Northwest Atlantic Fisheries Organization

Serial No. N2285

NAFO SCR Doc. 93/94

SCIENTIFIC COUNCIL MEETING - SEPTEMBER 1993

Combining Selectivities From Multiple Trouser Trawl Tows

by

D. L. Boulos[1], N. G. Cadigan[2], and W. M. Hickey[1]

[1] Fisheries Habitat and Management Branch, Dept. of Fisheries and Oceans
P. O. Box 5667, St. John's, Newfoundland, Canada A1C 5X1

and

[2] Science Branch, Department of Fisheries and Oceans
P. O. Box 5667, St. John's, Newfoundland, Canada A1C 5X1

ABSTRACT

An analysis of a trouser trawl experiment with multiple tows is presented. The standard method of analysis involves combining data from all tows and conditioning on the total catch. However, when the split of fish entering the experimental and control codends varies between tows, this method is inappropriate. A new model is proposed to analyze this type of data. Confidence bounds and likelihood ratio tests are developed as diagnostic tools for the analysis. The new model is also appropriate for the analysis of multiple subsampled tows without subsampling fractions. The new model provides a significantly better fit to the data.

INTRODUCTION

The trouser trawl (Cooper and Hickey, 1989; Walsh et al., 1992) has largely replaced the use of the covered codend as the preferred method for mesh selectivity studies. The trouser trawl consists of two codends, one made from small mesh and the other from a mesh for which selectivity is to be determined. Under the assumption that both codends encounter exactly the same distribution of fish lengths (i.e. an equal split), selectivity can be determined relative to the distribution of fish that are caught in the small mesh codend. However, observations indicate that the split of fish entering both codends is not always equal (Nicolajsen, 1988; Walsh *et al.*, 1992).

The SELECT method of Millar and Walsh (1992) corrects for the potentially unequal split. This method conditions on the total catch in both codends and estimates the selectivity for the experimental codend while estimating and controlling for an unknown proportion split.

A potential problem arises with the SELECT method when computing a common selectivity from multiple tows. It is possibile that the proportion split is not constant, but widely variable across sets. The parameter estimates for the common selectivity curve will be inaccurate if the data is combined without taking into account the unequal splits.

The SELECT method can be modified to take into account a different proportion split for each set while estimating a common selectivity for all sets. This paper presents the framework and methology for estimating this common selectivity curve.

MATERIALS AND METHODS

The standard SELECT method, which estimates the split of fish into both codends, is first presented (Millar and Walsh, 1992). The split (p), refers to the proportion of fish which enter the experimental codend from all the fish that enter the trouser trawl.

Let N_{l1} and N_{l2} be random variables representing the number of fish from length class l that are caught in the experimental and control codends, respectively. Let λ_l be the rate at which fish of length l enter the trawl. Let r(l) be the probability that a fish of length l is caught in the experimental codend, given that it enters this codend. Assuming that all the fish entering the control codend are caught,

$$E(N_{l1}) = pr(l)\lambda_l \text{ and}$$
$$E(N_{l2}) = (1-p)\lambda_l .$$

If N_{l1} and N_{l2} are Poisson random variables, then conditional on $N_{l1} + N_{l2}$, $N_{l1}/(N_{l1} + N_{l2})$, is a Binomial random variable with expectation:

$$E\left\{ \frac{N_{l1}}{N_{l1} + N_{l2}} \right\} = \frac{pr(l)}{1 - p + pr(l)} = \phi(l). \qquad (1)$$

Again, r(l) is the probability of a fish capture by the experimental codend, given that it enters this codend and $\phi(l)$ is the probability that a fish of length l is caught in the experimental codend, given that it is caught. r(l) is generally represented by the logistic function:

$$r(l) = \frac{e^{(a+bl)}}{1 + e^{(a+bl)}}.$$

ANALYSIS OF DATA FROM MULTIPLE SETS

If K tows are conducted in a selectivity experiment, then a common r(l) for all tows can be computed from all the data only if the splits are the same for each tow. Otherwise, the estimates for this common selectivity will be inaccurate.

Let k index tows (i.e.k=1,...,K). The model for the proportion retained ($\phi(l)$) in multiple tows with common selectivity but different splits is:

$$\phi_k(l) = \frac{p_k r(l)}{1 - p_k + p_k r(l)}. \qquad (2)$$

p_k in (2) allows the proportion split to vary with each tow; that is, (2) has a split for each tow. An alternative procedure, and one that is used in this paper, estimates a common split corresponding to the split for one tow and K-1 deviations from the common split for the remaining K-1 tows. In general,

$p_k = p + dp_i$, for k<K, and

$p_k = p$, if k=K,

where p is the common split and

dp_i is the deviation of tow i from p.

The split for the last tow is used arbitrarily as the common split. The varying split can be incorporated into the standard SELECT methodology by specifying K-1 categorical variables which index an increment to be added to the common split (i.e. for two tows, say $p_1 = 0.40$ = common split = p and $p_2 = 0.60$, $dp_1 = 0.20$, $p_k = p + dp_1 cat_1 = 0.4 + 0.2 cat_1$ (cat_1=1 if tow=2, cat_1=0 otherwise)).

When the selectivity data is based on subsampled catches , a modification of the standard method is required. Typically, subsampled catches are scaled by the sampling fractions and the standard method is applied or the methods developed in the previous paper by Cadigan and Hickey (MS-modified SELECT) can be used. However, sampling fractions are usually based on total weights caught

and Hickey show that the standard SELECT method can be applied without knowledge of the subsampling fractions. For multiple tows, the method developed here can also be applied to subsampled data without knowledge of subsampling fractions. This is advantageous if subsampling fractions are uncertain, unavailable or difficult to obtain.

The MS procedure involves a modification of (2):

$$\phi_k(l) = \frac{f_{k1}p_k r(l)}{f_{k2}(1-p_k)+f_{k1}p_k r(l)} \quad . \qquad (3)$$

Let $$p^*_k = \frac{f_{k1}p_k}{f_{k2}(1-p_k)+f_{k1}p_k} ,$$

then $$\phi_k(l) = \frac{p^*_k r(l)}{(1-p^*_k)+p^*_k r(l)} \quad . \qquad (4)$$

The f_{k1} and f_{k2} refer to the subsampling fractions by number or by weight for the experimental and control codends, respectively. Note from (4) that selectivity can be estimated even without the subsampling fractions, i.e. estimate (2).

Asymptotic standard errors for parameter estimates, where tow specific splits are estimated, are derived from the covariance matrix for parameter

Appendix of Millar and Walsh (1992). Recall that for K tows, there is a common split (p) for all tows and K-1 tow specific deviations ($dp_1,\ldots,dp_{k-1}$) which result in K+2 parameters and a (K+2) x (K+2) covariance matrix to be estimated.

The covariance matrix is first derived for the model with a split for each tow. Using basically the same notation of Millar and Walsh (1992), define I to be the information matrix and $\underline{\theta}'=(a, b, p_1,\ldots,p_k)$. The $COV(\hat{\underline{\theta}})=I^{-1}$, where $\hat{\underline{\theta}}$ is the maximum likelihood estimate of $\underline{\theta}$. Define $i_{j,k}$ to be the $(j,k)^{th}$ element of I, $i_{1,1}$, $i_{1,2}$ $i_{2,2}$ are computed exactly the same as in Millar and Walsh (1992). For each tow, the elements $i_{1,j+2}$, $i_{2,j+2}$ and $i_{j+2,j+2}$, j=1,...,K, are obtained by summing $i_{1,3}$, $i_{2,3}$ and $i_{3,3}$ in Miller and Walsh (1992) over the j^{th} tow. Note that $i_{j,k} = 0$ for $j > 3$, $k > 3$ and $j \neq k$. If fractions subsampled are used (i.e. the MS method), then $i_{1,j+2}$ and $i_{2,j+2}$ must be multiplied by $\partial p^*_j/\partial p_j$ and $i_{j+2,j+2}$ must be multiplied by $(\partial p^*_j/\partial p_j)^2$, $j=1,\ldots K$. In addition, p^*_k replaces p_k in all the formulas in the **Appendix** of Millar and Walsh (1992).

The parameter estimates for the model considered, where one common split and K-1 deviations (instead of K splits) are to be determined, can be obtained by a suitable linear transformation. Let $\hat{\underline{\theta}}^{*\prime}=(a, b, p, dp_1,\ldots, dp_{k-1})$ be the parameter estimates for the deviation model and $\hat{\underline{\theta}}'=(a, b, p_1,\ldots, p_k)$ be the parameter estimates for the K separate split model. The linear transformation $\hat{\underline{\theta}}^* = G\hat{\underline{\theta}}$ shows how $\hat{\underline{\theta}}^*$ can be formed from $\hat{\underline{\theta}}$. For the example considered in this paper (5 tows) G is a 7x7 matrix:

$$\underline{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} .$$

The covariance matrix for the estimates is then $\underline{GI}^{-1}\underline{G}'$.

Simultaneous confidence bounds for r(l) are generated to analyze the assumption that selectivity is constant between tows. These confidence bounds are interpreted such that the probability that a selectivity curve lies outside it's confidence bounds is at most α. Let $\underline{\theta}' = (a,b)$, where a and b are parameters from the logistic model for selectivity. The method involves constructing $(1-\alpha)100\%$ simultaneous confidence bounds for the Logit(r(l)) = a + bl using the asymptotic normality of the maximum likelihood estimates for a and b. The reverse Logit transformation is then applied to these bounds to yield simultaneous confidence bounds for the retention lengths or selectivities.

The procedure is highlighted as follows:

a) From asymptotic theory for maximum likelihood estimates,

$\hat{\underline{\theta}} \sim N(\underline{\theta},\Sigma)$, where Σ is the 2x2 covariance matrix for a and b.

Also, $$(\hat{\underline{\theta}} - \underline{\theta})' \, \Sigma^{-1} \, (\hat{\underline{\theta}} - \underline{\theta}) \sim \chi^2_{(2)} .$$

b) The following probability statements can be made:

$$1-\alpha = \text{Prob}\left[(\underline{\hat{\theta}} - \underline{\theta})' \, \Sigma^{-1} \, (\underline{\hat{\theta}} - \underline{\theta}) \le \chi^2_{\alpha(2)} \right],$$

$$= \text{Prob}\left[\max_{\underline{k}\neq 0} \frac{[\underline{k}'(\hat{\theta} - \theta)]^2}{\underline{k}'\Sigma\,\underline{k}} \le \chi^2_{\alpha(2)} \right], \quad \text{(Graybill, 1976)},$$

$$= \text{Prob}\left[-\sqrt{\underline{k}'\Sigma\,\underline{k}\,\chi^2_{\alpha(2)}} \le \underline{k}'(\underline{\hat{\theta}} - \underline{\theta}) \le \sqrt{\underline{k}'\Sigma\,\underline{k}\,\chi^2_{\alpha(2)}}, \text{ for all } \underline{k}\neq\underline{0} \right].$$

c) Let $d = \sqrt{\underline{k}'\Sigma\,\underline{k}\,\chi^2_{\alpha(2)}}$, then

$$\text{Prob}\left[\underline{k}'\underline{\hat{\theta}} - d \le \underline{k}'\underline{\theta} \le \underline{k}'\underline{\hat{\theta}} + d \right] \ge 1-\alpha.$$

This statement holds for any $\underline{k}\neq\underline{0}$. In particular, let $\underline{k}' = [1,1]$, then the simultaneous confidence bounds for r(l) are given by:

$$\text{Prob}\left[\frac{e^{(a+bl)-d}}{1+e^{(a+bl)-d}} \le r(l) \le \frac{e^{(a+bl)+d}}{1+e^{(a+bl)+d}} \right] \ge 1-\alpha$$

DATA

A selectivity study for Atlantic cod, Gadus morhua, using nominal 135 mm square mesh was conducted in NAFO Subdivisions 2J3KL. The study was accomplished in February, 1992 and was performed on the "ZANDVOORT", a 52 meter commercial stern trawler. This vessel carried a 47 m Hampijhan trawl which had been converted to a trouser trawl with a vertical divider panel and twin codends. The footrope length was 61 m and the headline length was 47 m. The mesh size was 160 mm (k.c.) in the wings, square and first belly and 160 mm in the twin extensions.

Five sets or tows were performed with this trawl and sampling was carried out by two fisheries representatives. Samples were obtained on deck when possible; however, when temperatures were below freezing, sampling was carried out on the ramp. The lengths were measured to the nearest cm and length frequencies were used to create three cm groupings on which the analysis was performed.

RESULTS

Figure 1 is used in analyzing the assumption of common selectivity among tows. The estimated selectivities for each tow (parameter estimates are in Table 1) and 95% confidence bounds are plotted as points and dotted lines, respectively. The solid line is

the estimated selectivity from the tow whose number is indicated in the upper left hand corner of each plot. If the selectivity from one tow exceeds the confidence bounds from another tow then that is evidence that the selectivities are different. It appears that selectivity is different between tows 9 and 13, 9 and 14, as well as 14 and 15. These are not statistical tests however, because the variability in the estimated selectivity indicated by the solid line is ignored. A better procedure will follow .

Four models are estimated, they are:

(i) Tow specific splits and selectivities using MS.

(ii) Tow specific splits and a common selectivity using MS, (3).

(iii) Tow specific splits and a common selectivity using the raw data, i.e. no sampling fractions, (2).

(iv) Common split and selectivity using MS.

The estimates from analysis (i) are presented in Table 1. The splits are quite variable, especially for tows 9 and 13. The estimates from analysis (ii) are presented in Table 2. Table 3 contains the estimates from analysis (iii) and is presented to illustrate that subsampling fraction are not required as (iii) produces the same common selectivity estimate as (ii). The only difference in Table's 2 and 3 is the interpretation of the p estimates. The parameter estimates for analysis (iv) are presented in Table 4.

The three estimates of selectivity for each tow are presented in Figure 2 (recall that the selectivity in (ii) and (iii) are the same). The common selectivity for tows 13 and 14 computed using (iv) appear to be inaccurate, while the selectivity estimated from (ii) is more similar to the tow specific fits from (i).

Statistical comparisons between fits are made with likelihood ratio tests or, equivalently, deviance tests (see Millar and Walsh, 1992). A schematic diagram of the hypotheses tested is presented in Figure 3. The deviance for (i) can be used as a goodness of fit test; it indicates whether the logistic model is appropriate. The p-value is large which suggests that the logistic model provides an adequate fit. The difference in deviances (χ^2) between (i) and (ii) is large and statistically significant, i.e. p-value < 0.05. This suggests that selectivity is statistically different between tows. Similarly, the p-value < 0.05 for the comparison of (ii) and (iv), indicating that the splits are not equal among tows.

DISCUSSION

A procedure has been presented to estimate a common selectivity curve from a trouser trawl selectivity study with multiple tows, and where the split varies among tows. The procedure may also be used to estimate a common selectivity from multiple subsampled tows even without the subsampling fractions. This could be advantageous if subsampling fractions do not exist, are difficult to obtain, or are measured poorly. The procedure should not be applied routinely if splits are not different between tows because unnecessarily estimating tow specific splits reduces the overall precision of parameter estimates. It is not recommended to stop obtaining subsampling fractions for this reason. Further research about subsampled experiments is required.

The confidence bounds presented are useful tools to analyze between tow variability in selectivity but they do not provide statistical tests. The likelihood ratio statistical tests give overall measures of between tow variability in selectivity; however, when the number of fish in a tow is large, even small differences in selectivities and splits will be significant. In this case, determining if selectivities are practically common among tows is subjective.

The data analysis is unsatisfactory in the sense that there appears to be between tow variability in selectivity and the methods developed here assume that selectivity is constant. Further work in the direction of Fryer (1992) is required. For this reason the estimates presented are recommended only for comparative purposes, not for general use.

REFERENCES

Cooper, C.G. and Hickey, W.M. 1989. Selectivity experiments with square mesh codends of 135, 140 and 155 mm. Fisheries Development and Fisherman's Services Division Project Report 154, 29pp.

Fryer, R.J. 1991. A model of between-haul variation in selectivity. ICES, Journal of Marine Science, 48: 281-290.

Graybill, F.A. 1976. Theory and application of the linear model. Wadsworth Publishing Company, Inc. Belmont, California, 704pp.

Millar, R.B. and S.J. Walsh. 1992. Analysis of trawl selectivity studies with application to trouser trawls. Fisheries Research, 13: 205-220.

Nicolajsen, A. 1988. Estimation of selectivity by means of a vertically split Nephrops trawl. ICES C.M., /B:9, 12pp.

Walsh, S.J., Millar, R.B., Cooper, C.G. and Hickey, W.M. 1992. Codend selection in American plaice: diamond versus square mesh. Fisheries Research, 13: 235-254.

Table 1. Tow specific split and selectivity parameter estimates and standard errors (in parentheses) for proportion retained. S.R. is the selection range, the range in fish lengths over which selection occurs. L50 is the 50% retention length, and the length of fish at which 50% are retained by a fishing gear and 50% escape.

Tow	a	b	L50	S.R.	p
7	-17.9 (2.00)	0.34 (0.05)	52.44 (1.52)	6.43 (0.86)	0.51 (0.11)
9	-15.6 (1.40)	0.31 (0.04)	50.82 (1.78)	7.14 (0.84)	0.34 (0.31)
13	-16.6 (1.25)	0.30 (0.03)	55.08 (2.35)	7.30 (0.78)	0.75 (0.10)
14	-21.4 (1.68)	0.40 (0.04)	54.17 (1.41)	5.55 (0.54)	0.49 (0.20)
15	-12.3 (1.05)	0.22 (0.03)	54.59 (3.62)	9.79 (1.34)	0.43 (0.37)

Table 2. Tow specific splits and common selectivity parameter estimates and standard errors (in parentheses), using fractions subsampled, for the proportion retained. S.R. is the selection range and L50 is the 50% retention length.

a	b	L50	S.R.	p	dp1	dp2	dp3	dp4
-15.7 (0.56)	0.29 (0.01)	54.58 (1.00)	7.63 (0.37)	0.55 (0.11)	0.02 (0.08)	-0.03 (0.08)	0.16 (0.08)	-0.16 (0.10)

Table 3. Tow specific splits and common selectivity parameter estimates and standard errors (in parentheses), without using fractions subsampled, for the proportion retained. S.R. is the selection range and L50 is the 50% retention length.

a	b	L50	S.R.	p	dp1	dp2	dp3	dp4
-15.7 (0.56)	0.29 (0.01)	54.58 (1.00)	7.63 (0.37)	0.92 (0.01)	-0.07 (0.02)	0.01 (0.01)	-0.03 (0.01)	-0.06 (0.01)

Table 4. Tow common split and common selectivity parameter estimates and standard errors (in parentheses), using fractions subsampled, for the proportion retained. S.R. is the selection range and L50 is the 50% retention length.

a	b	L50	S.R.	p
-16.0 (0.57)	0.29 (0.01)	54.38 (0.81)	7.45 (0.35)	0.55 (0.06)

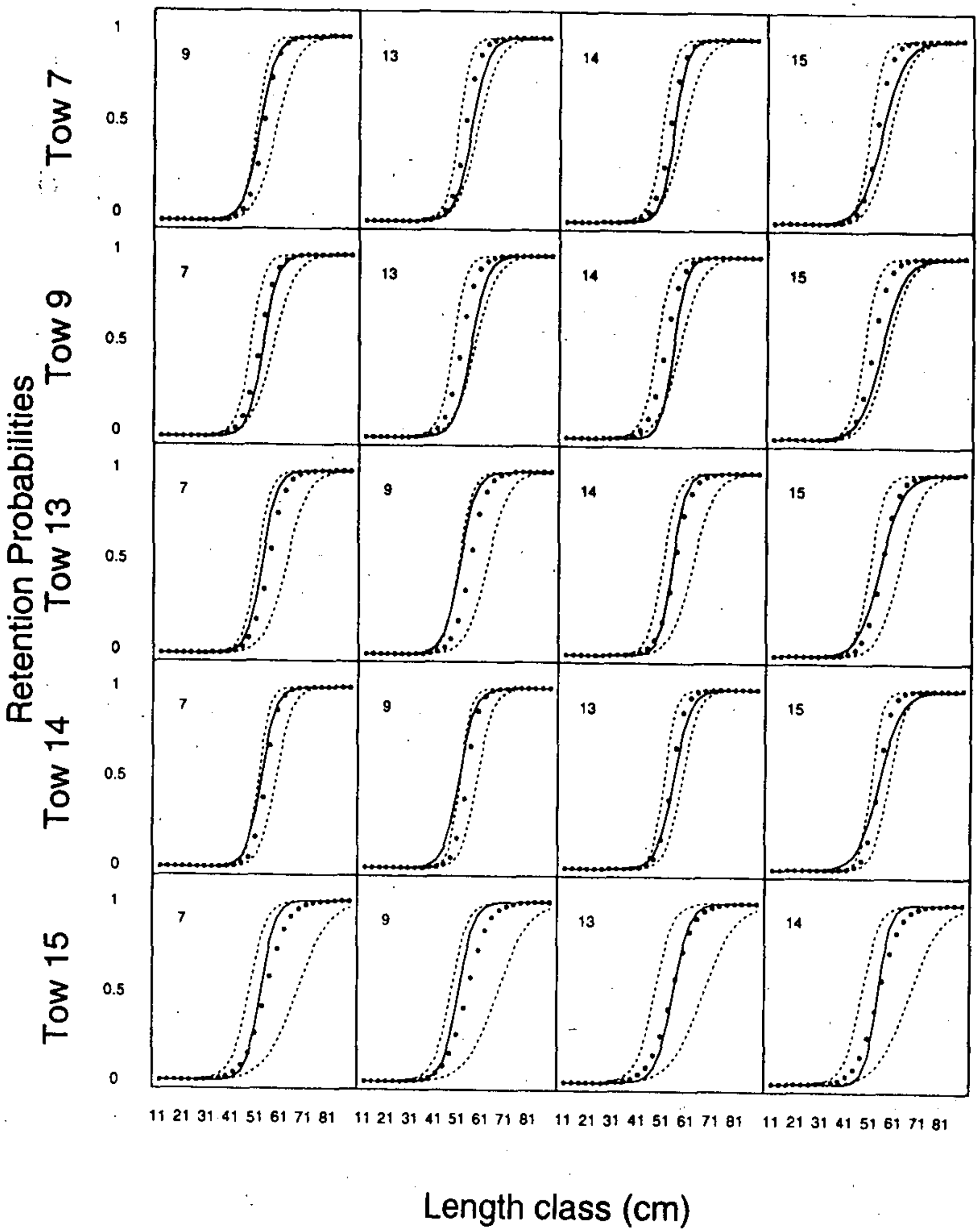

Figure 1. A tow by tow comparison of selectivity. The estimated selectivity (points) and 95% confidence bounds (dotted lines) for each tow are plotted in different rows and every column. The estimated selectivity (solid line) is plotted for each other tow, whose number is represented in the upper left hand corner.

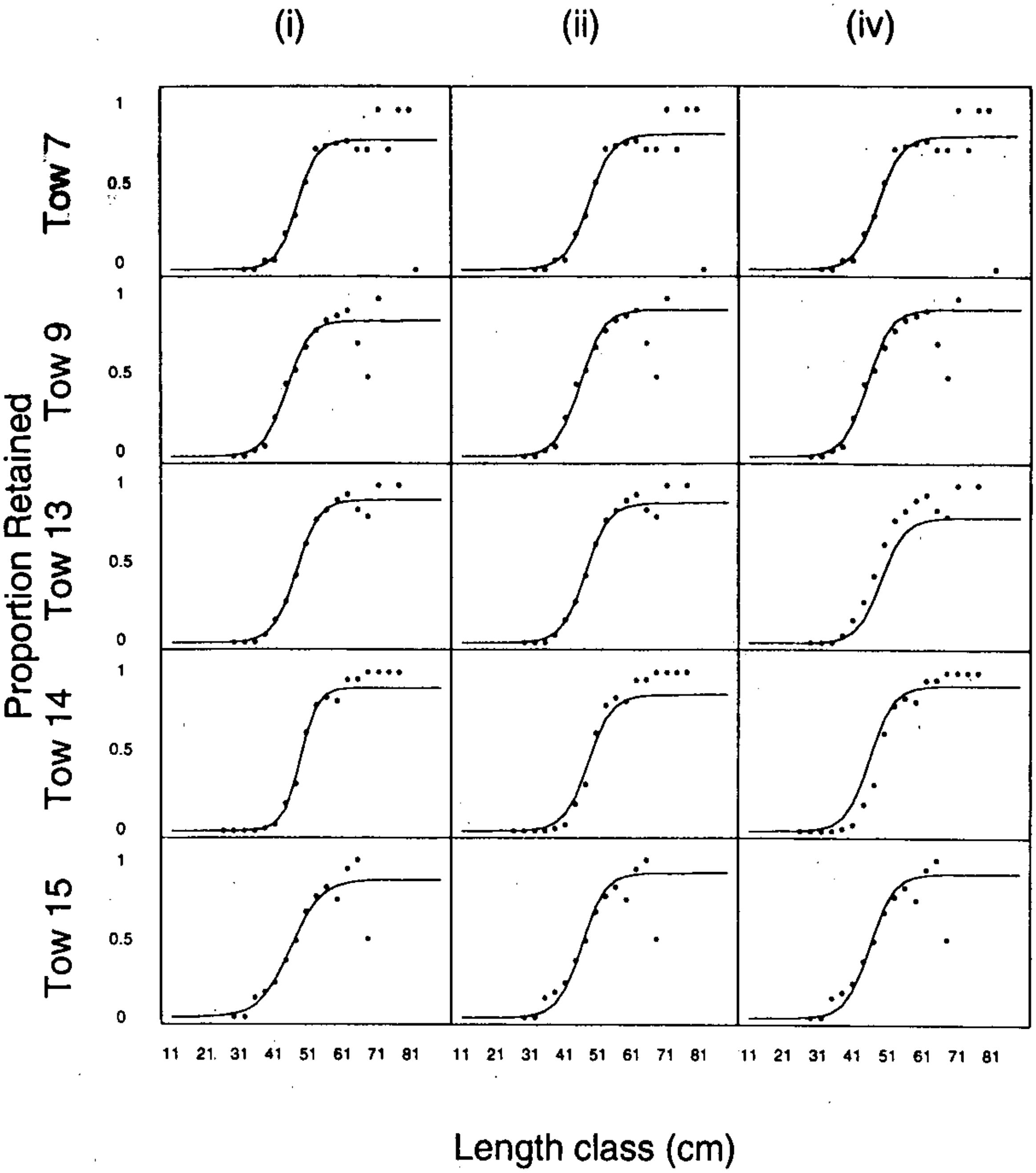

Figure 2. Estimated proportion retained (solid line) and observed proportion retained (points) for each tow and different models: (i) Tow specific splits and selectivities using MS. (ii) Tow specific splits and a common selectivity using MS, (3). (iv) Common split and selectivity using MS.

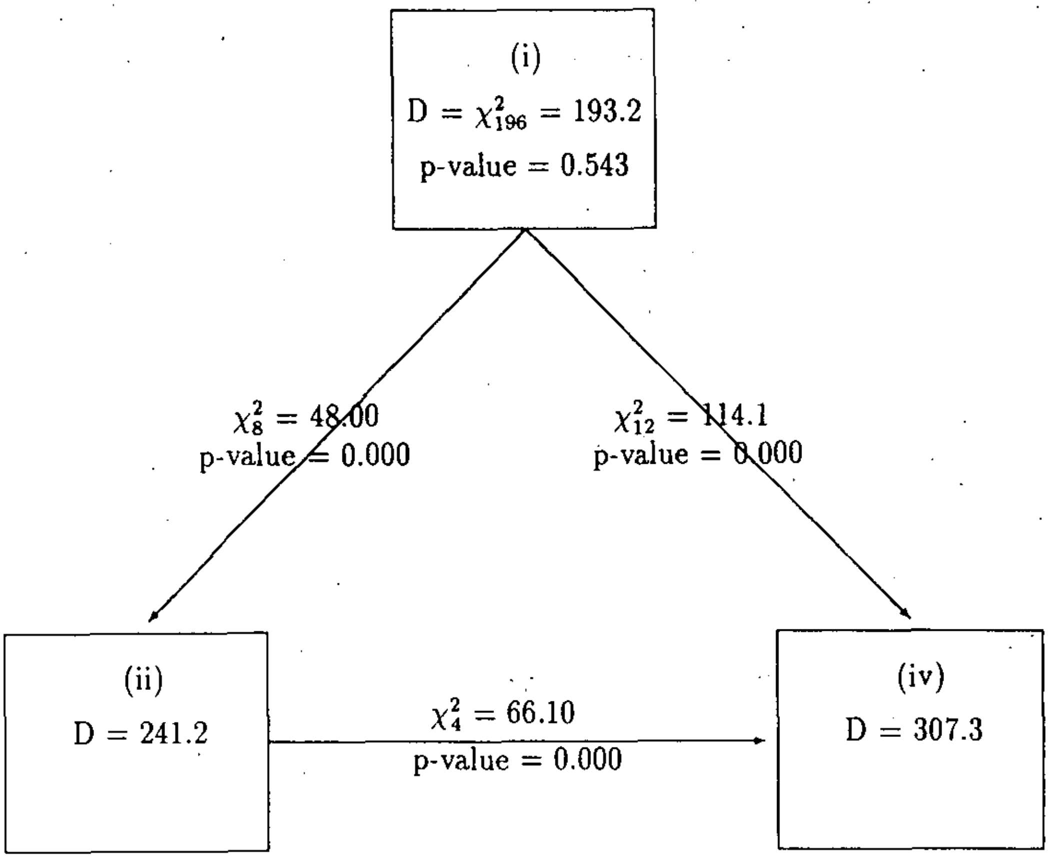

Figure 3. Likelihood ratio tests of model fits. The p-value comes from a Chi-square distribution with degrees of freedom as indicated. The models fit are: (i) Tow specific splits and selectivities using MS. (ii) Tow specific splits and a common selectivity using MS, (3). (iv) Common split and selectivity using MS.