ICES Journal of Marine Science, 60: 132–146. 2003
doi:10.1006/jmsc.2002.1331

Star acoustic surveys of localized fish aggregations

I. J. Doonan, B. Bull, and R. F. Coombs

Doonan, I. J., Bull, B., and Coombs, R. F. 2003. Star acoustic surveys of localized fish aggregations – ICES Journal of Marine Science, 60: 132–146.

"Stars" are an alternative design for acoustically acoustically surveying isolated fish aggregations in which transects cross at a centre point over an aggregation to form a star or wheel-spoke pattern. The method was devised particularly for aggregations of orange roughy (*Hoplostethus altanticus* Collett) to cope with the practical difficulties of manoeuvring a vessel towing a transducer over a small, dense school of fish. Several methods of analysing star patterns were considered including kriging and a method that ignored the spatial arrangement of the transects. However, simulations of roughy aggregations showed that the best approach was to transform from Cartesian to polar coordinates and then use standard statistical methods (polar method). The polar method was robust to shifts in the transect centre off the aggregation centre and aggregation movement in a random way. Variance estimation was best with a polar version of transitive kriging. Stars using the polar method were also better than the usual parallel transect design when transect numbers were low but the results were similar when six or more transects were used. However, in all the cases considered parallel transects consistently overestimated the variance. We conclude that star transects offer a robust and effective way of estimating the biomass of small, localized aggregations of fish that minimizes vessel time, and yields good precision.

Keywords: fish aggregations, orange roughy, star acoustic survey.

I. J. Doonan, B. Bull, and R. F. Coombs: NIWA, Wellington, New Zealand. Correspondence to I. J. Doonan: NIWA, PO Box 14-901, Kilbirnie, Wellington, New Zealand; tel: +64-4-386 0300; e-mail: i.doonan@niwa.co.nz

Introduction

Many fish species form isolated, localized aggregations up to two kilometres or so across. Particular examples are tunas under fish-aggregation devices (FADs) and deepwater species such as the orange roughy on seamounts. Acoustic techniques are often the best way of estimating the biomass of such aggregations but the survey design poses some problems. One approach is to use "stars" in which transects cross at a centre point over an aggregation to form a star or wheel-spoke pattern. Stars were first used by Depoutot (1987) for tunas under FADs to track density distributions over time and, more recently, Josse *et al.* (1999) have used various star configurations to study tuna distributions under FADs using an acoustic system with a hull-mounted transducer. Independently we have adopted star transects for estimating acoustically the biomass of orange roughy (*Hoplostethus altanticus* Collett), smooth oreo (*Pseudocyttus maculates* Gilchrist) and black oreo (*Allocyttus niger* James, Inada, and Nakamura) on seamounts and in spawning aggregations. Figure 1 shows an echogram of a spawning aggregation of orange roughy on Cameron's seamount at the eastern end of the Chatham Rise, New Zealand.

Acoustic surveys of these deepwater species in New Zealand waters are usually on spawning aggregations that form over a short period in mid-winter. There are a number of physical problems, including poor weather and great ranges, that make the use of a hull-mounted transducer frequently unsatisfactory (Do & Coombs, 1989). Then again seamounts often have steep slopes on which the bottom-shadowing effect is an important consideration. To minimize these and other problems we use a transducer mounted in a deep-towed underwater vehicle.

Since the focus of our surveys is almost always on estimating biomass for stock assessment we use a stratified, random survey design, with randomly spaced parallel transects normally (Jolly & Hampton, 1990). However, there are practical difficulties in steaming closely spaced parallel transects on small aggregations and this led us to adopt stars initially because they are easier to traverse and result in less wear-and-tear and

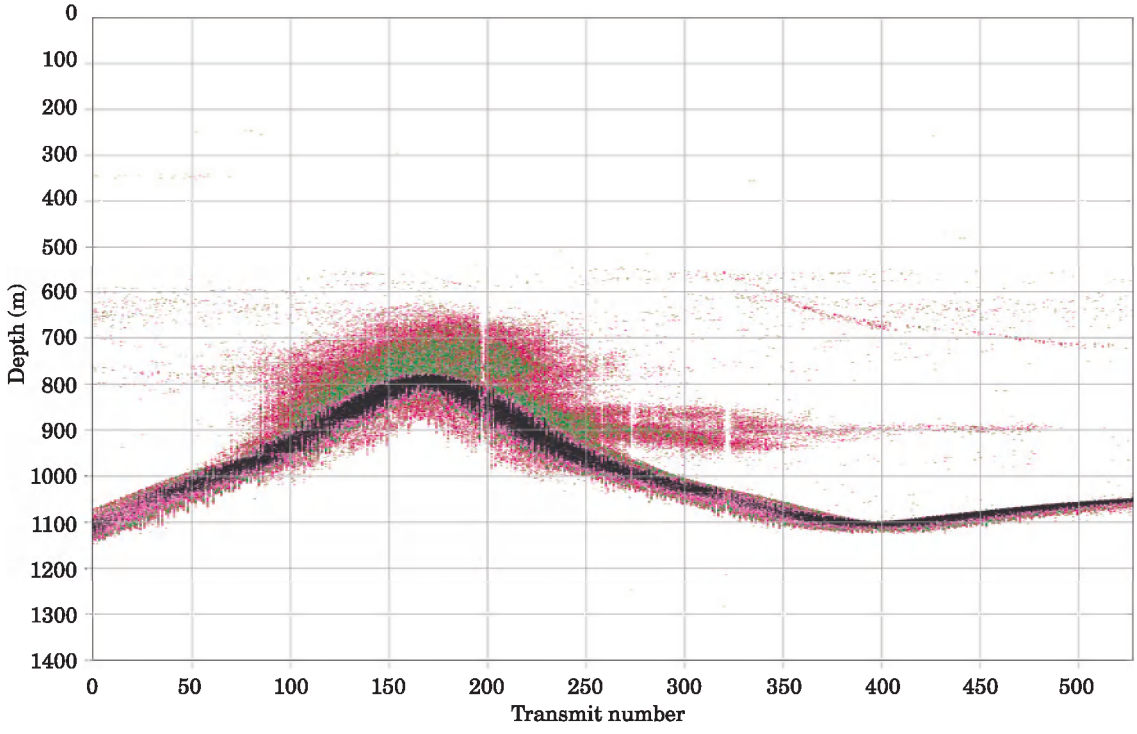

Figure 1. An example of an echogram on Cameron's seamount on the Chatham Rise, July 2000. The distance for 100 transmits is about 600 m.

risk to the equipment. Figure 2 shows both parallel- and star-transect tracks run over Cameron's seamount and it can be seen that stars give the more straightforward transit to the next transect. We have found, in fact, that stars can halve the turning time, yielding time savings of up to 40% and significantly increasing the number of transects that can be done. Josse *et al.* (1999) used star surveys for similar reasons in general and because of the need to achieve adequate coverage of aggregations in a limited time ($\sim$2 h) in particular.

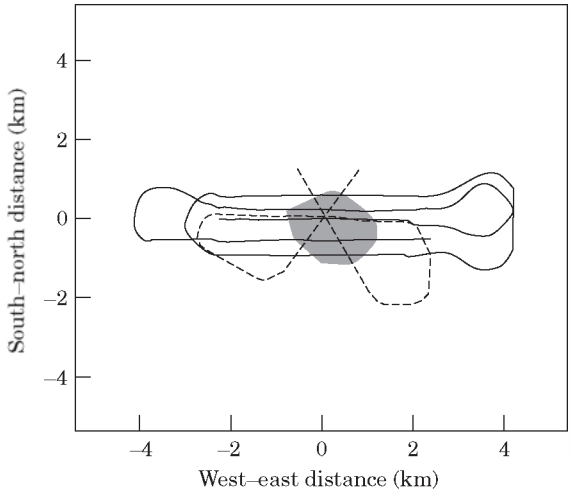

Figure 2. An example of a ship's tracks for star (2000 survey, dotted line) and parallel (1998 survey, solid line) transects on Cameron's seamount [centred on point (0,0)] on the north-east Chatham Rise. Shaded area is the seamount base at 1050 m; the top is at 784 m.

The precise geometry of the star is important in the subsequent analysis as will be discussed later. However, using a towed transducer inevitably means that it is very difficult to steam an accurate star in which all transects are evenly spaced and intersect at the same point. The transducer is typically towed on up to 2 km of cable and the effects of wind and currents mean that it is often offset from the vessel track. Thus, although the vessel may cross the centre of the school the towed transducer may not. Some compensation of the vessel track is possible so that the towed body passes close to the centre each time. In some instances we have exaggerated this effect with the aim of randomizing the transects with respect to the aggregation. The resulting design is the "offset star" in which transects cut through the school at random points but still on approximately equally spaced bearings (Figure 3). Even with perfect tracking any movement of the school will also induce an offset-star effect.

Thus, although an ideal star can be achieved using a hull-mounted transducer, if a towed system must be used

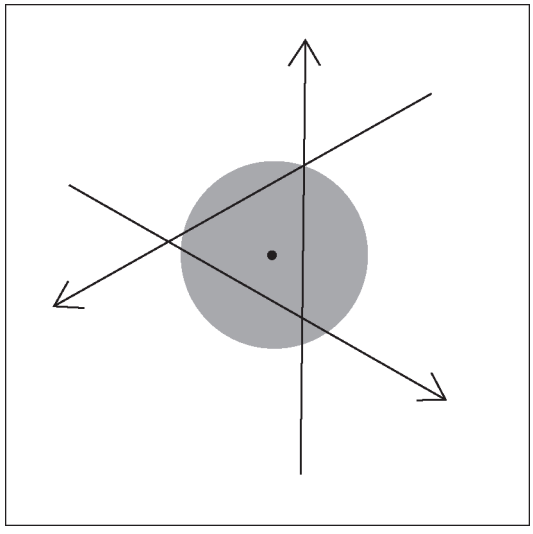

Figure 3. Transects for an offset star on a circular school.

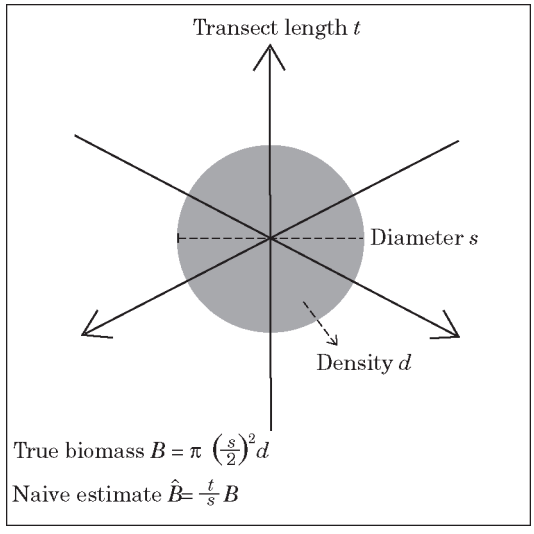

Figure 4. The naive estimate of biomass from a star on a circular school of constant density is biased high.

the ideal star can only be approximated. In this paper we derive four biomass estimators for star surveys and test them with a simulation study designed to mimic an orange roughy survey. These estimators are tested on both ideal and offset stars and we attempt to determine whether the latter provides an adequate approximation to the former for analytical purposes. We also use the simulations to compare star and parallel designs.

Methods

Biomass estimators from star surveys

The problem

Our initial "naive" approach to analysing stars was to treat them as if the transects were parallel. However, this is generally not valid except in the extreme case where the areal density is uniform within the school and the stratum area fits over the school exactly. There are two problems. First, if the centre of the star is placed at the centre of a fish school then density trends within the school will cause bias because the star over-samples the school centre. For example, if the density at the centre is relatively high the naive approach produces estimates that are biased high. The second problem is that the estimate of biomass increases as the stratum area increases and this introduces an increasingly positive bias. This can be demonstrated mathematically for a circular school of uniform density (Figure 4). If t is the transect length, s the school diameter, and d the weight density within the school, then the true biomass is

$$B = \text{school area} \times \text{mean weight density in school}$$

$$= \pi(s/2)^2 d$$

and the naive estimate is

$$\hat{B} = \text{star area} \times \text{average of transect mean densities}$$

$$= \pi(t/2)^2 (sd/t)$$

$$= (t/s)B.$$

The naive estimate has the undesirable property of being proportional to the transect length and there is a severe upwards bias if the transects are substantially longer than the school diameter. Thus, in general, this estimate is only correct if the transect length is equal to the diameter of the school and the weight density within the school is uniform.

Subsequently we adopted the intuitive approach of weighting data on the basis of position on the transect relative to the centre (Bull *et al.*, 2000), which avoids the biases of the naive approach. Here we formally derive this method mathematically by treating the transects as parallel in polar coordinates. We then propose an alternative estimate of variance using a polar-coordinate version of transitive kriging.

Another approach to analysing stars is to estimate the school area and the density within the school separately. We consider two approaches to this; first, a basic method in which the transects are truncated at the school edge – which defines the area – and using the density from the truncated portion and second, a two-stage kriging procedure in which the school bound is initially estimated using "indicator kriging" and then the density within the school is estimated using kriging.

The four estimators for biomass (polar, polar with transitive kriging, basic, and two stage kriging) were

compared using simulated data based on a series of idealized orange roughy aggregations with variations that are typically encountered on surveys. Figure 1 shows an echogram of a typical orange roughy aggregation on a seamount of the type reflected in the simulations. No attempt has been made to compare estimators over a broad range of conditions.

Basic method

A simple estimation method is to truncate transects at the edge of the school, estimate school biomass from each transect as area × mean density and average the results. Thus, if there are n transects, and after truncation the ith transect has length l_i and mean density d_i, the biomass estimate is:

$$\hat{B} = \frac{1}{n}\sum_{i=1}^{n}\left(\pi\left(\frac{l_i}{2}\right)^2 d_i\right).$$

This corrects the bias from areal expansion of biomass, but school density must be uniform for there to be no bias.

Kriging method

Kriging is a prediction method that explicitly takes spatial correlation into account and which divides the variation into large-scale (surface or trends) and small-scale variation (correlation and error variance). Cressie (1991) describes the basic theory. Several flavours exist. "Ordinary" kriging, to start with the simplest, assumes a constant mean over the region. This leads on to "universal" kriging, which estimates a mean surface and then to more complicated analyses. Spatial correlation is assumed to have some structure which is usually expressed in a semivariogram, that can be modelled as a function of the relative distance between two points, i.e., the correlation does not depend on location, at least in the estimation window around a point used to predict its value. Decomposing variation into large-scale and small-scale is not, in general, identifiable so that one person's surface trend can be another's correlation structure. Thus, using universal kriging need not necessarily work better than ordinary kriging even when a trend is involved. For example, Altman (2000) showed that ordinary kriging performed more or less as well as universal kriging in terms of mean squared error of predictions for some one-dimensional data sets she analyzed. Cressie (1991) gives an account of analyses on the wolfcamp aquifer data which shows qualitatively similar predicted surfaces of the piezometric head whatever method is used, i.e., ordinary, universal or median-polish kriging. Predictions are often done at a point but can also be for the total amount in a cell (block kriging). The correlation structure for the latter is linked to that for point predictions but it is modified to represent the average correlation within a cell and between cells.

For biomass estimation we used a two-stage method in which the extent of the mark was first estimated and then the mean density within the mark. Following Altman (2000) and Cressie (1991), we have only considered ordinary kriging here. The extent of the marks is estimated using indicator kriging (Rossi *et al.*, 1992). The acoustic measurements are categorized as MARK or NON-MARK on the basis of the positions of marks visible in the echogram. The indicator function I_{in_mark} is defined as 1 if in mark and 0 if not. The semivariogram of I_{in_mark} is estimated using the classical estimator of Matheron (1965) and a spherical semivariogram model with nugget is fitted using an approximate weighted-least-squares procedure (Cressie, 1991 pp. 61, 97). The survey area is divided into a square grid and ordinary kriging is carried out to estimate the probability that each square is in a mark. Squares for which the kriged probability is at least 0.5 are designated as MARK. In the second stage the mean density of the MARK squares is estimated using ordinary block kriging (Cressie, 1991 pp. 124). The estimated semivariogram of the biomass densities includes only "mark" measurements. As in the first stage a classical semivariogram estimator is used and a spherical variogram with nugget is fitted using weighted least squares. Block kriging is used to estimate the mean density over all of the grid squares identified as being MARK. The mean density is then multiplied by the total area of these squares to yield a biomass estimate.

Variance estimation for the kriging biomass estimate is problematic: the block-kriging process yields a variance but this does not take into account the error in the indicator krig or in the estimation and fitting of both semivariograms. Bootstrapping does not provide an alternative since no spatial-bootstrapping method is known (Cressie, 1991).

Exponential semivariogram models with nugget were also investigated but were found to fit simulated data poorly. In addition, we tried carrying out the density krig on the log-scale with bias correction, Cressie (1991), but preliminary investigations showed that the performance was relatively poor. In addition, the block-kriging method cannot be used on the log-scale, and it is not clear how else the variance for the estimated mean density in the mark could be estimated.

Polar method

In the polar method each density measurement is weighted proportionally to its distance from the centre of the star. A circular survey area is defined, centred on the centre of the star, with the smallest radius necessary to include the entire fish school. Transects which stop short of the edge of the survey area are padded at the ends with zero backscatter values until they reach the

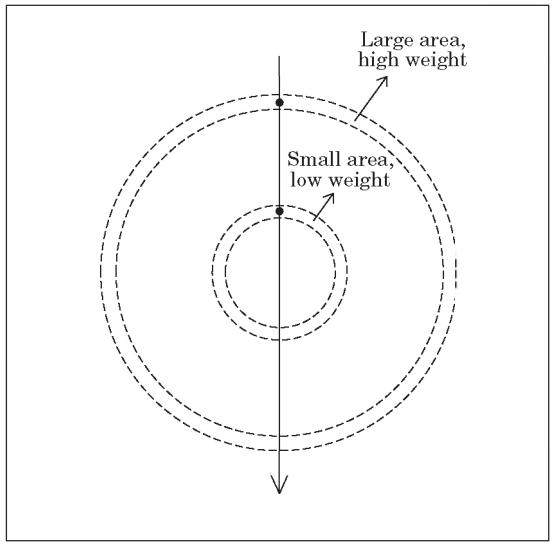

Figure 5. The polar method of estimating biomass for stars. Each backscatter measurement (black dots) is a sample from a ring (dashed circles) and is weighted in proportion to the area of the ring.

edge. If y_{ij} is the ith density measurement in the jth transect and r_{ij} the corresponding distance from the centre of the star, the weighted-mean biomass density for the transect, w_j, is given by

$$w_j = \sum_{i=-m}^{m} y_{ij}\left(\frac{r_{ij}}{\sum_{k=-m}^{m} r_{kj}}\right)$$

where we assume that there are m samples from the centre out to the edge, defining $r_{0j}=0$, and that we have nominated one side of the centre to take positive i indices and negative indices on the other side. If A is the size of the survey area and n the number of transects then the biomass estimate is:

$$\hat{B}=A\bar{w} \qquad (1)$$

and its variance is estimated by

$$\hat{V}(\hat{B})=A^2\hat{V}(\bar{w})=A^2\hat{V}(w)/n.$$

Each density measurement y_i is a sample from a circle of radius r_i and some thickness Δ (Figure 5). The area of the circle is $\pi(r_i+\Delta/2)^2-\pi(r_i-\Delta/2)^2$, which is proportional to r_i. Hence the measurements are weighted proportionally to the areas they sample from. In practice, the y_is are averaged over short sections of the transects rather than calculated for individual acoustic pings. The towed-body positions, and hence the r_i, are not known exactly. We estimate the path of the towed body by assuming that it passes over the centre of the star and that it moves at a speed equal to that of the towing vessel at any time.

The polar method is best understood by visualizing the star pattern in polar coordinates. In an ideal star, where all transects pass over the same centre point, the transects are parallel and evenly spaced in polar coordinates. Each straight-line transect in Cartesian coordinates becomes two parallel transects in polar coordinates, separated by 180° (Figure 6). As a result we can apply results for parallel transect designs to the analysis of stars. The biomass estimate for parallel transects can be recast as the double summation of the estimated biomass density over the length and width of the stratum area, i.e., as a 2-dimensional numerical integration. The same summation carried out in polar coordinates gives the polar method biomass estimate for a star.

Consider a rectangular stratum with n evenly spaced parallel transects each with m evenly spaced, biomass-density measurements y_{ij}. The biomass estimate is

$$\hat{B}=\Lambda\frac{1}{n}\sum_{j}\left(\frac{1}{m}\sum_{i}y_{ij}\right).$$

Let L and W be the dimensions of the stratum, so that $A=LW$. The stratum can be divided into nm rectangular cells each containing one density measurement, with length $\Delta l=L/m$, width $\Delta w=W/n$, and area $a=\Delta l\Delta w$. The estimate can then be recast as a double summation over cells, using each density measurement as an estimate of the mean density in the cell:

$$\hat{B}=\sum_{j=1}^{n}\sum_{i=1}^{m}(y_{ij}\cdot\Delta l\Delta w)$$

The true biomass is the equivalent double integral,

$$B=\int_0^W\int_0^L y(l,w)\,dl\,dw.$$

The same summation can be carried out in polar coordinates for a star. In polar coordinates, the true biomass is given by

$$B=\int_0^{2\pi}\int_0^R r\,y(r,\theta)\,d\theta\,dr$$

where R is the radius of the survey area, θ the angle, and r the radius. If there are 2n radial transects (from n star transects), and m evenly spaced density measurements, there will be 2nm rectangular cells in polar coordinates (r,θ). The spacing between measurements on a transect is R/m and the angle between evenly spaced radial transects is

$$\frac{2\pi}{2n}=\frac{\pi}{n}.$$

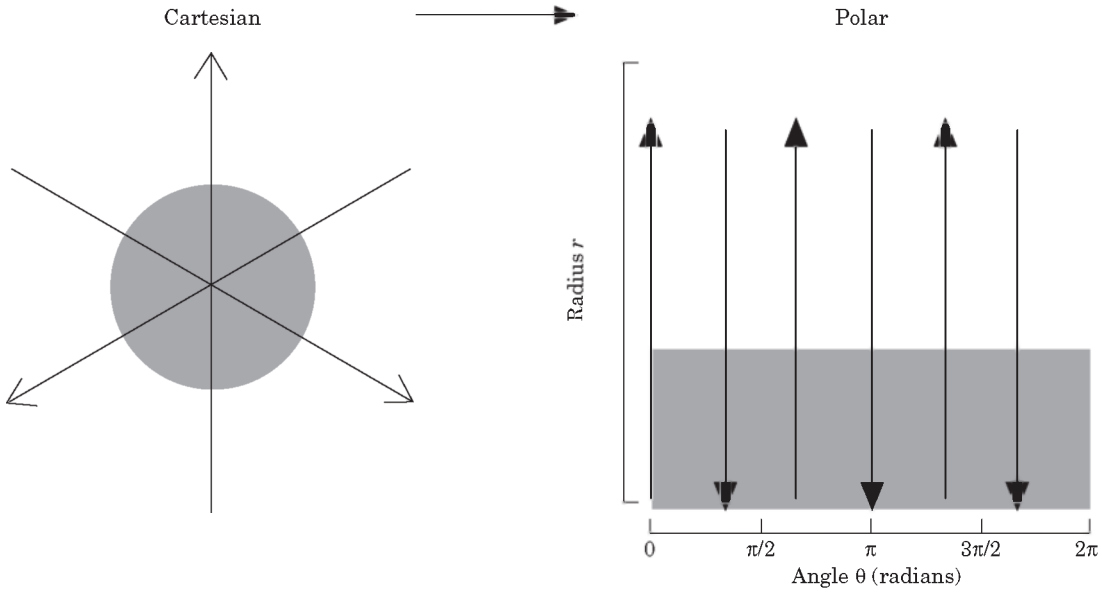

Figure 6. Representation of a star (left) in polar coordinates (right). Each star transect is converted to two radial transects. The radial transects are parallel and evenly spaced in polar coordinates.

The approximate biomass for the ij-th cell, using y_{ij} to approximate y(rθ) in the cell, is:

$$\int_{j\frac{\pi}{n}}^{(j+1)\frac{\pi}{n}} \int_{r_i-\frac{R}{m}}^{r_i+\frac{R}{m}} r\, y_{ij}\, d\theta\, dr$$
$$= y_{ij}\left[(j+1)\frac{\pi}{n} - j\frac{\pi}{n}\right]\frac{1}{2}\left[\left(r_i+\frac{R}{2m}\right)^2 - \left(r_i-\frac{R}{2m}\right)^2\right]$$
$$= y_{ij} r_i \frac{\pi R}{nm}$$

The total biomass is the sum over all the cells:

$$\hat{B} = \sum_{j=1}^{2n}\sum_{i=1}^{m}\left(y_{ij}\cdot\frac{\pi r_i R}{nm}\right)$$

and, since

$$R = 2\sum_{j=1}^{m} \Big/ r_j m,$$

multiplying the above by R/R gives

$$\hat{B} = \pi R^2 \frac{\sum_{j=1}^{2n}\sum_{i=1}^{m}\left(y_{ij} r_i \Big/ \sum_{k=1}^{m} r_k\right)}{2n}$$

Combining the 2n radial transects back into n star transects, nominating one to take negative indices, and noting that

$$\sum_{i=1}^{m} r_i = \frac{1}{2}\sum_{i=-m}^{m} r_i,$$

the biomass is also given by

$$\hat{B} = \pi R^2 \frac{\sum_{j=1}^{n}\sum_{i=-m}^{m}\left(y_{ij} r_i \Big/ \sum_{k=-m}^{m} r_k\right)}{n}$$

which is the polar-method biomass estimate (Equation 1).

A corollary is that the polar method is not valid for offset stars. If transects do not pass over the centre of the circle then they are not radial in polar coordinates (Figure 7), and the centre of the survey area is not represented in the double summation.

The double-summation method also requires that there be no missing data – the entire circular survey area must be filled with integration cells. We therefore truncate long transects and pad short transects with zeros so that the transects are all diameters of the same circle.

Polar method with kriging

Transitive kriging is a technique used to estimate biomass over an area that has been sampled on a regular one- or two-dimensional grid from a randomly distributed starting point (Petitgas, 1993; Bez *et al.*, 1997). The advantage of the method is that the variance estimates take account of the correlation structure of the

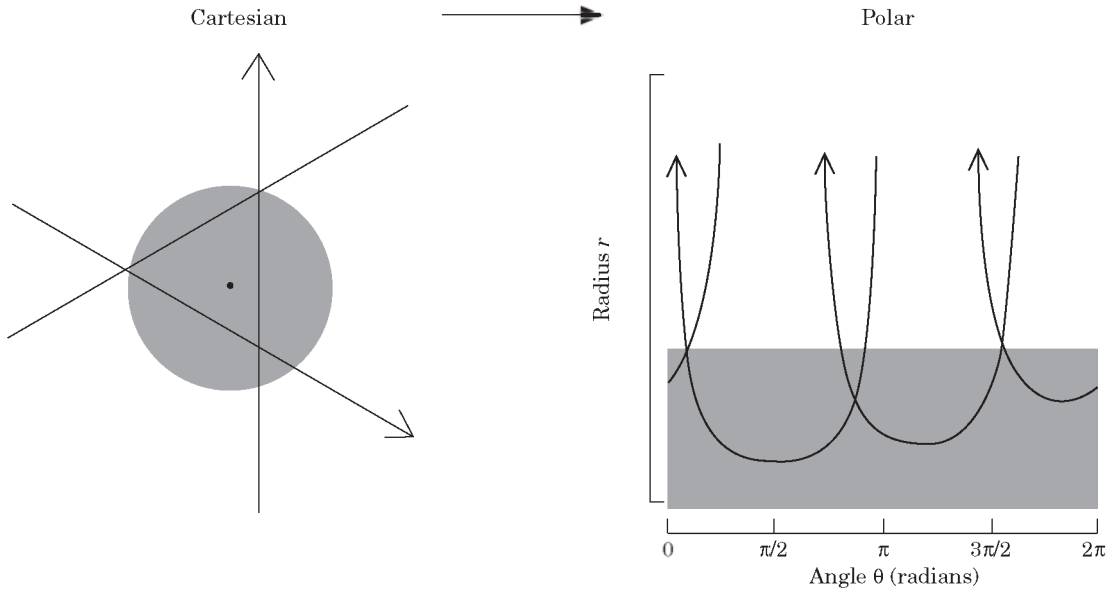

Figure 7. Representation of an offset star (left) in polar coordinates (right).

spatial distribution of the fish. Here, we develop the one-dimensional transitive kriging method in polar coordinates and apply it to derive variance estimates for the polar-method biomass estimator. The estimates take account of the correlations between adjacent transects caused by the spatial structure of the fish schools. Note that the biomass estimate is the same as that for the polar method above but the estimate of the variance is different.

If an area is sampled on a one-dimensional grid, e.g., values on a line that are made up of summations along a transect from a series of parallel transects, with spacing Δx and origin x_0, and $z(x)$ is the biomass density at x, the transitive kriging biomass estimate (Petitgas, 1993) is given by

$$\hat{B} = \Delta x \sum_{i=-\infty}^{\infty} z(x_0 + i\Delta x).$$

(The sum to infinity is used on the assumption that density is zero outside the survey area.) The true biomass is given by

$$B = \int_{-\infty}^{\infty} z(x)\, dx.$$

The correlation structure of the spatial distribution of the fish is expressed in terms of the covariogram

$$g(h) = \int_{-\infty}^{\infty} z(x)z(x+h)dx.$$

The integrated covariogram is equal to Q^2:

$$\begin{aligned}\int g(h)dh &= \iint z(x)z(x+h)dx\, dh \\ &= \int z(x)\left[\int z(x+h)dh\right]dx \\ &= B^2\end{aligned}$$

The variance of $\hat{B}$ is estimated, on the assumption that x_0 is uniformly randomly distributed within origin grid cell, as

$$\begin{aligned}V(\hat{B}) &= \frac{1}{\Delta x}\int_{-\infty}^{\infty} (B - \hat{B}(x_0))^2\, dx_0 \\ &= \left[\frac{1}{\Delta x}\int_{-\infty}^{\infty} (\hat{B}(x_0)^2)\, dx_0\right] - B^2 \\ &= \left[\frac{1}{\Delta x}\int_{-\infty}^{\infty} \left(\Delta x \sum_{i=-\infty}^{\infty} z(x_0 + i\Delta x)\right)\right. \\ &\quad \left.\left(\Delta x \sum_{i=-\infty}^{\infty} z(x_0 + i\Delta x)\right)dx_0\right] - \left(\int g(h)dh\right)^2 \\ &= \Delta x \sum_{i=-\infty}^{\infty} g(i\Delta x) - \int g(h)dh.\end{aligned}$$

The one-dimensional transitive kriging can be applied to the weighted-mean density estimates for the radial transects in the polar method. If the n radial transects are on regularly spaced bearings then they form an evenly spaced, one-dimensional grid in θ-space, with $\Delta\theta = 2\pi/n$. It could be expected that a pair of radial transects on similar bearings would yield more similar

results than a pair of widely spaced transects, and hence that the n weighted-mean density estimates would be correlated. The transitive kriging takes account of these correlations in its variance estimates.

Let $w(\theta)$ be the weighted-mean density for the radius at angle θ, so the observed radial transect densities are $w_1=w(\theta_0)$, $w_2=w(\theta_0+\Delta\theta)$ . . . $w_n=w(\theta_0+(n-1)\Delta\theta)$. Further, let $z(\theta)=w(\theta)A/2\pi$. Then the one-dimensional transitive kriging estimate on z is:

$$\hat{B}=\frac{2\pi}{n}\sum_{i=1}^{n} w(\theta_0+(i-1)\Delta\theta)$$
$$=A\bar{w}$$

which is the polar method biomass estimate $\hat{B}$. Now let $g(\theta)$ be the covariogram of z. The function $g(\theta)$ must be periodic, i.e., $g(\theta)=g(\theta+2\pi)$, and symmetric about 0. The limits of integration change accordingly:

$$\int_0^{2\pi} g(\theta)\,d\theta=B^2$$

and the variance of the biomass estimate is then

$$V(\hat{B})=\frac{2\pi}{n}\sum_{i=0}^{n-1} g(i\Delta\theta)-\int_0^{2\pi} g(\theta)\,d\theta.$$

An estimate of the covariogram is required to evaluate the variance. The empirical covariogram is estimated for angles which are integer multiples of the angle between transects,

$$\hat{g}(i\Delta\theta)=i\Delta\theta\sum_{k=1}^{n} z_i z_{(i+k)\bmod n}$$

where the $(i+k)$ mod n notation keeps the index to 1 through to n and because the z are periodic it indexes the same value as $i+k$.

The first term of the variance estimate is estimated from the empirical covariogram. For the second term it is necessary to estimate $g(\theta)$ for other angles so a parametric model is fitted to the empirical covariogram. The model should be periodic, positive definite, and the largest value, although other points can have this value, should occur at $\theta=0$. We use the cosine model:

$$g(\theta)=a_1+a_2\cos(a_3\theta)$$

which was found to give good fits to the simulated data. The model is fitted to $\hat{g}(k\theta)$ for k from 1 to n/2, using least squares. The fitting process does not use $\hat{g}(0)$ because it is inflated by the "nugget effect" i.e. small-scale random variation superposed on the correlation structure).

Simulation testing

A simulation study was carried out to test the performance of the four estimators of biomass described in this paper:

- polar method
- polar method with c.v.s calculated using polar transitive kriging
- basic method
- two-stage kriging.

The estimators were evaluated using simulated data. Biomass-density measurements were simulated at 10-ping intervals along star or offset-star transects. A set of simulation models of increasing complexity and "reality" was defined. We anticipated that all estimators would perform well with the simplest simulation models but that their performance would degrade to varying extents as the models became more complex and the estimation became more challenging. An estimator could be considered robust if it performed well with the most complex models.

The models formed a hierarchy, with an extra feature added at each level, in a progression towards real-world complexity. Model 1 was the simplest and most idealized. In models 2–5, the simulated schools became increasingly complex in their size, position and structure. Models 6–8 incorporated random fluctuation and temporal change. The parameters of the models were loosely based on real data from acoustic surveys of New Zealand seamounts (authors' unpublished data).

The models were:

(1) Uniform circular: A single circular school centred on the centre of the star was generated with constant biomass density. The school diameter was 1 km.
(2) Density profile: As (1), but the biomass density was highest at the centre of the school and decreased linearly to zero at its edge.
(3) Elliptical school: As (2), but the school was elliptical. The orientation was random: the lengths of the major and minor axes were 1.4 and 0.7 km.
(4) Off-centre school: As (3), but the position of the school was random. The centre of the school was displaced in a random direction from the centre of the star. The distance between the two was normally distributed with mean 0 and standard deviation 250 m.
(5) 3 small schools: As (4), but three smaller schools were added at random positions. The schools were distributed randomly around the hill at distances uniformly distributed from 0.75 to 1.25 km from the hilltop. They were circular, with areas one third that of the main school. The biomass density at the centres of the small schools was half that at the centre of the main school and decreased to zero at

the edges. The combined biomass of the small schools was therefore half that of the main school. The schools were allowed to overlap.

(6) Variation: As (5), but small-scale random variation was added. A multiplicative random error was applied to the biomass in each 10-ping section. The error was mixture-normally distributed, with a coefficient of variation of 0.3 with probability 0.9, or 1.0 with probability 0.1. This error distribution yields low levels of variation overall, with an occasional large "blip".

(7) Schools move: As (6), but the schools moved between transects. Each school moved in a different random direction, with the main school moving 80 m and each smaller school moving 160 m between each pair of transects.

(8) Transect variation: As (7), but a multiplicative random error was applied to each transect, with c.v. 0.3. This could represent the effects of diurnal variation in fish abundance or vertical distribution, or experimental error, perhaps related to weather conditions.

Sets of simulated biomass-density values for models 1, 5, and 6 are shown in Figure 8, alongside the associated biomass-density profiles. (Models 2–4 are not shown because they are simply steps towards model 5; models 7–8 are not shown because they include changes between transects that are difficult to represent.)

Two survey designs were simulated:

- a star with three 3 km transects;
- an offset star with three 3 km transects, each offset from the centre by a random distance (uniformly distributed on [0, 300] m).

For each of the 16 combinations of model and design, 500 sets of simulated data were generated. Each of the four estimators was applied to each dataset to produce a biomass estimate and c.v. The biomass estimates were expressed relative to the true biomass, which is the integral of the underlying biomass density profile. A value greater than 1 represents an overestimate.

The performances of the three biomass estimators (basic, polar, kriging) are expressed in terms of systematic bias, variability (standard deviation), and overall error (root-mean-squared error, RMSE). For models 6 and 8, additional tests with two and six transects were run.

The four variance estimators (basic, polar, polar with kriging, kriging) are also evaluated in terms of systematic bias, variability (standard deviation), and overall error (root-mean-squared error, RMSE), but only for three-transect designs under models 6 and 8.

A comparison of the star survey with the usual parallel transect design was also made using simulation from model 8 – the most complicated and realistic – with three and six transects. The survey area was defined as a rectangle, with transects running west–east. The north–south width of the survey area was equal to twice the north–south width of the main school at the start of the survey. (This width was chosen to be large enough to include all the schools in most simulations; negative bias could have been reduced slightly by increasing it further but this would have been at the cost of more variability and a higher RMSE.) The transects were evenly spaced from a random starting point. Total backscatter was estimated using the standard Jolly & Hampton (1990) estimator.

Results

The performances of the three biomass estimators (basic, polar, kriging) are summarized in Table 1. All three estimators performed well for the basic model 1: it would be of some concern if they did not. For offset stars the basic and polar methods were biased low.

Model 2, in which the density of the school decreased from the centre to the edge, introduced some problems for the basic and polar methods. The basic method had a large positive bias for true stars. This bias occurred because all the transects passed through the centre of the school where the density was highest. The polar and kriging methods are designed to cope with oversampling and did not suffer this bias. However, for offset stars, the polar method was biased low because transects can miss the centre, in which case this area is not represented in the weighted-mean density. This negative bias in the polar method for offset stars appears in all subsequent models.

By model 3, in which the school was elliptical, all estimators had developed a small amount of variability for stars and a moderate amount for offset stars. No new biases were introduced.

In model 4, where the school's position could be off-centre, the bias of the basic method for star designs was reduced since the transects need no longer pass through the centre of the school. On the other hand, the basic method gave highly variable estimates for model 4 and all higher models. The kriging method still performed well and the polar method performed well for stars but was biased low for offset stars, though less so than in the simpler models.

Model 5, which included three small additional schools, led to increased variability in all estimators since transects are likely to miss some schools entirely. No new biases were introduced.

The introduction of random variation in measured densities in model 6 had little effect on the basic and polar method but kriging performance was considerably worsened with increased variability and positive bias.

Model 7, in which schools could move, had little effect on any of the estimators.

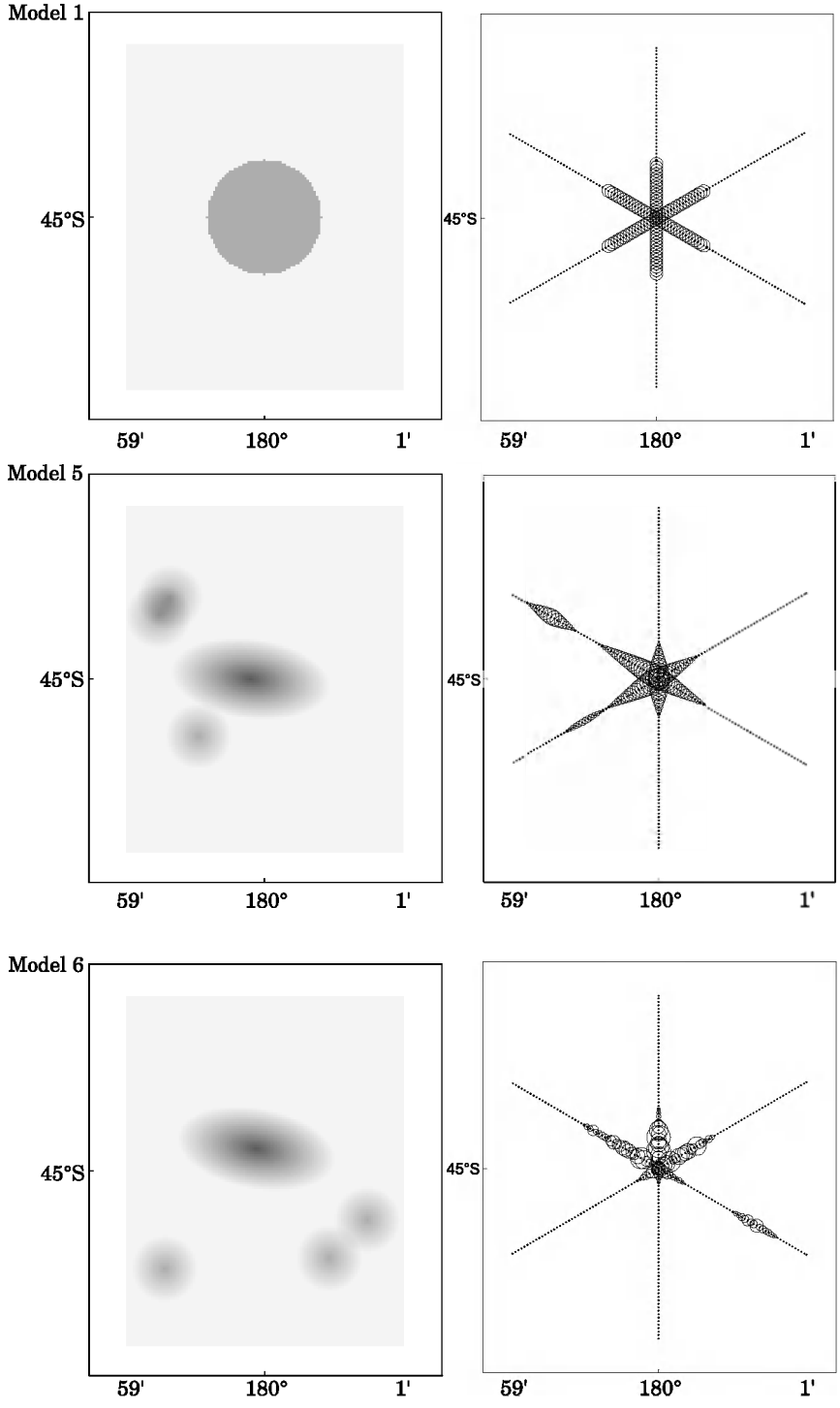

Figure 8. Comparison between simulation models 1, 5, and 6. Left-hand plots show the underlying backscatter-density profiles. Right-hand plots show simulated backscatter data for 3-transect star designs; circle diameters indicate mean-backscatter values in 10-ping sections. Model 1 is the basic, simplified version; model 5 allows a more realistic school structure and adds three small peripheral schools; model 6 includes random variation about the underlying density.

Table 1. Biomass estimate: performance of the basic, polar, and kriging estimators on simulated data using three transects, expressed in terms of bias, variability (standard deviation), and root-mean-squared error (RMSE). Based on 500 simulations per design per model. Note that the results for the polar method with kriging are the same as those for polar.

Estimator design		Basic		Polar		Kriging	
Model	Model name	Star	Offset	Star	Offset	Star	Offset
(a) Bias							
1	Uniform circular	−0.01	−0.13	0.00	−0.12	−0.07	0.01
2	Density profile	0.49	0.06	0.01	−0.25	0.04	0.01
3	Elliptical school	0.46	0.00	0.00	−0.24	0.03	0.04
4	Off-centre school	0.09	−0.14	0.00	−0.16	0.04	0.02
5	Three small schools	0.17	−0.01	0.01	−0.10	0.03	0.04
6	Variation	0.14	−0.03	−0.02	−0.09	0.14	0.11
7	Schools move	0.12	−0.03	0.00	−0.08	0.17	0.13
8	Transect variation	0.12	−0.09	0.00	−0.12	0.20	0.09
(b) Standard deviation							
1	Uniform circular	0.00	0.06	0.00	0.06	0.00	0.06
2	Density profile	0.00	0.18	0.00	0.12	0.00	0.06
3	Elliptical school	0.08	0.25	0.05	0.16	0.07	0.16
4	Off-centre school	0.36	0.34	0.10	0.26	0.11	0.20
5	Three small schools	0.39	0.39	0.25	0.29	0.26	0.29
6	Variation	0.41	0.41	0.25	0.31	0.37	0.37
7	Schools move	0.41	0.38	0.28	0.30	0.38	0.36
8	Transect variation	0.47	0.44	0.35	0.35	0.49	0.42
(c) RMSE							
1	Uniform circular	0.01	0.15	0.00	0.14	0.07	0.06
2	Density profile	0.49	0.19	0.01	0.27	0.04	0.06
3	Elliptical school	0.46	0.25	0.05	0.29	0.07	0.16
4	Off-centre school	0.37	0.36	0.10	0.30	0.12	0.20
5	Three small schools	0.42	0.39	0.25	0.31	0.26	0.30
6	Variation	0.43	0.41	0.25	0.32	0.39	0.38
7	Schools move	0.43	0.39	0.28	0.31	0.41	0.38
8	Transect variation	0.48	0.45	0.35	0.37	0.53	0.43

Table 2. Biomass estimation: performance of the basic, polar, and kriging estimators for 2-transect designs. Note that the results for the polar method with kriging are the same as those for polar.

Estimator design		Basic		Polar		Kriging	
Model	Model name	Star	Offset	Star	Offset	Star	Offset
(a) Bias							
6	Variation	0.15	−0.05	−0.01	−0.10	0.30	0.22
8	Transect variation	0.11	−0.07	−0.01	−0.12	0.30	0.22
(b) Standard deviation							
6	Variation	0.47	0.47	0.34	0.39	0.51	0.54
8	Transect variation	0.56	0.53	0.45	0.46	0.64	0.61
(c) RMSE							
6	Variation	0.50	0.47	0.34	0.40	0.59	0.59
8	Transect variation	0.57	0.53	0.46	0.46	0.71	0.66

The final addition, in model 8, of between-transect variability added substantial variability to each of the estimators.

In the simulations with two transects (Table 2), for models 6 and 8 none of the methods did well and all of them had high variability as one might expect. The kriging method especially was problematic and had substantial bias. With 6 transects (Table 3), all methods did substantially better.

The question of whether the estimated variances of the estimators accurately reflect the true errors was considered. If they do the estimated c.v.s for each model

Table 3. Biomass estimation: performance of the basic, polar, and kriging estimators for 6-transect designs. Note that the results for the polar method with kriging are the same as those for polar.

Estimator design		Basic		Polar		Kriging	
Model	Model name	Star	Offset	Star	Offset	Star	Offset
(a) Bias							
6	Variation	0.15	−0.04	−0.01	−0.11	0.10	0.09
8	Transect variation	0.01	−0.12	0.00	−0.06	0.15	0.11
(b) Standard deviation							
6	Variation	0.30	0.31	0.11	0.20	0.22	0.24
8	Transect variation	0.34	0.37	0.26	0.31	0.33	0.35
(c) RMSE							
6	Variation	0.34	0.31	0.11	0.23	0.26	0.24
8	Transect variation	0.34	0.39	0.26	0.31	0.36	0.36

and design should be close to the true RMSE. Estimated and true variabilities are compared in Figure 9, for models 6 and 8 – two of the more complex, realistic scenarios. Here, the polar method with kriging was generally best (Table 4), in terms of RMSE with the next-best estimate being up to 19% worse. Within the polar method the transitive krig estimate was 18% better on stars and 10% on offset stars (Table 4). The 2-phase kriging estimate of variance was always biased low at about half the true value.

Biomass estimates from the parallel-transect surveys for model 8 did substantially worse than the polar method when using three transects but the two performed similarly when using six transects (Tables 2, 3, & 5). Variance estimates for the parallel design were biased upwards, even with six transects and compared poorly with the polar values (Table 5 and Figure 9).

Discussion

To the best of our knowledge, this is the first use of star-transect patterns and of polar transitive kriging respectively to estimate the biomass of fish. Similar designs have been used to track tuna under FADs and they may have been used for exploratory mapping, but not for biomass estimation *per se*. The polar method gave the best results for the simulated orange roughy surveys for both stars and offset stars when modest amounts of variation were introduced. The variability in star estimates was mainly due to multiple schools (model 5) and transect variation (model 8). For offset stars the main sources of variability were the density profile (model 2), and transect variation (model 8). Thus, the simulation results show that the polar method is reasonably robust to the star centre being displaced from the school centre and the effects of random, fish-aggregation movement. There is a modest advantage in using transitive kriging to estimate variance for the polar method (Table 4). However, this was only the case for simulations in which the star centre did not coincide with the school centre which introduces some asymmetry into the mean values in the transects as they traverse around the star centre. If the centres do coincide then this method had no advantage because the simulated aggregation is isotropic.

Both types of star performed better than parallel transects when a low number of transects were used but the differences were minor when modest numbers were employed. However, the variance estimates from offset stars had a high bias and so were more poorly estimated than those from stars. The poor variance estimation from parallel transects could be improved by transitive kriging because the density has structure in the simulations, i.e., the highest density is in the middle of the aggregation decreasing towards the edges. This structure creates additional apparent variation over the aggregation that is incorporated into the sample variance estimate. Transitive kriging or, perhaps, kriging will tend to correct for the structural change and eliminate the apparent variation.

Because of the time constraints imposed by short spawning seasons and large areas needing to be covered we have tended to use only small numbers, normally between two and four, of transects on seamount surveys (Figure 2). Stars perform best for this number of transects and give better results overall than parallel transects in terms of survey practicality, least survey time and minimum variance,

In practice, as noted earlier, all star transects tend towards offset stars to a greater or lesser extent because of navigational inaccuracies and fish movements. Moreover, as described in the introduction, if a towed transducer is used an offset star is more or less inevitable. The offset star simulated here had a cross-over disc that was 600 m across compared to the 1000 m of the aggregation and was intended to reflect an orange roughy survey situation. For our most

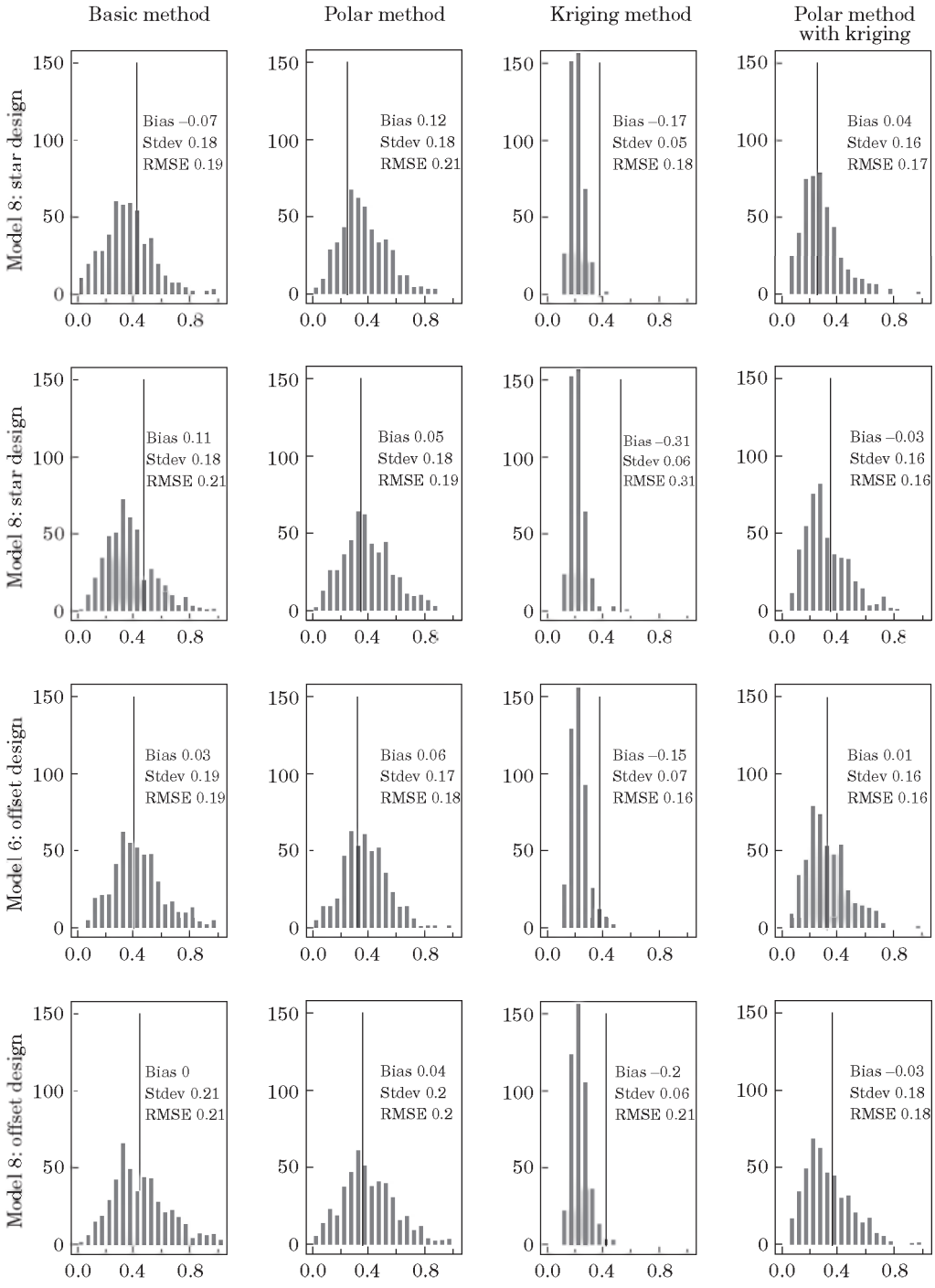

Figure 9. Variance estimation: comparison of estimates with true errors using 3 transects: estimated c.v.s (histogram) vs. true RMSE (vertical line) for each estimator, for models 6 and 8 with stars and offset stars. In the ideal case the histogram would be tightly distributed about the line. Bias, standard deviation, and RMSE of the c.v. estimates are given for each case.

realistic model (model 8) offset stars had a negative bias compared to true stars but this bias was swamped by the variance component so that the RMSE was similar for both. Hence, for the conditions simulated here using a star with a towed transducer will yield acceptable results.

Table 4. Variance estimation: performance of the polar and the polar with kriging estimators for 3-transect designs.

Estimator design		Polar		Polar with kriging	
Model	Model name	Star	Offset	Star	Offset
(a) Bias					
6	Variation	0.12	0.06	0.04	0.01
8	Transect variation	0.05	0.04	−0.03	−0.03
(b) Standard deviation					
6	Variation	0.18	0.17	0.16	0.16
8	Transect variation	0.18	0.20	0.16	0.18
(c) RMSE					
6	Variation	0.21	0.18	0.17	0.16
8	Transect variation	0.19	0.20	0.16	0.18

Table 5. Biomass and variance estimation: performance of the estimators for parallel design on model 8 using three and six transects.

	Biomass performance		C.v. performance	
	Number of parallel transects		Number of parallel transects	
	3	6	3	6
Bias	−0.12	−0.12	0.29	0.25
Standard deviation	0.48	0.30	0.25	0.20
RMSE	0.49	0.32	0.38	0.32

The simulation results show that at least three star transects should be carried out on each seamount. The precise number will depend on how many seamounts are to be surveyed and for what purpose. Three per seamount is sufficient for a typical orange roughy survey that aims to estimate the total biomass over many seamounts.

Simulations of roughy aggregations showed that the best approach is to transform from Cartesian to polar coordinates and then use standard statistical methods (polar method). The polar method was robust to shifts in the transect centre off the aggregation centre and aggregation movement in a random way, i.e., offset designs can be analysed using the polar method. Variance estimation was best with a polar version of transitive kriging. Stars using the polar method were also better than the usual parallel-transect design when transect numbers were low but the results were similar when six or more transects were used. However, in all cases considered parallel transects consistently overestimated variance. We conclude that star transects, by minimizing vessel time and yielding good precision, offer a robust and effective way of estimating the biomass of small, localized aggregations of fish.

Acknowledgements

Funding for this research was provided by New Zealand Ministry of Fisheries contract ORH199901. Allan Hicks provided helpful comments on the manuscript.

References

Altman, N. 2000. Krige, smooth, both or neither? Australian and New Zealand Journal of Statistics, 42: 441–461.

Bez, N., Rivoirbard, J., and Guiblin, P. H. 1997. Covariogram and related tools for structural analysis of fish survey data. *In* Geostatistics Wollongong '96, Volume 2, pp. 1316–1327. Ed. by E. Y. Baafi, and N. A. Schofield. Kluwer Academic Publishers, Netherlands.

Bull, P., Doonan, I., Tracey, D., and Coombs, R. 2000. An accoustic estimate of orange roughy abundance on the Northwest Hills, Chatham Risc, June–July 1999. New Zealand Fisheries Assessment Report 2000/20. 36 pp.

Cressie, N. A. C. 1991. Statistics for Spatial Data. Wiley, New York. 900 pp.

Depoutot, C. 1987. Contribution a l'etude des Dispositifs de Concentration de Poissons a partir de l'experience polynesienne, ORSTOM Tahiti Notes Doc. Oceanogr. No. 33, 170 pp.

Do, M. A., and Coombs, R. F. 1989. Acoustic measurements of the population of orange roughy (*Hoplostethus atlanticus*) on

the north Chatham Rise, New Zealand, in winter 1996. New Zealand Journal of Marine and Freshwater Research, 23: 225–237.

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

Josse, E., Bertrand, A., and Dagorn, L. 1999. An acoustic approach to study tuna aggregated around fish aggregating devices: methods and validation. Aquatic Living Resources, 12: 303–313.

Matheron, G. 1965. La Théorie des Variables Régionalisées et ses Applications. Masson et Cie, Éditeurs, Paris. 306 pp.

Petitgas, P. 1993. Geostatistics for fish stock assessment: a review and an acoustic application. ICES Journal of Marine Science, 50: 285–298.

Rossi, R. E., Mulla, D. J., Journel, A. G., and Franz, E. H. 1992. Geostatistical tools for modelling and interpreting ecological spatial dependence. Ecological Monographs, 62(2): 277–314.