Multistage cluster sampling design and optimal sample sizes for estimation of fish discards from commercial trawlers

Michelle Allen*, David Kilpatrick, Mike Armstrong, Richard Briggs, Grant Course, Nélida Pérez

*Department of Agriculture and Rural Development, Biometrics Division, Newforge Lane, Belfast, Northern Ireland BT9 5PX, UK
Department of Agriculture and Rural Development, Agricultural and Environmental Science Division, Newforge Lane, Belfast, Northern Ireland BT9 5PX, UK
Centre for Environment Fisheries and Aquaculture Science, Lowestoft NR33 7LH, UK
Centro Oceanográfico de Vigo, Apartado de Correos 1552, 36280 Vigo, Spain

Received 7 August 2000; received in revised form 15 December 2000; accepted 11 April 2001

Abstract

The fish discards surveys, conducted by Northern Ireland, Spain and England under the auspices of EC project 95/094, are multistage in design with a structure of hauls within trips within vessels. The data collected by on-board observers from these surveys enabled investigation into the optimal number of hauls, trips and vessels that require sampling to achieve target levels of precision for each gear type and country. Residual maximum likelihood (REML) was used to estimate the variance components between vessels, between trips within vessels and between hauls within trips within vessels, while examining the significance of the fixed effects of area, quarter and their interaction. To meet the REML requirement that the data are normally distributed, the extreme skewness of the data was corrected by a natural log transformation. The estimated lognormal variance components were used within the multistage formulae to estimate the optimal average sampling levels at each stage. The results showed that more resources than were available during the project were required to achieve a coefficient of variation of 20% for the estimation of discard levels for Northern Ireland twin-rig and Spanish pair trawls. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Fishery surveys; Fish discards; Multistage sampling; Optimum sample size; Variance components; Celtic Sea; Irish Sea; North Sea; Bay of Biscay

1. Introduction

Consider the general case where each unit in a population, referred to as the primary units, can potentially be sampled and further partitioned into a number of sub-units, the secondary units, which likewise can be sampled. If the sub-units within a selected primary unit give similar results then it is uneconomical to measure them all. Hence, greater precision will be obtained by measuring few sub-units over many units rather than vice versa. This procedure is known as two-stage sampling in which the primary unit is not completely measured but is itself sampled. Extending beyond two levels is known as multistage sampling (Cochran, 1960; Sukhatme and Sukhatme, 1970).
Within fisheries research multistage cluster sampling has been applied to determine optimal allocation for age–length keys to improve catch at age estimates, for example Schweigert and Sibert (1983), Lai (1987) and Horppila and Peltonen (1992). The latter highlighted the advantage of multistage sampling, which is to greatly reduce the variation of the estimate while collecting less data. This was achieved by adjusting the original sampling scheme involving measuring 500 fish from each of 10 loads to one involving measuring 34 fish from each of 46 loads.

Another application, due to Tamsett and Janacek (1999) and Tamsett et al. (1999), used discards data from the North and Irish Seas to investigate possible stratification schemes and optimal numbers of fish per basket to be measured, respectively. In both papers, the error distribution in the models used to estimate variance components was assumed to be binomial. Over-dispersion occurred indicating that the residuals were not constant and, as acknowledged in Tamsett and Janacek (1999), there existed “... a wide (logarithmic) variation in quantities of fish caught during trips”. Although the over-dispersion was taken into account when calculating sample sizes, in this study it will be shown how the use of lognormal variance components avoids the need for this.

Estimation of discarding practices of commercial fisheries is important because discarding of both target and non-target species occurs and may cause a reduction in average landings in the longer term. Quantifying discards of commercial fisheries thus enable them to take account of in stock assessments. Surveys to estimate the quantities of fish discarded from commercial fishing vessels typically use multistage sampling comprising up to six levels each of which contribute to the variability:

<table>
<thead>
<tr>
<th>Stage</th>
<th>Units</th>
<th>Population size</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>Vessels</td>
<td>V</td>
<td>v</td>
</tr>
<tr>
<td>Second</td>
<td>Trips</td>
<td>T</td>
<td>t</td>
</tr>
<tr>
<td>Third</td>
<td>Hauls</td>
<td>H</td>
<td>h</td>
</tr>
<tr>
<td>Fourth</td>
<td>Boxes</td>
<td>B</td>
<td>b</td>
</tr>
<tr>
<td>Fifth</td>
<td>Fish length</td>
<td>L</td>
<td>l</td>
</tr>
<tr>
<td>Sixth</td>
<td>Fish age</td>
<td>A</td>
<td>a</td>
</tr>
</tbody>
</table>

where V is the number of vessels in the fleet; T, H, B, L and A are the average number of trips per vessel, hauls per trip, boxes per haul, fish lengths per box and fish aged per length, respectively, in the fleet. For simplicity we have assumed that T, H, B, L and A are constant. The lower case equivalents (v, t, h, b, l and a) are the corresponding numbers in the samples. Under the latest discards project, EC 98/095 which commenced July 1999, Northern Ireland are now collecting data on boxes and it is anticipated that the analysis will be extended to all six levels identified. However, for the present study the available data only permits investigation of the first three stages.

Currently, on-board observers collect the data and during trips of many hauls of short duration, it may be impossible to sample all hauls, particularly when single observers are used. By taking account of the multistage sampling design and partitioning the overall variability over the various stages, optimum sample sizes can be estimated. This then provides guidance to the on-board observers regarding the minimum number of hauls which need to be sampled to achieve a target precision with little further gain attained by measuring any more.

When tackling a sampling problem there are two main statistical issues to be addressed:

1. The probability method to be used to select the units for sampling, that is, selection with equal probability or with probability proportional to size (pps). Theoretically with multistage sampling a combination of probability methods can be used, for example selecting primary units with equal probability and secondary units with pps. However, based on the results of Allen et al. (2001) pps did not give sufficient gain in precision over equal probability selection to justify its extra complexity for the fleets considered in this paper. Hence in this study the vessels, trips and hauls are assumed to be selected with equal probability.

2. The optimal sample size required. This can either be to maximise the precision of the estimate subject to a specified cost, \(C_0\), or to achieve a specified precision expressed as the variance, \(\text{Var}_{\text{target}}\), of the estimate. As cost data were not available it was possible to investigate only the latter option. Rather than use a specified variance it is more
meaningful to consider a range of coefficients of variation (CVs).

Data from the following fleets were used to demonstrate the application of multistage cluster sampling design: Northern Ireland midwater and twin-rig trawlers; Spanish demersal heavy and pair trawls; and English pair trawls, unspecified otter and _Nephrops_ trawls. The Northern Ireland midwater, Spanish demersal heavy and English pair and unspecified otter trawlers target demersal fish. The Northern Ireland twin-rig, Spanish demersal heavy and English _Nephrops_ trawlers target _Nephrops norvegicus_. The Spanish pair trawlers target mainly blue whiting.

2. Methods

2.1. Sources of data

Detailed descriptions of the discards surveys for each of the countries are given in the report to the Commission of the European Communities Directorate General for Fisheries (Contract EC 95/094). Allen et al. (2001) gives a brief overview of the methods used for trip selection. Table 1 details the sample and population data of the number of vessels, trips and hauls by country and gear type. Figs. 1–3 show the haul positions for each gear type, and the area definitions used in the analysis, for Northern Ireland, Spain and England, respectively. The area definitions were provided by the fishery scientist from each research laboratory. Discards data were collected by Northern Ireland on approximately 50 species; by Spain on four spot megrim, megrim, black angler, white angler, hake, blue whiting, mackerel, horse mackerel and _Nephrops_; and by England on cod, haddock and whiting.

Spain had information on only the total number of demersal heavy and pair trips and on the total number of trips for the fleet. It was assumed that the survey data were representative of the fleet when estimating the number of hauls for demersal heavy trips at fleet level. The data gathered on pair trawls are day boats with the mean number of hauls per trip from the sample tending towards one (Table 1). For this reason the data for pair trawls were analysed at trip level and in the instance when a trip trawled in two areas this was recoded as two trips.

2.2. Estimation of the variance components

The basic datum is the number of all fish discarded per hour at haul level except in the case of Spanish pair trawlers when the basic datum is the number of all fish discarded per hour at trip level. Individual species were not considered as the project is concerned with the discarding of non-target as well as target species. The data are highly variable and subject to a number of sources of variation. The standard statistical procedure to analyse such data is analysis of variance (ANOVA) which partitions the total variation in the data into its various sources thus:

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample and population data for the duration of the survey—number of vessels, trips and hauls</td>
</tr>
<tr>
<td>Duration of survey</td>
</tr>
<tr>
<td>April 1997–August 1998</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>January–December 1997</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>June 1997–August 1998</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Estimated number of hauls.

*Data analysed at trip level see Section 2.1 for reason.
From these estimates are obtained of the variance components $\sigma_{VT}^2$, $\sigma_{V}^2$ corresponding to between hauls, between trips and between vessels, respectively. The relative sizes of these provide a measure of the contribution of the various sources to the overall variation in the data.

However, in the current situation where the data are unbalanced with respect to different numbers of hauls per trip and trips per vessel, standard ANOVA cannot be applied. The situation is further complicated by the presence of fixed terms, which contribute to the overall variability. These include quarter, area fished and their interaction. Hence, the residual maximum likelihood (REML) procedure available in Genstat (Payne et al., 1993) was used to analyse the data set. This procedure can be used when data are categorised, unbalanced and subject to variation at different levels or strata. REML analysis can be thought of as a generalisation of ANOVA. If the data are perfectly balanced then the results of REML and ANOVA are exactly the same. The fixed effects are analogous to treatment effects and the random effects to the block effects of an ANOVA (Robinson, 1987; Searle et al., 1992).

There are two stages to REML analysis—first the variance components are estimated, then the treatment means. The variance components are estimated using a maximum likelihood method under the condition that the residual variation is greater than zero and the remaining variance parameters are greater than or equal to zero. The treatment means are estimated by generalised least-squares using varying weights by applying information on the sources of variation. These estimates are based on the assumption that the
Fig. 2. Haul positions of (a) demersal heavy (+) and (b) pair trawls (×) sampled during the Spanish discards survey (January–December 1997) and area definitions.

Fig. 3. Haul positions of (a) pair (×) and Nephrops (+) trawls and (b) unspecified otter trawls (+) sampled during the English discards survey (June 1997–August 1998) and area definitions.
data are normally distributed. The significance of the fixed terms within the model can be assessed using the Wald statistic, defined as the sum of squares for a fixed effect divided by the residual variance. The Wald statistic has an asymptotic chi-squared distribution with degrees of freedom equal to those of the fixed effect under scrutiny. The null hypothesis of the fixed effect being equal to zero is tested. For unbalanced data the Wald statistic is affected by the order of fitting the fixed terms as the sum of squares for a fixed effect is determined ignoring all terms fitted later in the model. To properly assess the significance of the fixed effects, the terms need to be rotated (Payne et al., 1993).

2.3. Application of REML analysis to the data

The discarding rate data, defined as the number of all fish discarded per hour (ndph) at haul level (except in the case of Spanish pair trawlers when it was taken at trip level), are significantly \((P < 0.001)\) positively skewed. A natural log transformation, and in some instances the addition of a constant, was applied to normalise the data and eliminate the skewness (Table 2). The constants were determined whenever skewness of zero was achieved. Separate REML analyses were carried out for each gear type, as the gear types exhibit different discarding practices, and country. The fixed effects of area (Figs. 1–3, as devised for each country by the relevant fishery scientist), quarter, and the interaction area \(\times\) quarter were tested for each gear type and country.

2.4. Optimal sample size

The estimated lognormal variance components were used in the formulae below to calculate the optimal average number of hauls per trip per vessel that require sampling to achieve a target precision. The multistage formulae for equal units at each level were used although the data were unbalanced, with unequal number of trips and hauls for each vessel. This could theoretically be taken into account when calculating optimal sampling allocation. However, this requires access to all individual trips and hauls of those vessels sampled throughout the survey period which was not possible.

The sample mean is given by

\[
\hat{y} = \frac{\sum\sum\sum y_{ij}}{vth}
\]

and the variance is

\[
Var(\hat{y}) = \left(1 - \frac{V}{V}ight) \frac{S^2}{y} + \left(1 - \frac{T}{T}ight) \frac{S^2}{y} + \left(1 - \frac{H}{H}ight) \frac{S^2}{y}
\]

The optimum value for the number of vessels to sample, \(v_{opt}\), corresponding to achieving a target

<table>
<thead>
<tr>
<th>Country</th>
<th>Gear</th>
<th>ndph</th>
<th>ln(ndph)</th>
<th>ln(ndph + d)*</th>
<th>Standard error of skew value</th>
<th>nobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Ireland</td>
<td>Midwater(^b)</td>
<td>2.572***</td>
<td>−2.483***</td>
<td>0</td>
<td>0.243</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>Twin-rig</td>
<td>1.466***</td>
<td>−0.326</td>
<td></td>
<td>0.258</td>
<td>87</td>
</tr>
<tr>
<td>Spain</td>
<td>Demersal heavy(^b)</td>
<td>8.186***</td>
<td>−1.302***</td>
<td>0</td>
<td>0.115</td>
<td>447</td>
</tr>
<tr>
<td></td>
<td>Pair(^b,c)</td>
<td>4.680***</td>
<td>−1.659***</td>
<td>0</td>
<td>0.414</td>
<td>32</td>
</tr>
<tr>
<td>England</td>
<td>Pair(^b)</td>
<td>2.803***</td>
<td>−2.680***</td>
<td>0</td>
<td>0.388</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Unspecified otter(^b)</td>
<td>3.454***</td>
<td>−2.802***</td>
<td>0</td>
<td>0.120</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td>Nephrops</td>
<td>1.358***</td>
<td>0.222</td>
<td>0.524</td>
<td></td>
<td>19</td>
</tr>
</tbody>
</table>

\(^*\) \quad P \leq 0.05; \quad ** \quad P \leq 0.01.

\(^d\) \quad d = 2.71, 5.77, 1.35, 11 and 3.825 for Northern Ireland midwater, Spanish demersal heavy, Spanish pair and English pair and English unspecified otter trawls.

\(^b\) \quad ndph = 0.001 when ndph = 0, ndph is the number of all fish discarded per hour.

\(^c\) \quad Data analysed at trip level see Section 2.1 for reason.
The variance of $V_{\text{target}}$ is

$$v_{\text{opt}} = \frac{S_Y^2 + (1 - (t/T))(S_{VT}^2/t) + (1 - (h/H))(S_{VTH}^2/th)}{\text{Var}_{\text{target}} + (S_Y^2/V)}$$

(3)

The optimum values for $t$ and $h$, which are independent of $V_{\text{target}}$, are

$$t_{\text{opt}} = \sqrt[3]{ \frac{c_t(S_{VT}^2 - (1/H)S_{VTH}^2)}{c_t(S_Y^2 - (1/T)S_{VT}^2) - 1} }$$

(4)

and

$$h_{\text{opt}} = \sqrt[3]{ \frac{c_hS_{VTH}^2}{c_h(S_{VT}^2 - (1/H)S_{VTH}^2)} }$$

(5)

corresponding to minimising the total cost $C$ where $C = c_vv + c_tv + c_vth$ and $c_v$, $c_t$, and $c_h$ are the costs associated with sampling a vessel, trip and haul, respectively (Cochran, 1960; Sukhatme and Sukhatme, 1970).

One remaining problem is that the variance components are on a logarithmic scale and the target CVs are on an arithmetic scale. Hence it was necessary to convert the target variance, $V_{\text{target}}$, to a logarithmic scale. It is possible to show that the squared CV of the mean on the arithmetic scale is approximately equal to the variance of the mean on the logarithmic scale (Finney, 1941; Julious and Debarnot, 2000).

### 3. Results

The chosen models from which the variance components were derived are given in Table 3. The significance when including the fixed effects is assessed by the change in deviance from a basic model comprising the general mean and the random effects. Residuals from the fitted models showed no apparent trends when plotted against actual values. The magnitude of the residuals from the Spanish data indicated greater variability. Table 4 shows the lognormal

<table>
<thead>
<tr>
<th>Country</th>
<th>Gear</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Ireland</td>
<td>Midwater</td>
<td>$\ln(\text{ndph} + 2.71) = a + \text{area} + \text{vessel} + \text{vessel} \times \text{trip} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
<tr>
<td></td>
<td>Twin-rig</td>
<td>$\ln(\text{ndph}) = a + \text{vessel} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
<tr>
<td>Spain</td>
<td>Demersal heavy</td>
<td>$\ln(\text{ndph} + 5.77) = a + \text{vessel} + \text{vessel} \times \text{trip} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
<tr>
<td></td>
<td>Pair*</td>
<td>$\ln(\text{ndph} + 1.35) = a + \text{area} \times \text{quarter} + \text{vessel} + \text{vessel} \times \text{trip}$</td>
</tr>
<tr>
<td>England</td>
<td>Pair</td>
<td>$\ln(\text{ndph} + 11) = a + \text{area} + \text{vessel} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
<tr>
<td></td>
<td>Unspecified otter</td>
<td>$\ln(\text{ndph} + 3.825) = a + \text{area} + \text{vessel} + \text{vessel} \times \text{trip} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
<tr>
<td></td>
<td>Nephrops</td>
<td>$\ln(\text{ndph}) = a + \text{vessel} + \text{vessel} \times \text{trip} + \text{vessel} \times \text{trip} \times \text{haul}$</td>
</tr>
</tbody>
</table>

* Data analysed at trip level see Section 2.1 for reason.

<table>
<thead>
<tr>
<th>Country</th>
<th>Gear</th>
<th>$S_Y^2$</th>
<th>Standard error</th>
<th>$S_{VT}^2$</th>
<th>Standard error</th>
<th>$S_{VTH}^2$</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Ireland</td>
<td>Midwater</td>
<td>0.261</td>
<td>0.257</td>
<td>0.300</td>
<td>0.200</td>
<td>0.378</td>
<td>0.061</td>
</tr>
<tr>
<td></td>
<td>Twin-rig</td>
<td>1.301</td>
<td>0.697</td>
<td>–</td>
<td>–</td>
<td>0.629</td>
<td>0.101</td>
</tr>
<tr>
<td>Spain</td>
<td>Demersal heavy</td>
<td>0.504</td>
<td>0.451</td>
<td>1.160</td>
<td>0.465</td>
<td>2.727</td>
<td>0.197</td>
</tr>
<tr>
<td></td>
<td>Pair*</td>
<td>0.202</td>
<td>1.387</td>
<td>4.198</td>
<td>1.666</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>England</td>
<td>Pair</td>
<td>0.859</td>
<td>0.955</td>
<td>–</td>
<td>–</td>
<td>0.475</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td>Unspecified otter</td>
<td>0.478</td>
<td>0.175</td>
<td>0.157</td>
<td>0.073</td>
<td>0.541</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>Nephrops</td>
<td>0.450</td>
<td>0.424</td>
<td>0.211</td>
<td>0.285</td>
<td>0.124</td>
<td>0.062</td>
</tr>
</tbody>
</table>

* Data analysed at trip level see Section 2.1 for reason.
variance components and standard errors of the variance components.

It was not possible to substitute costs into Eqs. (4) and (5) and use the \( h_{\text{opt}} \) and \( t_{\text{opt}} \) values in Eq. (3) for \( v_{\text{opt}} \). Instead two alternative approaches were used:

1. For the average number of hauls per trip was kept constant at the average from Table 1 and the number of sampled trips per vessel ranging from 1 to 10, Eq. (3) was evaluated for CVs of 10, 15 and 20% to give \( v_{\text{opt}} \).
2. For a constant CV of 20%, both the average number of trips per vessel and hauls per trip were varied and Eq. (3) was again evaluated.

It should be noted that a preferred level of precision is not being advocated in this study. The CVs have been chosen with the intent of illustrating the method only.

The operating curves (Figs. 4–9) for each gear type and country give values for \( v_{\text{opt}} \) using these two approaches where the value given for \( v_{\text{opt}} \) is calculated for the duration of the survey. Once a curve approaches its asymptote there is little precision gain in measuring more trips.

### 3.1. Northern Ireland

The variance component, vessel \( \times \) trip, for twin-rig trawls, was found to be negative, possibly due to the lack of replicated trip data per vessel. This component was removed from the basic model. Therefore, it was necessary to assume that only 1 trip would be sampled for vessels carrying out twin-rig trawls. This reduces the design to a two-stage sampling strategy for twin-rig trawls at the between vessel and the between haul levels. For midwater trawls the fixed effects of area, quarter and area \( \times \) quarter interaction were significant \((P < 0.05)\), however, only the addition of area and area \( \times \) quarter into the basic model provided a very highly significant \((P < 0.001)\) decrease in the change in deviance. For twin-rig trawls although area was highly significant \((P \approx 0.003)\) there was not a significant \((P > 0.05)\) reduction in the deviance from the basic model.

Fig. 4(a), for midwater trawls, shows that a CV of 20% is achieved by sampling five hauls per trip and either one trip for each of 13 vessels or two trips for nine vessels. Hence, overall more trips are required for an average of two trips per vessel than one trip. For two trips per vessel and five hauls per trip it is necessary to sample 32 and 16 vessels, respectively, to achieve target CVs of 10 and 15%. For a target CV of 20%, if one trip per vessel is sampled with an average 10 or 15 hauls per trip then, for midwater trawls 12 vessels should be sampled in each case (Fig. 4(b)). Thus only a small reduction in the number of sampled vessels is required when increasing the average number of hauls per trip from 5 to 10 or 15 for a target CV of 20%.

For twin-rig trawlers, there is again little reduction in the number of vessels when sampling more than three hauls per trip. Hence, if sampling only three hauls per vessel, to achieve target CVs of 10, 15 or

![Fig. 4. Optimal number of vessels for (a) five hauls and varying the average number of trips per vessel and CV and (b) CV of 20% and varying the average number of trips per vessel and hauls per trip for Northern Ireland midwater trawls.](image-url)
20%, a total number of 44, 31 or 23 vessels, respectively, need to be sampled (Fig. 5).

3.2. Spain

For demersal heavy trawls area and area \times quarter were highly significant \((P < 0.01)\), however, these fixed effects did not significantly \((P > 0.05)\) reduce the change in deviance when compared to the basic model although area was borderline significant \((P \approx 0.08)\). For pair trawls area \times quarter was borderline significant \((P \approx 0.05)\) with a highly significant \((P \approx 0.004)\) decrease in the change in deviance from the basic model. The fixed effect quarter for either gear type was not significant \((P > 0.05)\).

For demersal heavy trawlers sampling on an average seven hauls per trip requires either one trip for 39 vessels or two trips for 25 vessels to achieve a CV of 20%. Decreasing the target CV to 10 or 15% requires 86 or 43 vessels, respectively, with two trips per vessel (Fig. 6(a)). If the average number of hauls is increased to 10 or 15 then a target CV of 20% requires 24 or 23 vessels with two trips per vessel (Fig. 6(b)). Thus increasing the average number of sampled hauls per trip from 7 to 10 or 15 for a target CV of 20% gives only a slight reduction in the number of sampled vessels.

Achieving a CV of 20% for pair trawlers requires either one trip for 96 vessels or two trips for 50 vessels representing approximately the same number of sampled trips in both cases. Decreasing the CV to 10 or 15% with two trips per vessel requires 144 or 80 vessels, respectively, to be sampled (Fig. 7).
3.3. England

For pair trawls the variance component vessel × trip is negative, which could be due to the lack of trip replication per vessel within the available data set. This component was removed from the basic model with the result that only the optimal number of hauls per vessel can be examined for this gear type. For pair trawls the fixed effect quarter was very highly significant \( (P < 0.001) \) with a very highly significant \( (P < 0.001) \) decrease in the change in deviance from the basic model. For unspecified otter trawls the fixed effects area and area × quarter were highly significant \( (P \approx 0.004 \text{ and } P \approx 0.003, \text{ respectively}) \) but only the addition of area into the basic model had a significant \( (P \approx 0.003) \) decrease in the change in deviance.

The operating curve for English pair trawls was approaching its asymptote at approximately three hauls per vessel. To achieve CVs of 10, 15 and 20\% when sampling three hauls per vessel requires 16, 13 and 10 vessels, respectively.

For unspecified otter trawls, the target CV of 20\% is achieved when sampling seven hauls per trip and either one trip for 14 vessels or two trips for 13 vessels indicating that more effort is required when sampling two trips per vessel. Decreasing the target CV to 10 or 15\% with one trip per vessel requires 50 or 24 vessels, respectively (Fig. 8(a)). If the average number of sampled hauls is increased to 10 or 15 and one trip per vessel is sampled then 13 vessels are required for either scenario to achieve a CV of 20\% (Fig. 8(b)).

For *Nephrops* trawls the target CV of 20\% requires two hauls per trip and either one for 15 vessels or two trips for 12 vessels. Again more effort is required when sampling two trips per vessel. For two hauls per trip and one trip per vessel, CVs of 10 or 15\% are achieved by sampling 44 or 24 vessels, respectively (Fig. 9).

4. Discussion

4.1. Magnitude of variance components and associated standard errors

In considering optimum sample sizes in multistage cluster sampling, the general principle is that if the sub-units within a selected primary unit give similar results then, for a fixed number of sampled sub-units, greater precision will be obtained by measuring a few sub-units over many primary units. In the current context this is true if the relative magnitude of the variance components decreases from vessel through...
vessel × trip to vessel × trip × haul. This would also be reflected in the operating curves if the optimal number of vessels per trip does not decrease substantially over the first few units (number of trips per vessel) on the x-axis. The Spanish pair trawl results are an example of the relative magnitude of the variance components increasing sharply from vessel to vessel × trip. It is visible in the shape of the operating curve (Fig. 7) as the optimal number of vessels, when considering the first few units (number of trips per vessel) on the x-axis, are decreasing rapidly.

The greatest variability can be seen at haul level with the exception of Northern Ireland twin-rig and English pair and Nephrops trawls. For Northern Ireland twin-rig and English pair trawls it was not possible to estimate the vessel × trip variance component due to the lack of trip replication per vessel but the variance components indicate that there is greater variability between vessels. For English unspecified otter trawls the vessel × trip variance component is small in relation to the other variance components indicating consistency in the discarding practices between trips within vessels where area is a significant fixed effect.

The standard errors of the between vessel variance components are relatively large in several cases, for example Spanish pair trawlers (Table 4), and are based on only a few degrees of freedom, indicating that caution should be exercised in interpreting them. This also occurs to a lesser extent with the between trip variance components, for example English Nephrops trawlers (Table 4) where information was available on only two vessels which were sampled more than once. This has implications for ensuring adequate vessel replication in future sampling schemes, and is discussed further in Section 4.5.

4.2. Significance of tested fixed effects and stratification

Tamsett and Janacek (1999) found that landing port, area fished and quarter were significant factors affecting discarding rates and included them as random effects. In this investigation area, quarter and their interaction were considered as fixed effects as it was required to examine the specific effects of area and quarter.

The lack of available data over time may be one reason why the fixed effect quarter did not significantly (\(P > 0.05\)) reduce the change in deviance from the basic models except in the case of English pair trawls. This also prohibited testing the random effect year in the analysis. When testing the significance of the fixed effects area or area × quarter, which was not possible in the case of English pair and Nephrops trawls as the sampled vessels fished in one area only, either or both were significant when tested using the Wald test. However, in the cases of Northern Ireland twin-rig and Spanish demersal heavy trawls this did not result in a significant decrease in the change in deviance from the basic models.

The significance of the fixed terms can indicate possible stratification schemes. In the discards context the danger is one of over stratification in that an on-board observer will waste time trying to find a suitable vessel subject to numerous constraints rather than boarding the first suitable vessel randomly chosen from the population regardless of any constraints. However, post-stratification can be applied to the data.

Throughout the discards project EC 95/094 Spain used a stratified random sampling scheme, stratifying by gear, quarter and port of departure. Bearing in mind that the Spanish pair fleet are day boats and port of departure will determine the area fished the highly significant (\(P ≈ 0.004\)) interaction of area and quarter confirmed the need for a stratified random sampling scheme. For the Spanish demersal fleet area was borderline significant (\(P ≈ 0.08\)). Further data collection is necessary before any definitive conclusions can be made about the demersal fleet. Spain has continued implementing this stratified approach in the current discards project EC 98/095.

Northern Ireland or England (Centre for Environment Fisheries and Aquaculture Science, CEFAS) did not use a stratified approach during the discards project EC 95/094 but used random sampling with equal probability and with pps sampling, respectively. CEFAS sampled English vessels landing into ports along the north east coast of England. The results from this study would indicate that post-stratification by area for Northern Ireland mid-water trawlers and English unspecified otter trawlers is appropriate.

Northern Ireland are continuing to sample the whole fleet using random sampling with equal probability under the current discards project EC 98/095 and will post-stratify by gear and, depending on data collection, quarter and area. The Sea Fish Industry Authority
(Seafish, based at Plymouth and Hull, England) have replaced CEFAS as members of discards project 98/095. Seafish are using a stratified approach in the main otter and beam trawl fleets operating from Newlyn, Looe and Plymouth by anticipated area to be fished (north or south of the Cornish Peninsula) and historical effort (mean number of hours fished taken over the previous 3 years) for each quarter. Using this stratification scheme the objective is to select trips for sampling with equal probability. CEFAS have now adopted a random sampling scheme with equal probability for the entire English and Welsh fleet. The selected vessels are approached in draw order for possible sampling. If the vessel is fishing in the North Sea or Skagerrak the next available trip is sampled otherwise the vessel is not sampled. Data are then raised using effort sampled and effort reported ratios to quarterly level. Effort is defined to be hours spent towing.

4.3. Optimal number of hauls, trips and vessels

It is preferable if on-board observers can gather data on all hauls while on a trip as there will always be a gain in precision by sampling all hauls. However, this may not be possible if the time required to process one haul exceeds the time interval between hauls, or if hauls are made within the observer’s rest period. The data collected from the fleets considered in the present paper indicate that in many cases, there is little to be gained in terms of precision by logging more than a certain number of hauls in a trip. This suggests that there would be some benefit in transferring observers between vessels at sea, although there are important safety issues in such operations and sampling may no longer be random.

Based on the available data the results indicate that one trip per vessel should be sampled except for Spanish pair trawls. For this fleet the results indicate that the same amount of resources, in terms of total observed trips, are required to sample an average of either one to two trips per vessel. At the time of the survey there were 35 Spanish vessels carrying pair trawls which suggests that in order to achieve a CV of 20% three to four trips per vessel are required (Fig. 7), assuming it is possible to achieve complete coverage of all vessels.

The results indicate that more resources, than were available in the discard surveys considered, are required to obtain a CV of 20% for Northern Ireland twin-rig and Spanish pair trawls. For England, a redirection of some sampling effort away from vessels using unspecified otter trawls towards pair and Nephrops trawls may have gone some way to achieving a target CV of 20% within available resources. It should be noted in the case of English pair and Nephrops trawlers the results are based on small samples.

4.4. The use of fisher self-sampling

One way to reduce the costs of sampling while increasing the sample size would be to introduce a fisher self-sampling scheme possibly using an on-board observer for the first trip sampled for a selected vessel in order to ensure consistency of measurement. Of the data sets analysed a fisher self-sampling scheme would be most suited to Spanish pair trawls as the trip usually lasts 1 day with the average number of hauls per trip tending towards one. Here the results show that considerably more effort, than was available during project EC 95/094, would be required to achieve a CV of 20%.

However, when fisher-self sampling was tested by Spain for their purse seine fleet it was rejected by the vessel owners to avoid problems whereby the crew carrying out the sampling were paid different amounts from the rest of the crew. Another potential problem was that the number of discarded species recorded by the skippers in the self-sampling scheme was found to be less than those recorded by the observers (Pérez et al., 1996).

In Northern Ireland fisher self-sampling has been used to estimate quantities of whiting and Nephrops discarded in the Northern Ireland Nephrops fishery since the early 1980s. These estimates are included in the annual ICES assessments of the stocks (e.g. Anon., 1997, 1999). Skippers of selected vessels (Briggs, 1985) supply the samples used to produce these estimates. Data taken from four to six trips per month are analysed. Fisher self-sampling has not been attempted in other Northern Ireland fisheries.

England (Seafish) investigated the use of fisher self-sampling for otter, beam offshore and beam inshore trawls in areas VII d–f and h during 1997 and 1998 (Seafish Consultancy Report No. 160). This was carried out as part of the Seafish Channel discard and
effort survey programme funded by the Ministry for Agriculture, Food and Fisheries (MAFF), which commenced in 1995. The data from six trips during which self-sampling was carried out were analysed and compared with data collected by the observers from the same or similar trips and were found to be consistent and in agreement. Of the six trips, four were carried out without observers on board. The observers found that the crews were enthusiastic about the self-sampling trials which undoubtedly contributed to the success of the sampling.

4.5. Sampling with partial replacement

If it can be established that the data collected using fisher self-sampling are reliable then the procedure and costs for providing crew with training in the use of this procedure will need to be considered. One possibility would be to select a core set of vessels and train their crew who, since they would be regularly engaged in self-sampling, would need only periodic checking for adherence to procedures. Since sampling could no longer be considered random, adjustments would be needed to the estimation procedures. In Allen et al. (2001) the use of sampling with partial replacement (SPR) is discussed as a possible sampling scheme for future discard surveys. SPR is an extension of the regression estimator and is based on the repeat sampling on either some or all occasions of the core set of vessels while supplementing the resulting data with samples from the remaining part of the fleet not part of the core set. SPR can be used to both estimate the current state and investigate trends (Cunia and Chevrou, 1969; Scott and Kohl, 1994). Omule (1984) illustrates how SPR can be used in conjunction with multistage sampling. However, for SPR to give gains in statistical precision, there needs to be correlation in the vessels’ discarding rate from one period to the next. At this stage the lack of available data consistently collected over time precludes the investigation of this option.

5. Conclusions

Rather than taking the figures quoted in the results as definitive answers to the question of optimal sample size, this study has attempted to suggest a methodology for analysing discards data in order to better inform fishery scientists for future improvements to discards surveys. The precision of the variance component (Table 4) indicates that caution should be exercised when interpreting the results. With further data collection the estimates could be improved and would permit further testing of the fixed effects as well as the random effect year.

While, for the majority of fleets analysed in this study an average of one trip per vessel would appear optimal, without some vessel replication it would be impossible to investigate this further. Some vessel replication would facilitate strong working relationships with the industry and establish which vessels could be used in a possible future SPR approach.

There will always be gains in precision, however slight, when collecting data beyond optimal levels. Based on current available data, the approach in this paper identifies the sample sizes required at each level to collect sufficient data to achieve a desired level of precision. It also illustrates what is practical and realistic to accomplish when planning future surveys taking account of monetary, time and workforce restrictions.

Acknowledgements

This study has been carried out with the financial support of the European Commission. The study neither reflects the views of the Commission nor its future policy in this area. The willing co-operation of the Northern Ireland, Spanish and English fishing industries in this programme is gratefully acknowledged. The authors would like to thank all scientific staff who carried out the field work in particular Tom Woods and Natasha Bunn. Contributions from John Cotter (CEFAS, Lowestoft), Spike Searle (Seafish, Plymouth) and Bill Lart (Seafish, Hull) were very much appreciated.

References


