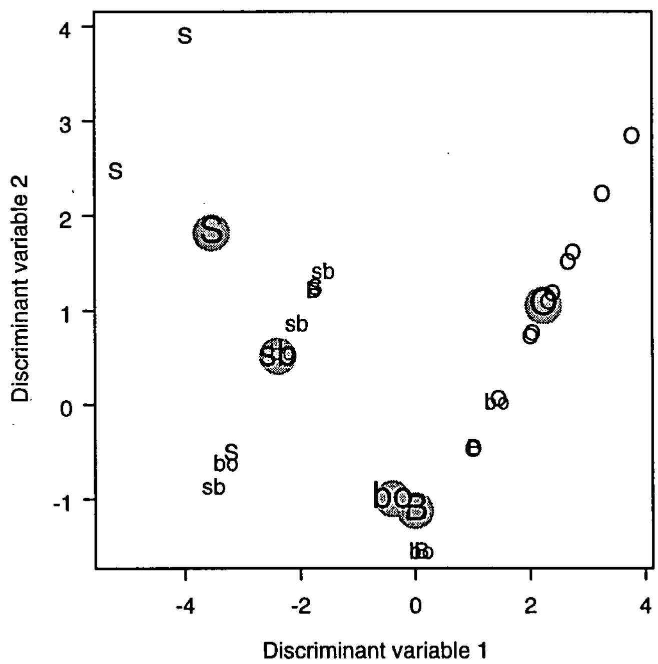

Figure 7: The first two discriminants for the MARS flexible discriminant analysis, with group means circled

(iv) BRUTO flexible discriminant analysis

Figure 8 shows the plots of the first two discriminants using the BRUTO algorithm. The model fit is again poor in comparison to that of Fisher's method, and has a higher misclassification rate. However, again clear discrimination between the single species groups is evident. Less obvious is the discrimination between combined groups of species. As with the MARS model, there is some evidence of over-fitting. Again misclassification is high (30%), though lower than that for the MARS algorithm.

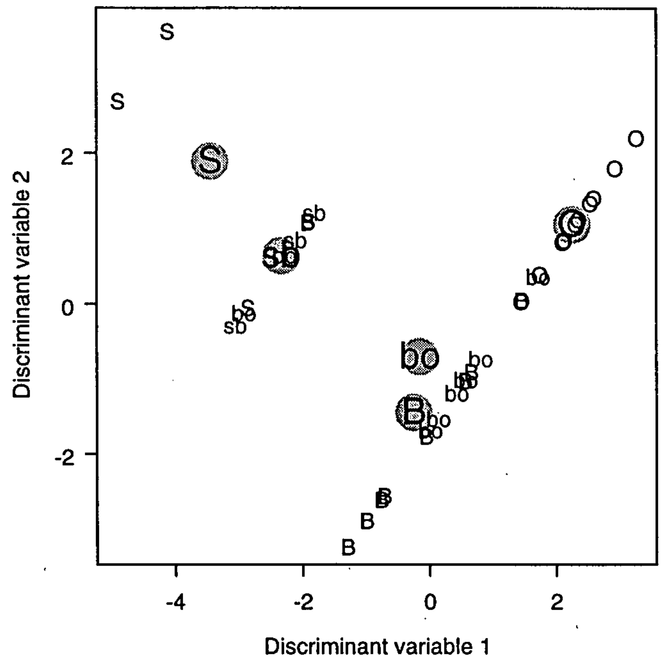

Figure 8: **The first two discriminants for the BRUTO flexible discriminant analysis, with group means circled**

3.3 Classification trees

A classification tree seeks to construct a "tree" or digraph, where each node splits the data into groups according to some decision rule, using recursive partitioning. Each decision rule can be based on either a categorical or continuous variable. Classification trees are implemented in S-Plus as the function `tree` (Mathsoft 1997). Classification trees have an advantage over other forms of classification analysis in that they allow a mixture of both numeric and categorical data as inputs, as well as being less sensitive to non-additive behaviour (Mathsoft 1997). In addition, they can be more easily interpreted as decision rules.

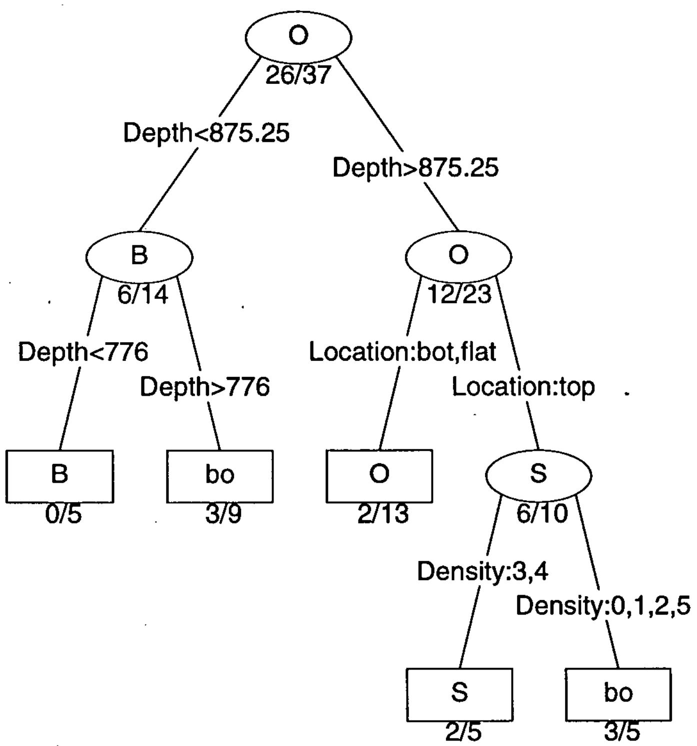

Figure 9: Classification tree showing the nodes (ovals), terminal nodes (rectangles) and missclasification at each terminal node.

Figure 9 shows the classification tree estimated from the acoustic data, with 5 terminal nodes. Three variables are selected for classification: depth, location and density. Type and height are not selected. Depth clearly was considered an important predictor variable, as shown by its location at the top of the classification tree, however, the misclassification rate of the tree is high.

4. CONCLUSIONS

A wide range of statistical methodologies are available for investigating predictive models and classification. This report investigates a small proportion of them. Possible techniques for future work include multinomial logistic regression, generalised additive models, neural networks and regression trees. A requirement, however, for all methods is adequate data. Small data sets such as the one used in this report allow only limited investigations.

The discriminant analysis results suggest that a decision rule based on gross parameters of acoustic marks may be useful in predicting species composition. The degree of separation between groups of species is promising. The classification tree is less convincing. While the non-linear forms of discriminant analysis should provide better fits (especially considering that the explanatory variables are unlikely to have a normal distribution), the lack of data precluded a proper comparison.

The use of trawling for establishing ground truth is problematic. Tows may not, in reality, be very representative of the species composition of an acoustic mark. However, it is difficult to see a way around this problem. Of more practical concern, therefore, are the differences that are hidden by the use of coarse species categories. Smooth and black oreos are known to sometimes congregate as juveniles and adults separately. If these concentrations have different properties, then this would introduce a further complication into any predictive analysis. This analysis used a simple classification algorithm for the composition of each classification category. With more data, a greater number of categories and a more complex algorithm could be included into the analysis.

The comparison of the predictive qualities of the acoustic data on species composition is usually carried out using cross validation techniques. The principle is that a training data set (typically a sample chosen at random from the complete set of data) is used to generate models. These models are then tested for their predictive qualities on the remaining data. Without a large data set, comparison between models is not possible. In addition, future research will need to consider a risk or cost function. Such a function should allow specification of the cost or risk of a misclassification of an acoustic mark. This is a necessary requirement for comparing predictive methodologies.

5. ACKNOWLEDGEMENTS

Thank you to the large number of people within NIWA and on the *R.V. Tangaroa* who were responsible for collecting the data used within this report, including Ian Doonan who also provided useful advice. This project was funded by the Ministry of Fisheries under project code OEO9701

6. REFERENCES

Cordue, P.L. 1996: A nomenclature for marks occurring on echograms. NIWA internal report (fisheries) 257. 21 p. NIWA, Wellington.

Fisher, R.A. 1936: The use of multiple measurements in taxonomic problems. *Annals of Eugenics 7*: 179–188.

Freidman, J. 1991: Multivariate adaptive regression splines (with discussion). *Annals of Statistics 19(1)*: 1–141.

Hastie, T., Tibshirani, R. & Buja, A. 1994: Flexible discriminant analysis by optimal scoring. *JASA*: 1255–1270.

Mathsoft 1997: *S-PLUS Guide to Statistics.* Data Analysis Products Division, Mathsoft, Seattle WA. 877 p.

Ripley, B.D. 1996: Pattern recognition and neural networks. Cambridge University Press, Cambridge. 403 p.

Venables, W.N. & Ripley, B.D. 1994: Modern Applied Statistics with S-PLUS. 2nd edition. Springer-Verlag, New York. 462 p.

Appendix

Table 4: The observations, detailing the species composition, acoustic mark parameters and classification group

Species composition (%)			Mark	Mark	Mark	Mark	Mark	Group
smooth	black	other	type	density	height	location	depth	
0.36	0.49	0.16	layer	5	20	top	878.0	Smooth/Black
0.87	0.11	0.02	layer	3	40	top	900.0	Smooth
0.94	0.04	0.02	layer	4	10	top	901.0	Smooth
0.84	0.11	0.05	plume	3	70	top	991.0	Smooth
0.22	0.72	0.06	plume	4	160	top	1010.0	Black
0.76	0.20	0.04	plume	5	60	top	1012.0	Smooth
0.50	0.44	0.06	plume	4	100	top	1023.0	Smooth/Black
0.00	0.75	0.25	layer	5	20	flat	628.0	Black
0.00	0.77	0.23	layer	5	50	flat	662.5	Black
0.00	0.81	0.19	layer	5	20	flat	694.5	Black
0.00	0.96	0.04	plume	5	120	bottom	772.0	Black
0.00	0.38	0.62	layer	4	20	flat	795.0	Black/Other
0.01	0.55	0.44	layer	5	40	flat	828.5	Black/Other
0.01	0.94	0.05	layer	4	20	flat	844.0	Black
0.02	0.91	0.07	layer	4	50	flat	855.0	Black
0.00	0.75	0.25	layer	2	20	flat	690.0	Black
0.06	0.43	0.51	layer	2	10	flat	780.0	Black/Other
0.05	0.64	0.31	layer	0	20	flat	795.0	Black/Other
0.07	0.85	0.08	layer	1	20	flat	843.5	Black
0.01	0.44	0.55	layer	1	10	flat	845.0	Black/Other
0.01	0.34	0.65	layer	0	10	flat	872.5	Black/Other
0.16	0.49	0.35	layer	0	5	top	894.0	Black/Other
0.04	0.00	0.96	layer	0	5	flat	946.5	Other
0.03	0.19	0.78	layer	1	5	flat	947.0	Other
0.01	0.76	0.23	layer	1	5	flat	947.5	Black
0.05	0.44	0.51	layer	1	20	top	968.5	Black/Other
0.11	0.32	0.57	layer	2	10	bottom	978.5	Black/Other
0.04	0.11	0.85	layer	1	20	flat	980.5	Other
0.25	0.65	0.10	layer	2	15	top	988.0	Smooth/Black
0.02	0.00	0.98	layer	1	15	flat	1022.5	Other
0.02	0.01	0.97	layer	2	10	bottom	1025.0	Other
0.06	0.02	0.92	layer	0	10	flat	1046.0	Other
0.01	0.06	0.93	layer	2	5	flat	1051.5	Other
0.06	0.00	0.94	layer	0	20	flat	1072.5	Other
0.01	0.00	0.99	layer	2	3	flat	1079.0	Other
0.01	0.00	0.99	layer	0	5	flat	1118.0	Other
0.00	0.00	1.00	layer	0	10	flat	1157.0	Other