CCAMLR Science, Vol. 15 (2008): 1–34

AN INTEGRATED STOCK ASSESSMENT FOR THE PATAGONIAN TOOTHFISH (*DISSOSTICHUS ELEGINOIDES*) FOR THE HEARD AND MCDONALD ISLANDS USING CASAL

S.G. Candy✉ and A.J. Constable
Department of the Environment,
Water, Heritage and the Arts
Australian Antarctic Division
203 Channel Highway, Kingston 7050
Tasmania, Australia
Email – steve.candy@aad.gov.au

Abstract

An integrated stock assessment for the Patagonian toothfish (*Dissostichus eleginoides*) for the Heard and McDonald Islands (CCAMLR Division 58.5.2), using CASAL and data consisting of multi-year random stratified trawl survey (RSTS) abundance estimates by length bin, commercial catch-at-length data, standardised CPUE series for the trawl grounds, and tag releases and recoveries by length bin, is described. The annual surveys are spatially representative of the main plateau, where juvenile fish are found, but are of relatively low intensity in effort compared to the commercial shots. In contrast, the commercial shots are very restricted in space, consisting of three main grounds. The model implemented in CASAL is a simple, single-sex, single-area population model, but spatial complexity in the fishery was modelled using separate fishing selectivity functions for each ground by gear (trawl and longline) combination. Various combinations of dataset weighting were investigated using haul-level estimated effective sample sizes with, additionally, iteratively estimated process error for the catch-at-length data and the inclusion versus exclusion of the tag data. A key uncertainty is the number of ages fully selected by the main survey series. With all the data included in the model, age-4 and 5 fish are fully selected, whereas when the survey data have greater influence and without the tag data, the selectivity of these ages was reduced. The method of quantifying process error used in this assessment removes 'systematic lack-of-fit' (SLOF) from population/fishery model predictions. Extension of this method of estimating process error to RSTS abundance data and commercial catch CPUE data is given, but incorporation of process error for the RSTS data was not considered appropriate, since SLOF could not be removed to an acceptable degree. The issues concerning the effect of the tension between survey data and mark–recapture data on parameter estimation are discussed.

Résumé

Le présent article est une description de l'évaluation intégrée des stocks de légine australe (*Dissostichus eleginoides*) aux îles Heard et McDonald (Division 58.5.2 de la CCAMLR) réalisée à partir de CASAL et de données consistant en des estimations d'abondance par lots de longueurs, provenant d'une campagne d'évaluation stratifiée aléatoire au chalut (RSTS) pluriannuelle, en données commerciales de capture selon la longueur, en séries de CPUE normalisée pour les lieux de pêche au chalut et en données de pose et de récupération de marques par lots de longueurs. Les campagnes d'évaluation annuelles offrent une représentation spatiale du plateau principal, sur lequel sont observés les juvéniles de poissons, mais elles ne déploient qu'un effort de pêche relativement peu intense, par comparaison avec les poses commerciales. Par contre, les poses commerciales sont très limitées dans l'espace, ne concernant que les trois principaux lieux de pêche. Le modèle appliqué dans CASAL est un modèle de population simple, fondé sur un seul sexe et une seule zone, mais la complexité spatiale de la pêcherie est modélisée au moyen de fonctions de sélectivité de la pêche, différentes pour chacun des lieux de pêche par combinaison d'engins (chalut et palangre). Diverses combinaisons de pondérations de jeux de données sont étudiés au moyen de la taille effective des échantillons estimée au niveau du trait avec, de plus, l'erreur de processus estimée de manière itérative pour les données de capture selon la longueur et l'inclusion, par rapport à l'exclusion, des données de marquage. L'une des incertitudes clés réside dans le nombre d'âges pleinement sélectionnés par la principale série de campagnes d'évaluation. Lorsque toutes les données sont entrées dans le modèle, les poissons des âges 4 et 5 sont pleinement sélectionnés, alors que lorsque les données des campagnes d'évaluation ont une plus grande influence et que l'on omet les données de marquage, la sélectivité de ces âges est réduite. La méthode de la quantification de l'erreur du processus utilisée dans cette évaluation supprime "le défaut d'ajustement systématique"

(SLOF) des prévisions des modèles de population/pêcherie. L'extension de cette méthode d'estimation de l'erreur de processus aux données d'abondance de RSTS et aux données de CPUE de la capture commerciale est donnée, mais l'insertion de l'erreur de processus pour les données de RSTS n'a pas été considérée comme appropriée, du fait que le SLOF n'a pas pu être supprimé jusqu'à un niveau acceptable. Les questions concernant l'effet de la tension entre les données des campagnes d'évaluation et les données de marquage–recapture sur l'estimation des paramètres font l'objet d'une discussion.

Резюме

Описывается комплексная оценка запаса патагонского клыкача (*Dissostichus eleginoides*) в районе о-вов Херд и Макдональд (Участок АНТКОМа 58.5.2), проводившаяся на основе CASAL и данных, включающих оценки численности по интервалам длины, полученные по многолетним случайным стратифицированным траловым съемкам (RSTS), коммерческие данные о длинах в улове, стандартизованные ряды CPUE для траловых участков, а также данные о выпусках и повторных поимках по интервалам длины. Ежегодные съемки в пространственном отношении дают представление об основном плато, где находится молодь, но интенсивность усилия при них ниже, чем при коммерческих постановках. Однако коммерческие постановки очень ограничены в пространственном отношении, т.к. проводятся только на трех основных участках. Модель, выполненная в CASAL, представляет собой простую модель популяции для одного пола и одного района, а пространственная сложность промысла моделировалась с помощью отдельных функций промысловой селективности для каждого участка в соответствии с сочетанием промысловых снастей (трал и ярус). Различные варианты взвешивания наборов данных анализировались с использованием рассчитанных эффективных размеров выборки на уровне улова с добавлением итерационно рассчитанной ошибки обработки данных о длинах в улове и включением/исключением данных мечения. Основную неопределенность представляет количество возрастов, полностью отобранных основным съемочным рядом. При включении в модель всех данных рыба в возрасте 4 и 5 отбирается полностью, тогда как при большем влиянии съемочных данных и без данных мечения селективность этих возрастов меньше. Используемый в этой оценке метод количественного определения ошибки обработки устраняет «систематическое несоответствие» (SLOF) из прогнозов модели популяции/промысла. Описывается использование этого метода определения ошибки обработки для данных о численности данных RSTS и данных CPUE по коммерческим уловам, однако было решено, что включать ошибку обработки RSTS не следует, т.к. невозможно в приемлемой степени избавиться от SLOF. Рассматриваются вопросы, связанные с воздействием несоответствий между съемочными данными и данными мечения–повторной поимки на оценку параметров.

Resumen

Se describe una evaluación integrada del stock de austromerluza negra (*Dissostichus eleginoides*) de las Islas Heard y McDonald (División 58.5.2 de la CCRVMA), utilizando el modelo CASAL y datos que comprenden las estimaciones multianuales de la abundancia por intervalo de tallas de prospecciones de arrastre estratificadas aleatoriamente (RSTS), los datos de talla de las capturas comerciales, las series cronológicas de la CPUE normalizada para los caladeros de arrastre, y los datos de liberación y recuperación de marcas por intervalo de tallas. Las prospecciones anuales son representativas del área de la plataforma principal, donde se encuentran peces juveniles, pero el esfuerzo es relativamente menor comparado con la intensidad del esfuerzo de los lances comerciales. Por el contrario, el área donde se efectúan los lances comerciales es muy limitada, sólo tres caladeros principales. El modelo ejecutado con CASAL es un modelo demográfico simple, para un solo sexo y una sola área, aunque se utilizaron funciones de selectividad para una combinación de artes (arrastre y palangre) por separado para cada caladero para simular la complejidad espacial en la pesquería. Se estudiaron diversas combinaciones de ponderación de los conjuntos de datos utilizando tamaños efectivos de muestra estimados de los lances individuales, agregando el error de tratamiento estimado repetidamente para los datos de frecuencia de tallas de la captura, e incorporando o excluyendo los datos de marcado. Una duda importante es el número de edades seleccionadas totalmente por la serie de la prospección principal. Cuando todos los datos han sido incluidos en el modelo, los peces de 4 y 5 años de edad están totalmente seleccionados, mientras que cuando los datos de la prospección tienen más peso y no se incluye los datos de marcado, la selectividad de estas edades disminuye. El método utilizado en esta evaluación para calcular el error de tratamiento elimina la 'falla sistemática del ajuste' (SLOF) de las predicciones del modelo de población/pesquería. Se presenta una ampliación de este método de estimación del error de tratamiento para los datos de abundancia de RSTS y los datos de la CPUE de la captura comercial, pero la incorporación del error de tratamiento para los datos RSTS no se consideró apropiada, pues no se pudo reducir SLOF a un nivel aceptable. Se consideran los problemas relacionados con el efecto de la tensión entre los datos de la prospección y los datos de marcado y recaptura en la estimación de parámetros.

Keywords: random stratified trawl surveys, catch-at-length frequencies, catch and effort indices, tag data, process error, lack of fit, CCAMLR

Introduction

Patagonian toothfish (*Dissostichus eleginoides*) is a long-lived, slow-growing species (Gon and Heemstra, 1990) which is harvested by longline and trawl fisheries in the vicinity of Heard and McDonald Islands (CCAMLR Division 58.5.2). Assessments of long-term annual yield for *D. eleginoides* in Division 58.5.2 were first undertaken in 1996 (SC-CAMLR, 1996) using the approach adopted for this species by the CCAMLR Working Group on Fish Stock Assessment (WG-FSA) in 1995. This approach used random stratified trawl surveys (RSTS) of young fish to estimate abundance of juvenile fish using the CMIX method and software (de la Mare, 1994; de la Mare et al., 2002). The estimates of cohort strength were then projected for 35 years using a population model, incorporating growth, mortality, maturity and fishing removals and associated selectivity functions using the generalised yield model (GYM) (Constable and de la Mare, 1996). The results using the GYM are presented by Welsford et al. (2006a).

Assessments of fish stocks using all available data are becoming common-place with the advent of maximum likelihood and Bayesian statistical tools that can incorporate diverse datasets (Maunder, 2003; Butterworth et al., 2003; Hillary et al., 2006). CASAL (Bull et al., 2005) is one such tool that makes the integration of these datasets relatively straight forward within a generalised stock assessment package. CASAL is a software package for carrying out 'integrated' stock assessments (i.e. integration of all relevant datasets in parameter estimation), that has been used to assess long-term annual yield according to the precautionary approach of CCAMLR, described above, in assessments of Antarctic toothfish (*D. mawsoni*) (Dunn et al., 2004, 2005; Dunn and Hanchet, 2006, 2007) and *D. eleginoides* (Hillary et al., 2005, 2006; Agnew et al., 2007). The primary difference in this approach to that used in the GYM is that many parameters can be estimated simultaneously within the model, including the pre-exploitation median spawning biomass, rather than having to estimate the parameters individually, which does not account properly for the correlation between parameters. In this way, the integration of projections across uncertainty in parameter values uses sample sets of parameters that are consistent with the data and are appropriately correlated. This approach is advantageous when parameters are difficult to estimate in isolation, such as those in fishing selectivity functions.

This paper provides an integrated assessment of stock status and recruitment variability for *D. eleginoides* in Division 58.5.2. A feature of this assessment is that it used a number of contrasting datasets, including multiple fisheries catch-at-length proportions, fisheries-independent research survey data (RSTS described above) and mark–recapture data from the different fisheries. Although variable in their respective lengths of time series, these different 'views' of the stock provide insights into the variety of issues that need to be addressed in toothfish assessments, notably the age-specific spatial structure in the stock, the potential biases that may arise from the individual datasets and the processes that might be used to effectively weight the respective contributions of these datasets for parameter estimation.

The assessments presented here have evolved over recent years. They update the model used at WG-FSA in 2006 using data from the 2007 season as well as 2006 data not available for WG-FSA in 2006. In particular, this paper incorporates refinements to previous assessments of this stock to resolve some of the earlier difficulties in this assessment including: (i) estimation of the coefficient of variation (CV) for length given age, (ii) use of non-informative priors for year-class strength parameters, (iii) separate selectivity parameters used for the pre-2006 compared to the 2006–2007 fishing seasons for the main trawl ground, (iv) separate selectivity parameters for the late (within-year) season compared to the combined early (within-year) seasons (see below for a description of the within-year seasons) for the main trawl ground, and (v) the use of an improved method of determining effective sample size (ESS) for commercial catch-at-length data as described in Candy (2008).

The CV for the normal distribution for length-at-age, CV_{VB}, is required to convert length frequencies to age frequencies in CASAL. Previously (Constable et al., 2006a, 2006b), this was obtained independently of CASAL from the fit of the von Bertalanffy growth model to length-at-age data (i.e. estimated parameter σ in Candy et al., 2007, Table 1). In order to investigate the sensitivity of predictions of age structure to CV_{VB} this parameter was estimated using CASAL.

In 2007, the CCAMLR Working Group on Statistics, Assessments and Modelling (WG-SAM) (SC-CAMLR, 2007b, paragraph 4.6(ii)) noted that the CV in year-class strength (YCS – number of annual age-1 recruits relative to the median pre-fishery age-1 annual recruitment, R_0), CV_R, was found in sensitivity trials to be largely determined by the CV provided as a prior in previous integrated models (Constable et al., 2006a, 2006b). The calculation of long-term yield can be very sensitive to CV_R in that a high variability in recruitment can result in the depletion rule being the binding part of the decision rule, rather than the escapement rule. For this reason a non-informative prior distribution was used for YCS parameters.

A stock-recruitment relationship was not employed; instead, historical annual recruitments varied about R_0 as described above. Recent assessment work by Hillary et al. (2006) for the South Georgia toothfish stock that employed a Beverton-Holt stock-recruitment relationship with fixed (i.e. not estimated) steepness parameter of 0.8 (sensitivity analyses varied this parameter using values of 0.7 and 0.9) indicated that the range of estimated spawning stock biomass (SSB) that influence recruitment (i.e. SSBs corresponding to the steeply inclining phase of the relationship) are well below any current annual estimates of SSB. Therefore, the stock recruitment relationship was likely to have had only a small effect on average recruitment for SSBs corresponding to the plateau section of the relationship. For this reason incorporation of stock-recruitment relationship was not considered necessary in the current estimation of YCS parameters and R_0.

An important consideration in integrated models is to ensure that data are given appropriate weights in the objective function. The length composition data (proportions-at-length) from commercial catches, with sample sizes determined as the number of fish measured, can potentially swamp other datasets in the analysis. The method of calculating ESS (i.e. appropriately reducing the above sample sizes) assuming the length composition data follow the multinomial distribution is described in Candy (2008). Candy (2008) reported that this method, when compared to simulations of between-haul heterogeneity in proportions-at-length using the Dirichlet-multinomial distribution, gave the best estimates of ESS over the complete range of simulated between-haul heterogeneity. In addition, Candy (2008) gave a method of accounting for process error in catch-at-length data taking into account systematic lack-of-fit (SLOF) using a generic model for SLOF. This method was used for catch-at-length data and the method of estimating process error variance as part of the process of fitting the generic model for SLOF was extended to all datasets.

While a key sensitivity of the model was uncertainty in the natural mortality rate, specified by the parameter M (see Constable et al., 2006a, 2006b) the effect of changing M was not considered with M kept fixed at 0.13 year^{-1} as recommended in SC-CAMLR (2006). The maximum age was set to 35 years since the model of Candy et al. (2007) gives unrealistically high average length-at-age as age is increased much past 35 years because of the lack of available data for ages above 35 years that define the asymptotic phase of the growth.

Methods

Population model

The population of *D. eleginoides* was modelled as an age-structured population with 35 separate age classes, from age 1 to 35 years with the last age class being a plus class. It is a single-area (i.e. single closed population), multi-fishery model that does not discriminate between sexes. The growth parameters were based on a von Bertalanffy growth model that incorporates an early age (<5 years old) adjustment as described in Candy et al. (2007). Because of the young-age adjustment, growth was modelled using a vector of predicted mean length-at-age as input to CASAL. The estimate of CV of length given age, CV_{VB}, was obtained by Candy et al. (2007) as 0.1 and this value was used as the initial value in the joint estimation of this parameter with the other model parameters.

Two fishing methods are used in Division 58.5.2 – trawling and longlining. The annual cycle was split into three seasons: (s1) 1 December–30 April, (s2) 1 May–30 September, and (s3) 1 October–30 November. These seasons were structured to accommodate the main longline fishery occurring during the period from May to September.

There are now three main fishing grounds (Figure 1), Grounds B, C and D, with trawling restricted to Grounds B and C and the majority of longline fishing carried out in Grounds C and D. Each combination of gear by ground was considered a separate fishery in the model in order to include some of the spatial complexity of the region within the single-area model in CASAL. Thus, fishing selectivities in the model were really vulnerabilities, which are a combination of availability and gear selectivity; differences in age composition in the different fishing grounds were accommodated in this approach. Hereafter, the term 'selectivity' is used to be consistent with the terminology of the functions in CASAL but should be interpreted as vulnerability. Here, as in Constable et al. (2006b), selectivities were estimated only for each ground by gear combination since previous analyses indicated that selectivity function parameters could be pooled across seasons while maintaining very similar selectivity curves to those obtained without pooling. The two exceptions to this were for trawl in Ground B where a separate selectivity function was fitted in season 3 versus combined season 1 and 2 for fishing years prior to 2006, and the most recent fishing years of 2006 and 2007 were given a single selectivity function that was distinguished from the pre-2006 selectivities. The age-structured population model combined with the fishery removals and selectivities is denoted the 'population/fishery' model for the remainder.

Data

A number of datasets were used to estimate model parameters. Some data were used as direct inputs to the model and other data were included as observations to fit model parameters. These datasets (updated as far as possible for the 2007 fishing season and including 2006 data) included:

- abundance by length bins for groupings of annual surveys (observations);
- series of legal removals (catch) partitioned by fishery and season (input data);
- a series of estimated illegal, unreported and unregulated (IUU) removals (input data);
- a standardised catch-per-unit-effort (CPUE) series for the two commercial trawl fishing grounds (observations);
- length-frequency composition for the commercial fisheries in Grounds B, C and D (observations);
- numbers of tag–releases (input data) and tag–recoveries for each fishery and fishing year, as well as estimates of the number of fish 'scanned' for tags for each fishery in which the recaptures occurred (observations).

Surveys

Following the review of Welsford et al. (2006b), the surveys were grouped into five groups with the main group, Survey Group 1, comprising surveys from 2001, 2002 and 2004 to 2007 (Welsford et al., 2006b). The other surveys were treated individually because of their differences in design, timing and type of vessel being used – 1990, 1993, 1999 and 2003 (Survey Groups 3, 4, 2 and 5 respectively). The surveys carried out in 1992 and 2000 are excluded because of their poor sampling design in terms of toothfish assessment (see review by Welsford et al., 2006b). Annual surveys comprised between 120 and 160 hauls, each of which were of approximately 30 min duration, that were randomly located within strata that cover the main plateau where juvenile fish are found (Nowara and Lamb, 2007; SC-CAMLR, 2007a, Appendix L). Compared to the effort in the commercial trawl and longline fisheries, the survey represents only a small fraction of fishing effort in any year.

A double-normal plateau (DNP) selectivity function (using four parameters – Bull et al., 2005) was estimated for each survey group and was calculated as $f(x)$, for age x as:

$$\begin{aligned} f(x) &= a_{\max} 2^{-\left[(x-a_1)/s_L\right]^2} && ; x \le a_1 \\ &= a_{\max} && ; a_1 < x \le a_1 + a_2 \\ &= a_{\max} 2^{-\left[\{x-(a_1+a_2)\}/s_U\right]^2} && ; x > a_1 + a_2 \end{aligned}$$

where parameters to be estimated are (s_L, s_U, a_1, a_2, $a_{\max}$). In all cases $a_{\max}$ was not estimated but set to 1. The lower bound on a_2 was varied by either: (i) setting a lower bound of 0.1 (i.e. when setting the bound close to zero numerical problems for estimation occurred); or (ii) by fitting a double-normal (DN) selectivity function if this parameter hit the lower bound of 0.1, which is equivalent to the constraint $a_2 \equiv 0$.

Survey Group 1 was assumed to fully observe the stock that was vulnerable to the survey fishing gear, as quantified by the fitted selectivity function for the survey, (Welsford et al., 2006b) and was therefore, for all but one of the models described, given a catchability of $q = 1$. The population/fishery model that estimated q included the tag data to allow this estimation to be successful. The other surveys differ from Survey Group 1 in design and survey vessel. For these surveys, q was estimated with the exception that a common q was estimated for the early surveys (i.e. years 1990 and 1993).

The estimate of abundance of fish by length bin and the corresponding CV were obtained using a bootstrap procedure (Constable et al., 2006b), retaining the stratification and length composition in a haul during the bootstrap. Note that CVs that were greater than 5 were reduced to 5 and abundances of zero were given a nominal value of 1 and CV of 1 so that the likelihood could be calculated. There was only a single zero value and no CVs greater than 5 for Survey Group 1 so that these nominal values had negligible influence on model predictions.

Length bins of 50 mm were used with bins outside the 300 to 1 100 mm range removed to reduce the otherwise very large number of zero abundances as detailed by Constable et al. (2006b).

Catches

Historical legal catches for each fishery and estimated IUU removals were used as known catch (Table 1) (SC-CAMLR, 2006, Table 1). The estimated IUU series and legal removals were used (SC-CAMLR, 2006, Table 1) as fixed values assumed known without error.

Standardised CPUE

The method for standardising catch-and-effort time series data is described in Candy (2004) and was used to provide a CPUE series of yearly values, and corresponding CVs for each estimated value, for each of the main trawl grounds (Grounds B and C) up to and including 2007. These standardised CPUE values were used as a series of relative abundance observations. The catchability constant (q_CPUE) was an estimated parameter calculated separately for each of the two CPUE series.

Commercial catch length composition

Random length samples were obtained from commercial catches and binned by observers in 10 mm bins. In the model, these length-frequency (LF) data were aggregated into 100 mm bins. The length distributions were obtained as a proportion of catch in 100 mm length bins from 200 to 1 900 mm along with the associated sample size. To account for over-dispersion of the LF data relative to a multinomial distribution, the actual total number of fish across bins sampled from all hauls in the fishing method, ground, fishing year and season were replaced with an estimated ESS. The method of determining the ESS was that described in Candy (2008) using the fit of a gamma generalised linear model (GLM) to the empirical CV of the haul-by-haul proportions across bins.

For each fishery (i.e. fishing method by ground combination) a separate age-dependent selectivity function was fitted but with additional selectivity functions fitted for the pre-2006 compared to the 2006–2007 fishing seasons for Trawl Ground B, and for the late (s3) season compared to the combined early (s1, s2) seasons also for Trawl Ground B. Either the DNP model or the DN function were fitted using the same rationale as for the surveys (i.e. when a_2 hit the lower bound of 0.1).

Mark–recapture data

The mark–recapture data have been updated since Constable et al. (2006a) to include recaptured tagged fish that had not been measured for length and to exclude fish that were likely not to have mixed with the local population. In addition, the mark–recapture data have been better utilised in analyses by providing estimates of the over-dispersion parameters for the tag data as well as estimates of detection probability by fishery and season for recaptured tagged fish, estimation of tag-loss rates for dart tags, and improved estimates of the number of fish scanned for tags (Appendix 1).

For tags released from 1998 to 2007, separate tag categories were given for each combination of year and fishery in which fish were caught and released. The releases and recaptures were therefore aggregated across season within fishing year and fishery. In each case the number of tags released was given along with the proportion of tags in each 50 mm size bin. Tag-related mortality was set to 0.1 (Agnew et al., 2006). Fish recaptured in the same year and season or recaptured within 60 days of their release date were excluded from both release and recapture numbers. These restrictions were placed on the mark–recapture data in order to ensure that a reasonable level of random mixing of the retained tagged fish with the untagged population occurred at least at the local population level. In Constable et al. (2006a) the release data were aggregated across seasons for the trawl fisheries. Figure 1 shows the mark–recapture data for the main commercial grounds while Table 2 gives total numbers of fish released and recaptured by year given the above

Table 1: Total removals (tonnes) for Division 58.5.2 by fishery and season. Minor catches outside the main grounds are not included in this assessment. As a result, the total removals here may not be the same as the total removals for Division 58.5.2.

Fishery[1]	Season[2]	Fishing year											
		1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007[3]
Trawl B (f2)	Early (s1)	0	0	0	2969	1962	1028	290	147	897	1128	69	386
Trawl B (f2)	Mid (s2)	0	1813	2737	311	1310	1229	923	1593	612	116	616	154
Trawl B (f2)	Late (s3)	0	0	350	1	248	523	568	610	502	0	604	0
Trawl C (f3)	Early (s1)	0	0	0	111	34	9	1	3	17	8	13	15
Trawl C (f3)	Mid (s2)	0	0	341	26	1	80	926	213	114	782	403	16
Trawl C (f3)	Late (s3)	0	0	173	0	0	51	4	5	76	0	42	0
Longline C (f5)	Mid (s2)	0	0	0	0	0	0	0	94	211	161	386	217
Longline D (f6)	Mid (s2)	0	0	0	0	0	0	0	176	306	380	241	66
Longline E (f7)	Mid (s2)	0	0	0	0	0	0	0	0	25	95	0	0
IUU (IUU)		3000	7117	4150	427	1154	2004	3489	1274	531	265	112	0
Total fishery removals		3000	8930	7751	3845	4709	4924	6201	4115	3291	2935	2486	854
Surveys		0	0.3	1.5	93	7	45	35	13	65	21	12	12

[1] Names in brackets are the codes used to refer to the fishery in CASAL.
[2] Names in brackets are the codes used to refer to the season in CASAL.
[3] Catches for part of season completed.

Table 2: Summary of tagging data used in fit of the *a2-tags* and *a2-tags-calpe* population/fishery models.

Year of release:	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Number released[1]:	992	704	1168	1413	1268	1397	1260	1378	2130	956
Year of recapture	Number recaptured									
1998	2									
1999	58	6								
2000	24	71	48							
2001	10	19	94	74						
2002	10	2	65	56	84					
2003	3	1	11	34	134	73				
2004	2	1	8	10	36	110	116			
2005	1	0	1	0	12	23	109	18		
2006	1	1	1	1	4	18	22	79	73	
2007	0	0	0	0	0	0	5	13	141	66
Total recaptures	111	101	228	175	270	224	252	110	214	66

[1] Number of tags released and recaptured excludes tags for which recaptures occurred within the same fishing season (early, mid, late) as that of the release for the same year of release and/or those recaptured within 60 days of the release date.

restrictions. In total, 12 666 fish were released and 1 751 recaptured given the above restrictions. Without these restrictions the corresponding numbers of released and recaptured fish (excluding tagged fish caught outside Division 58.5.2) were 15 190 and 3 131 respectively (Welsford et al., 2007).

The number of tag–recaptures in each fishery and length bin were treated as observations and contributed log-likelihood components to the overall maximum likelihood estimation procedure for population and fishery selectivity parameters. For recaptured fish that were not measured for length, often because they were processed before observers had a chance to measure them, their recapture length was estimated from their release length, the time at liberty and the Fabens form of the von Bertalanffy growth model described in Candy et al. (2007). The bias correction described by equation (8) of Candy et al. (2007) was applied to the predicted recapture lengths, although this correction is only slight being less than 1%. For fish that had neither release nor recapture lengths measured, but had a trunk length at recapture, this length was used to predict total length. More detail on the methods used to estimate tag-detection rates and the dispersion parameter for the tag-data likelihood function are described in Appendix 1.

The other refinement to the recapture data compared to that used for the work described in Constable et al. (2006a) was in the calculation of the number of scanned fish. The number of scanned fish is the sample size of fish of which a fraction are 'observed' to be tagged. This is simply the sum across hauls within a CASAL fishery and recapture year of the estimated number of fish caught by length bin. This estimate involves disaggregating the estimated total number of fish caught in the haul to length bins using the subsample of the catch for which random length-frequency data is obtained. The estimated total number of fish caught is obtained by dividing the haul total catch weight by the mean weight of fish in the haul. However, when there are few fish measured for weight in a haul, the accuracy of the predicted total number of fish caught can be poor. To overcome this, the aggregate length-frequency data and weighed subsamples were obtained for each cruise and the mean weight of fish by length bin and the length bin proportions were used to obtain a weighted estimate of mean weight of fish. For hauls with less than five fish measured for weight, the cruise-level mean weight described above was used for the haul. Haul-level numbers of scanned fish by length bin were then aggregated across hauls within each combination of fishery and year.

Mean detection rates for recaptures calculated using the methods described in Appendix 1 for the fisheries for surveys (f1), Trawl Ground B (f2), Trawl Ground C (f3), and all of the longline fisheries combined were 0.9848, 0.9799, 0.9449 and 0.9930 respectively.

Model parameters and estimation

The parameter set to be estimated by CASAL consisted of B_0 pre-exploitation median spawning stock biomass, CV_{VB}, 22 YCS parameters (i.e. for years 1983 to 2004 inclusive), and a total of

37 selectivity parameters across survey groups and fisheries by grounds. Nuisance q-parameters are also estimated by CASAL but these, unlike the above parameters, do not directly affect the behaviour of the population/fishery model. The parameter, R_0, is the stock's average recruitment in the absence of fishing which is obtained in CASAL from B_0 assuming the pre-fishery stock is in equilibrium. Therefore, the annual historic recruitment series is obtained by multiplying R_0 and the estimates of YCS.

The weight-to-length relationship, maturity ogives, and the length-at-age vector are given in SC-CAMLR (2007a, Appendix L, Table 3).

Five variations of the population/fishery model were fitted. First, the model labelled *a2-ess* involved all the data except the tagging data. The second model, *a2-tags*, included the tagging data. The effect of the assumption of $q = 1$ for Survey Group 1 was explored by further varying this model to estimate catchability, q, of this main survey group, labelled *a2-tags-q1*. Previous work indicated that the CASAL estimation of this q simultaneously with the other parameters, particularly B_0, was only successful if alternative information on absolute abundance to the Survey Group 1 data was available and, in this case, this was provided by the tag data. The above population/fishery models used data weightings implicit in the log-likelihood definition for each dataset with process error set to zero and incorporating: (i) CVs of abundance estimates for each length bin for the survey data with estimates assumed to be lognormally distributed, (ii) CVs of the standardised CPUE series estimates data with estimates assumed to be lognormally distributed, (iii) multinomially distributed catch-at-length frequencies with ESS by fishery and year determined by the method of Candy (2008) for haul-level heterogeneity in frequencies, and (iv) binomially distributed number of tag returns by length bin conditional on number of scanned fish with a common over-dispersion parameter (Appendix 1). All of the above estimates of CV, ESS and the over-dispersion parameter were obtained prior to fitting the population/fishery model. The fourth population/fishery model fitted involved the effective down-weighting of the commercial catch-at-length data using the method of incorporating process error for this data described by Candy (2008). This method scales down each ESS by dividing by the fishery-specific over-dispersion parameter, which is obtained as a population/fishery model lack-of-fit statistic, and therefore requires a minimum of two CASAL fits in order to calculate this statistic. This population/fishery model, which also excluded the tagging data, is labelled *a2-calpe*. The final model variation examined was to carry out the above procedure for down-weighting the commercial catch-at-length data from the fit of the *a2-tags* model to give the *a2-tags-calpe* model. The justification for this re-weighting in the last two models is given under 'Discussion' after the method of handling process error and the results are described. The incorporation of process error, and thus data re-weighting, for all datasets used for the *a2-tags* model (i.e. survey data and CPUE data as well as the catch-at-length data) was carried out but this re-weighting was considered inappropriate for reasons discussed later.

Likelihoods and prior distributions

All parameters were assigned 'uniform' prior distributions as defined in Bull et al. (2005). Using their definition of the objective function as the negative log of the posterior probability of the parameter set, this prior distribution's contribution is set to zero so that the objective is simply the negative log-likelihood. Log-likelihoods corresponding to probability density functions assumed for each dataset were that of the lognormal (survey abundance data by length bin, standardised CPUE series), multinomial with sample size given by estimated ESS for each fishery and year combination (catch-at-length data), and binomial with overdispersion parameter (tag–recapture numbers by length bin with binomial sample sizes given by corresponding number of scanned fish). Standard definitions of these log-likelihoods were used as described in Bull et al. (2005).

Informative priors were used in the Constable et al. (2006a, 2006b) for B_0 and YCS using CASAL's uniform-log prior for B_0 (i.e. mildly informative) and lognormal prior for the YCS parameters with lognormal mean of $\mu = 1$ and CV of $CV_{YCS} = 1.1$ (i.e. highly informative if the CV is set relatively small compared to that obtained purely from the likelihood). If p defines a parameter in general and $\pi(p)$ its prior density function then the uniform-log prior adds the component $-\log\{\pi(p)\} = \ln(p)$ to the negative log-likelihood while the lognormal prior adds the component $-\log\{\pi(p)\} = \log(p) + 0.5[\log(p/\mu)/s + s/2]^2$ where $s = \sqrt{\log\left(1 + CV_{YCS}^2\right)}$ (Bull et al., 2005, p. 87). Both these informative priors for B_0 and YCS parameters were replaced here with $-\log\{\pi(p)\} = 0$. This is because the estimate of recruitment variability was found in Constable et al. (2006b) to be highly influenced by the value of CV_{YCS}. Since there is virtually no independent

Table 3: Results of assessments of stock status of *Dissostichus eleginoides* in Division 58.5.2 using CASAL. B_0 is the MPD estimate of the pre-exploitation median spawning stock biomass, CV_{VB} is the coefficient of variation for length-at-age, SSB status 2007 is the ratio of the CASAL prediction of SSB in 2007 to B_0, and R_0 is the MPD estimate of mean age-1 recruitment prior to exploitation (1981).

Model	Description	B_0 (SE)	CV_{VB} (SE)	SSB status 2007	R_0 (million)
a2-ess	Model *a1-50-notag-cl* in Constable et al. (2006b) + refinements	125 219 (5 806)	0.0977 (0.0008)	0.725	4.538
a2-calpe	*a2-ess* + C-a-L process error	152 332 (7 751)	0.0966 (0.0020)	0.818	5.525
a2-tags-calpe	*a2-calpe* + tag data	87 518 (1 625)	0.0702 (0.0016)	0.554	3.224
a2-tags	*a2-ess* + tag data	82 181 (1 303)	0.0954 (0.0008)	0.521	2.983
a2-tags-q1	*a2-ess* + tag data+ estimate *q* for SG1	78 314 (4 059)	0.1082 (0.0474)	0.470	2.817

information on recruitment variability available for this stock, subjectively setting a value of CV_{YCS} in a lognormal prior was considered to be unacceptable. Parameter estimates are referred to as maximum posterior density (MPD) estimates but in the case of all parameters being given a log-prior of zero (i.e. CASAL's 'uniform' prior) these are simply equivalent to maximum likelihood estimates.

Process error and systematic lack-of-fit

Assessment methods for the Ross Sea (Dunn et al., 2007) and South Georgia fisheries (Hillary et al., 2006) incorporate estimates of process error using an iterative procedure of fitting CASAL and then using lack-of-fit statistics to determine process error. These estimates of process error are then applied with an updated CASAL run and this two-step procedure is repeated until there are only small changes in the estimates of process error (A. Dunn, pers. comm.). However, Candy (2008) noted that for process errors to be random deviations between observed data and model fit, any SLOF, either across length bins or across years, should first be removed. This iterative method of estimating process error with the refinement of first removing SLOF was investigated as described below.

Fitting a simple SLOF model was achieved here by fitting to the deviations between observed and CASAL-fitted values both linear and quadratic terms in the continuous values of each of length bin and year including the interaction between years and length bins for these terms. This was done for the catch-at-length proportions and Survey Group 1 abundances. The method used to fit the SLOF model and the downward adjustment to the ESS due to residual process error is described for catch-at-length data in Candy (2008). For the Survey Group 1 abundances, since these are assumed to be lognormal, the SLOF model was fitted to the CASAL residuals on the log scale taking into account individual year by length bin CVs.

For CPUE data only a continuous-year SLOF model was fitted since there are no length bins, while for the single-year survey groups a continuous-bin SLOF model was fitted and process errors obtained as lognormal CVs in each case. For both these types of data and the multi-year survey (Survey Group 1) the SLOF analyses and calculation of process error took into account the individual year CVs by appropriately weighting the SLOF regressions (Appendix 2).

Results

Stock assessment

Table 3 summarises for each population/fishery model the maximum likelihood estimates (with standard errors (SEs) where available from CASAL) of median pre-exploitation spawning biomass (B_0), the status of the spawning stock in 2007 relative to B_0, median pre-exploitation Age 1 recruitment (R_0), and the estimate of CV of length given age, CV_{VB}.

Table 4 gives the estimates of the parameters for the DNP selectivity function for Survey Group 1 and their SEs. Parameter estimates only for Survey Group 1 are presented for brevity and because this survey group is considered to be most influential in providing estimates of YCS parameters. However,

Table 4: Estimates of selectivity parameters in Survey Group 1 and catchability of the survey groups in assessments of stock status of *Dissostichus eleginoides* in Division 58.5.2 using CASAL.

Model	Description	Selectivity parameter estimates Survey Group 1 (SE below estimate)				Survey group q estimate[1]				
	Double-normal with plateau	s_L	s_U	a_1	a_2	SG3 1990	SG5 1993	SG2 1999	SG7 2003	SG1
a2-ess	Model *a1-50-notag-cl* in Constable et al. (2006b) + refinements	0.024 (0.002)	4.586 (0.151)	2.465 (0.041)	1.839 (0.326)	0.304	0.304	3.468	0.843	1[1]
a2-calpe	*a2-ess* + C-a-L process error	0.021 (0.001)	4.938 (0.101)	2.261 (0.107)	0.639 (0.063)	0.1904	0.1904	1.9923	0.526	1[1]
a2-tags-calpe	*a2-calpe* + tag data	0.467 (0.068)	4.592 (0.109)	3.216 (0.098)	2.640 (0.150)	0.390	0.390	2.625	0.677	1[1]
a2-tags	*a2-ess* + tag data	1.164 (0.154)	4.573 (0.200)	4.009 (0.156)	1.726 (0.374)	0.545	0.545	4.434	0.997	1[1]
a2-tags-q1	*a2-ess* + tag data + estimate q for SG1	0.025 (0.120)	5.991 (0.128)	2.279 (0.239)	0.783 (0.382)	0.750	0.750	2.875	0.711	1.312

[1] Catchability q set to 1 for Survey Group 1 (2001, 2002 and 2004 to 2007).

Figure 2(a) shows graphically the fitted DNP and DN selectivity curves for the *a2-tags-calpe* model for all survey groups and commercial fisheries. Figure 2(b) shows the fitted selectivity function to Survey Group 1 for each of the population/fishery models corresponding to the parameter estimates given in Table 4. The shape of the left- and right-hand limbs of the curves are controlled by the normal density standard deviations given by s_L and s_U respectively. The age at which maximum selectivity of 1 first occurs is given by parameter a_1 and the width of the plateau was described earlier as determined by parameter a_2. The curves in Figure 2(a) show the distinct differences in how the surveys and trawl and longline activities overlap with the stock, notably that the surveys observe the youngest fish (less than age 5), the trawl fishery concentrates on larger but immature fish, and the longline fishery concentrates on larger fish again but with few mature fish. The notable exception is for the last two fishing seasons in Trawl Ground B for which the fitted selectivity function (*Sel_f2_s2r*) indicates that fish younger than 5 years have been selected.

Catchability, q, estimates are given in Table 4 for all other survey groups noting that q was fixed to 1 for Survey Group 1 for all but the *a2-tags-q1* model. When q was estimated for this survey group without the tag data being present, the estimation, particularly for B_0, became unstable (see 'Discussion').

Tables 3 and 4 show that, for the non-tag models, the estimate of B_0 increased substantially for model *a2-calpe* compared to model *a2-ess.* At the same time the length of the plateau of the Survey Group 1 selectivity, parameter a_2, decreased substantially for model *a2-calpe* down to a value of 0.64, thereby reducing the selectivity on ages 4 and 5. Survey Group 1 is the primary dataset for determining the magnitude of recruitment (q = 1) and, therefore, B_0. As a result, the estimate of B_0 will be sensitive to shifts in the selectivity of this survey group if catchability, q, is fixed. To investigate this further, for q fixed at 1, the a_2 parameter was fixed at 2 and model *a2-calpe* refitted. The resulting estimates of B_0 and R_0 were substantially reduced to 105 502 tonnes (SE = 3 772) and 3.860 million respectively. The implications of this key sensitivity of the assessment to the estimation of Survey Group 1 selectivity is discussed later.

When comparing *a2-tags* and *a2-tags-q1* which has q fixed at 1 and estimated at 1.3 respectively, it can be seen from Tables 3 and 4 that although the estimate of B_0 is similar for these two models, the estimated selectivity parameters are quite different. Therefore, it can be deduced that estimates of selectivity function, B_0 and q parameters for Survey Group 1 are inextricably linked. The other key point to note from Table 3 is that only for the *a2-tags-calpe* model did the estimate of CV_{VB} vary significantly (both statistically and practically) from the 0.1 estimate obtained by Candy et al. (2007). This parameter effectively determines, along with the length-at-age mean vector, how length-binned data (both abundances and relative frequencies) are disaggregated to age classes. With a smaller value of CV_{VB} there is greater potential to obtain a larger number of consecutive age classes that contribute to a given length bin.

Model fits

Figure 3 shows the contribution to the total value of the objective function (i.e. sum of the values of the negative log-likelihood for each dataset in the model) for each population/fishery model and each dataset in that model. CASAL estimation, in this case maximum likelihood, attempts to minimise the total value of the objective function, so smaller values indicate a better fit.

The comparison of the objective function values obtained from CASAL for the catch-at-length proportions between the *a2-calpe* and *a2-tags-calpe* models and any of the other models are uninformative because the data have been changed between models due to the change in ESS values in the *a2-calpe* and *a2-tags-calpe* models. However, for Figure 3 the objective function values for the catch proportions and the *a2-ess*, *a2-tags* and *a2-tags-q1* models have been scaled as if they had the same ESS values as the *a2-tags-calpe* model to allow this comparison to be valid.

By comparing objective values in Figure 3 across models within each dataset, the tension between the datasets obtained from commercial fishing (i.e. catch-at-length proportions and tag–recaptures) and the Survey Group 1 data can be seen:

(i) tag–recaptures: the better fit of the *a2-tags* model compared to that of the *a2-tags-calpe* model shows that when the commercial catch-at-length data are given less weight (by using lower ESSs in the latter model), the fit to the tag data is worse.

(ii) Survey Group 1 abundances: the better fit of the *a2-calpe* and *a2-tags-calpe* models compared to that of the *a2-ess* and *a2-tags* models respectively, shows that when the commercial catch-at-length data are given less weight (by using lower ESSs in the former models), the fit to the survey data is improved.

(iii) Catch_proportions: the fit has been adjusted for differences in ESS to correspond to fitted numbers per bin using the same ESS as the *a2-tags-calpe* model. The *a2-calpe* and *a2-tags-calpe* models fit worse than the *a2-ess* and *a2-tags* models respectively, since the former two models giving less weight to these data.

(iv) Other_surveys abundances: the fit is also very similar across all models for the other_surveys indicating that this data is not very informative for discriminating between models (since the fit has been constrained by the estimated q's and selectivity functions fitted separately for each year of survey, with the exception of the pooled estimate of q for the 1990 and 1993 surveys).

Points (i), (ii) and (iii) indicate, along with the results in Tables 3 and 4, that the tag data and commercial catch-at-length data are 'pulling' the parameter estimates (particularly B_0 and Survey Group 1 selectivity parameters) in a similar direction whereas the Survey Group 1 data is pulling these parameters in the opposite direction.

Diagnostics plots of the fits to the different datasets are shown for the *a2-tags-calpe* model in Figures 4 to 9. The fits to these data for the other three models (not shown) were visually quite similar as expected from the similarity of objection function contributions across models shown in Figure 3 (i.e. accounting for the discussion on the effect of the reduced ESS in two of the models). Figure 4 shows the fit to the Survey Group 1 abundance data. Figure 4(a) shows observed and fitted numbers while Figure 4(b) shows deviations of log of observed minus log of fitted abundance along with 95% confidence bounds determined from the CVs of observed (i.e. stratified random sample mean) abundances. Using the log scale prevents high abundances from dominating the visual perception of lack-of-fit as is the case in Figure 4(a). Figure 4(c) shows deviations as in Figure 4(b) but overlayed with the fitted SLOF trend and its approximate 95% confidence bounds. Fitted values in Figure (4a) show a consistent underestimation compared to observed abundances for the length bins that contain most of the fish. This indicates that either the abundance of young fish as determined by other datasets is not as high as that observed in the surveys, or that the survey selectivity for these fish has been underestimated. This underestimation was not rectified by estimating catchability, q. When q was estimated in the fit of the *a2-tags-q1* model because the estimate of q was greater than 1 (Table 4) then, as expected, the fitted abundance for these length bins was closer to the observed values (graph not shown) for some of the survey years. However, overall the fit was only very slightly better than the corresponding model that fixed q at 1, that is the *a2-tags* model (Figure 3).

Figure 5 shows the fit in that model for the remaining 'single-year' surveys. The estimate of q obtained for each of the early surveys (Table 4) shows that the 1999 survey (Survey Group 2) was likely to be overestimating the abundance of recruits while the other surveys (1990, 1993 and 2003) were underestimates.

Figure 6 shows the fit to the commercial length-frequency data for the main trawl fishery (Ground B) for within-year season 2. For the sake of brevity, only the fit to this set of length-frequency data is given as a demonstration, but the fits were generally very similar in quality both across datasets and across models apart from the slightly worse fit for the *a2-calpe* and *a2-tags-calpe* models (as Figure 3 demonstrates in this last case). Figure 6(a) shows the observed and fitted proportions while Figure 6(b) shows the deviance residuals for the SLOF model fitted to observed and predicted frequencies (i.e. proportions multiplied by ESS) (Candy, 2008). The bounds in Figure 6(b) are based on the 95% critical value of a chi-square distribution with single degree of freedom. Since deviance residuals are not independent, these bounds could be slightly conservative (i.e. wider) compared to those based on the true distribution of these residuals. Determining the true distribution is not possible due to the difficulty in rigorously attributing degrees of freedom to the SLOF model fit since this model treats predicted proportions as known constants and not as estimates obtained in part from the same observations.

Figures 7 and 8 show the standardised CPUE series versus the fitted trend from the population/fishery model for each of the trawl grounds respectively. Note that the standardised CPUE series in each case was obtained from the haul-by-haul data combined across all three seasons based on the standardisation model given by (Candy, 2004) and updated using data up to and including 2007. The contribution to the objective for the CPUE data was relatively small in each case due to the generally large CVs of the standardised estimates.

Figure 9 shows the fit to the tag–recapture numbers for the *a2-tags-calpe* model for releases and recaptures restricted to Trawl Ground B (f2). Figure 10 shows the estimates of YCS from 1983 to 2004.

Process error

As an example of the effect of between-haul heterogeneity on ESS, the method of Candy (2008) for catch-at-length data gave an ESS of 2 756 for 2003 in season 2 of the Trawl Ground B fishery when there were actually 13 089 fish measured for length from 490 hauls. When process error was estimated using the CASAL-fitted values, the fit of the SLOF model, and the residual Poisson deviance, the ESS was reduced from 2 756 to 518 at the first iteration of the two-step CASAL/process error estimation procedure for the *a2-tags* model. At the second iteration, the ESS decreased to 485. At the same time the lognormal process error CV for Survey Group 1 was 1.225 at the first iteration and decreased to 1.193 at the second iteration. However, the diagnostic graph (not shown) corresponding to Figure 4(c) for both iterations showed, similarly to Figure 4(c), that only for 2002 could the deviations about the smooth SLOF trend be considered random across bins. Clearly for the other years there is still non-random lack-of-fit remaining after fitting the simple SLOF model so that attributing these deviations entirely to process error is statistically invalid since process error by definition should be random (Candy, 2008). In practical terms, the incorporation of the resultant 'contaminated' process error variance into the CASAL fit will result in the survey data being overly down-weighted. The effect of this will be to exacerbate the SLOF for the survey data. In contrast, Figure 6(b) demonstrates that, for deviance residuals for the SLOF model fit to the catch-at-length frequencies for the Trawl Ground B, season 2 fishery, any bias or a consistent trend away from the zero line is not obvious. A similar conclusion can be drawn for the corresponding graphs (not shown) for the other fisheries/seasons. For this reason the two-step CASAL/process error estimation procedure, as applied to all datasets, was not used. However, as described earlier, the ESS calculated for the catch-at-length proportions, due to haul-level heterogeneity, was reduced by the effect of process error in the re-weighting of these data from the *a2-ess* and *a2-tags* models to the *a2-calpe* and *a2-tags-calpe* models respectively, using the first iteration of the two-step procedure described above (e.g. in the above example the ESS values used for the *a2-tags* and *a2-tags-calpe* models were 2 756 and 518 respectively).

Discussion

This paper follows preliminary work in 2005 and 2006 in developing an integrated assessment for *D. eleginoides* in Division 58.5.2. It uses all relevant and available data in the assessment, including surveys, fishery catch-at-length data, standardised CPUE series for trawl grounds, and length-binned mark–recapture data. A number of alternative data-weighting schemes were investigated. The weighting schemes, based on ESSs for the catch-at-length data, combined with measures of uncertainty of the other data that were obtained prior to the estimation carried out using CASAL (and therefore did not include process error) have a formal statistical basis. The inclusion of process error to down-weight the catch-at-length data can also be justified statistically, given the caveat that systematic lack-of-fit has been adequately modelled. Using residual lack-of-fit in this way is commonly used in statistical procedures to account for over-dispersion (see for example, Candy, 2002), however, reconciling residual lack-of-fit across diverse datasets, such as those considered in integrated fishery assessments, presents an extra challenge compared to the statistical analysis of a single type of data. The inclusion/exclusion of the tag data is an extreme form of data weighting. There is now little scope for substantially refining the assessment given the current model structure and datasets. The results of the different models presented here show that the spawning stock status may range from around 50 to 80% of pre-exploitation median abundance. The assessment scenario for use in managing catch limits in this fishery will depend on which datasets are chosen to reflect attributes of the stock and the relative weight given to each dataset.

Use of commercial data

Not surprisingly, the commercial length-composition data have a substantial influence on the assessment. The combination of fishing gears with seasons and grounds, and the obviously different selectivity functions estimated for each, indicates a complex spatial relationship between the fishery overall and the stock. Indeed, the changing nature of the proportions-at-length for the trawl fishery in Ground B, the main fishing ground, when compared to the Survey Group 1 dataset indicates that, particularly in the last two years of the series, either the age-specific distribution of the stock has changed (where this could be due to temporal or spatial changes or both) or that the behaviour of this fishery relative to the stock has altered.

The calculation of effective sample size gives a substantial reduction in weight of the commercial catch-at-length data. This is justified because of the need to account for spatial and temporal heterogeneity in fishery behaviour as well as the impact of different sample sizes per haul in the number of fish being measured. Candy (2004) found that there was significant small-scale (i.e. subarea) area and area by year variation in CPUE in Trawl

Ground B. This targeting is likely to introduce clustering of hauls and therefore positive small-scale spatial correlation between catch-at-length data at the haul level. The method (Candy, 2008) of determining an appropriate ESS for these data which reduces the multinomial sample size of number of fish measured per fishing year, season and fishery assumes that the hauls are independent. A theoretical method of incorporating haul-level spatial correlation into the calculation of an ESS is not available, so an alternative implemented here is to use residual lack-of-fit to estimate process error and, as a rather crude approximation, scale down the naïve ESS using process error expressed as an overdispersion parameter (Candy, 2008). An alternative would be to incorporate random effect terms into the fishing selectivity functions in an analogous way to the incorporation of random subarea-by-year interactions in the log-linear CPUE model given by Candy (2004). CASAL cannot currently incorporate random effects. Also, some of these issues may be better resolved when more ageing of the historical otolith collection is carried out (Welsford and Nowara, 2007).

The standardised CPUE series for each trawl ground do not influence estimation to any practical degree due to the very high uncertainty for the annual estimates resulting in a negligible contribution to the objective function. The generally poor fit, particularly in Trawl Ground B, where the decline in observed standardised CPUE is not replicated by the model, remains of concern. Figure 10 shows a recent increase in values of YCS for 2000 and 2001 which results in an increase in predicted abundance of corresponding age-6 and age-5 cohorts in 2006 respectively, and similarly for 2007 when these cohorts are age 7 and 6 respectively. The increased abundance of these age classes which are fully, or close to fully, selected by the trawl fishery would explain why predicted CPUE in Figures 7 and 8 increases in 2006 and 2007. One possibility to be explored in the future for improving the incorporation of CPUE data is to ensure that in the estimation a direct relationship exists between standardised CPUE data and catch length-composition data, which are different attributes of the same set of samples at the haul level. At present, the standardised CPUE and the length-composition data are not directly linked at the haul-level for parameter estimation. The haul-by-haul CPUE and the length-composition data ideally would be standardised within the fit of the integrated assessment model so that, when the hauls are combined, they both reflect attributes of the same component of the stock.

Use of survey data

The survey data are the only fishery-independent data and were thus able to be obtained by stratified random sampling. Therefore, Survey Group 1 (years 2001, 2002 and 2004 to 2007) is assumed to provide unbiased absolute estimates of recruitment (Welsford et al., 2006b) given that a catchability of 1 is correct to a reasonable approximation. For models that incorporate the tag data it was possible to estimate q for Survey Group 1 (model *a2-tags-q1*). The estimate (Table 4) was 1.3 compared to the model that fixes q at 1 (model *a2-tags*). However, there was a corresponding significant change in selectivity parameters (also see Figure 2b) between these two models for Survey Group 1 (Table 4) which resulted in some compensation for the estimated q since the estimate of B_0 for both models was quite similar (Table 3). Therefore, given the comparison of the results for the two models with tag data, fixing the catchability, q, at 1 for Survey Group 1 was considered reasonable. When tag data was not used, this q could not be reliably estimated since estimation for the *a2-ess* model became unstable with wildly oscillating estimates of B_0 during the iterative search to minimise the objective function. Even though convergence was eventually achieved in some runs using different starting values for parameters, these resulted in very different estimates of B_0. The combination of these behaviours strongly suggests that a search could easily have been 'trapped' in the parameter space around a local minimum and, in general, the likelihood surface is relatively flat with respect to this parameter.

These results obtained by comparing model fits cast doubt on the automatic assumption that estimation of q as a nuisance parameter gives an estimate of true catchability. Alternatively, estimation of q can be seen simply as a device for improving the fit to absolute abundance data. This can follow since the lack of fit when q is fixed at 1 may not be due to the actual catchability of the survey being substantially different to 1 but could be due to the influence of other datasets or an inadequate structure and/or parameterisation of the population/fishery model. The evidence in this case supports the conclusion that tensions between datasets is the issue with data obtained from commercial hauls indicating a lower abundance than the survey data. For q to be a reliable estimator of catchability requires the assumption that these other datasets, after manipulation within the population/fishery model, also accurately quantify the stock at the time of sampling.

The degree to which a length bin is estimated to be selected by the trawl fishery will be dependent on the errors in the observations, which were relatively large, combined with the relative influence of other data, including other length bins in the survey data and different datasets. The consistent underestimation of fish in the surveys (observed abundances are greater than those expected from the population model) could be due to a number of factors. These include the possibility that the estimate of biomass, B_0, is too low or the estimated selectivity of the older ages in the survey is too low (i.e. increasing the B_0 or selectivities could give expected abundances in the length bins of older fish in the surveys equivalent to the observed abundances). Such errors could be due to the influence of other datasets on parameter estimates, as mentioned above, or the effect of too high a value for M or M may in reality vary substantially with age. In addition, the disaggregation of length bins to age classes using the length-at-age vector with lognormal distribution with estimated CV_{VB} may be overly simplistic, and so on. However, it is clear (Tables 3 and 4) that when the influence of commercial length-composition data is reduced (*a2-calpe* versus *a2-ess* and *a2-tags-calpe* versus *a2-tags*), the estimate of B_0 is increased and for the former of these models, that exclude the tag-data, selectivity for ages above 4 years is reduced (Figure 2b). Inclusion of datasets or external information to provide greater contrast for estimating selectivity of the surveys will be very important for accurately estimating B_0.

Of primary importance in interpreting predicted abundance over years is to note that this depends strongly on number of recruits in each year, which is given by the YCS estimate multiplied by R_0, their projected growth, and the disaggregation of length bins to age classes. Inaccuracy in this disaggregation could be largely responsible for the strong oscillation in YCS estimates between 1989 and 1997 seen in Figure 10. Therefore, the recruitment trend for this period may be spurious.

Use of mark–recapture data

The issue of excluding versus including the tag data is equally, if not more, difficult to reconcile than the weighting applied to commercial length-composition data, particularly when the mark–recapture data, to date, primarily have been obtained from the trawl fishery in Ground B. The inclusion of both length composition data and tag returns for the fisheries generally results in lower stock size estimates and correspondingly a greater selectivity of age-5 and 6 fish in Survey Group 1. This can be seen when the tag data are included (cf. *a2-ess* versus *a2-tags*) in the fit and when the influence of the catch-at-length data is down-weighted by scaling the ESS using process error (cf. *a2-calpe* versus *a2-tags-calpe*).

The fits to the observed tag data show that the tag–recaptures (Figure 9) would suggest that the stock size is even less than that estimated by the models that include the tag data or that the estimated selectivities for these fisheries are greater than they should be. This tension between the tag data and the survey data remains to be reconciled but may in part be explained by the dominance of the data from the fishery in Trawl Ground B in the analysis.

As noted earlier, the tension between the survey dataset and the datasets derived from fishing operations (catch at length, mark–recapture) in their influence on parameter estimates could be a product of the different spatial coverage of the respective operations combined with poor mixing of the population between grounds. Clearly, if the stock is well mixed across both the fishing grounds and the remaining unfished areas, then the surveys and the fishing operations will provide an assessment for the whole area. However, as the mark–recapture data indicate (Figure 1) that the grounds may be relatively isolated from one another, the assessment based on tag data will be drawn towards a sum of the stocks in each major fishing ground rather than for an aggregate of the whole area, the latter of which is better represented by the surveys.

Further work could consider the effect of the retarding of growth due to 'tagging shock' (Hillary et al., 2006; Agnew et al., 2007). The adjusted von Bertalanffy growth model (Candy et al., 2007), used to project released/recaptured fish into larger length bins when the required increment is predicted to be achieved, does not account for tag shock. In Candy et al. (2007) predicted growth from their model overestimated the annual growth rate of recaptured fish (approximately 40–50 mm year^{-1} versus 30–40 mm year^{-1} respectively, see Figure 5 of Candy et al., 2007).

Concluding remarks

This paper presents an integrated assessment using CASAL for *D. eleginoides* in Division 58.5.2 which is an improved assessment compared to those presented previously. However, as expected, the assessment was sensitive to the inclusion of different datasets, including the implicit weight given to each dataset via measures of uncertainty about data inputs, and the choices of parameters used in the stock assessment.

In terms of statistical methods, this assessment has applied rigorous statistical weighting procedures for the fishery datasets commonly used in integrated assessments. The implemented method of Candy (2008) for calculating ESS of the catch-at-length data in order to account for haul-level variability is superior to the method of Dunn et al. (2007) and greatly superior to the method of McAllister and Ianelli (1997) (see Candy, 2008). In addition, the method of quantifying process error applied here first removes SLOF from model predictions obtained from CASAL. Finally, the appropriate weighting of the SLOF regressions for lognormal 'observation' data (i.e. survey abundance estimates and CPUE series) with individual CV values, in order to estimate process error, is described (Appendix 2) and implemented in this assessment. However, for reasons given, the extension of incorporation of process error to these lognormal observation data was not satisfactory for this assessment because of the failure to adequately remove SLOF for the survey abundance data. This meant that after the first iteration of the CASAL/process error estimation procedure, the main survey series was overly down-weighted by process error variance since this variance was 'contaminated' with SLOF. For this reason process error variance (i.e. overdispersion) was only used to reduce the effective sample size for the catch-at-length data.

The validity of the high degree of influence that the commercial catch-at-length data has on CASAL's estimation of historic stock structure, due to the large number of hauls and measured fish for commercial fishing relative to those for the main survey series (RSTS Group 1), was explored via the incorporation of haul-level estimation of ESS combined with process error estimation to down-weight these data as mentioned above. The fact that commercial hauls, in contrast to the RSTS shots, are not a random stratified sample of the spatial extent of the stock results in a degree of uncertainty about the representativeness of the commercial data in reflecting stock structure in an unbiased way, particularly if the non-random behaviours of the vessels vary from one year to the next. Under these circumstances one could ask whether selectivity should be estimated at all, or whether catches-at-length in the model should simply be extracted from the population according to the proportions-at-length. This would be a valid procedure when the catch-length composition is well estimated simply by length-frequency sample proportions.

Improved prediction of historical recruitment trends compared to that given in Figure 10 and an overall improved understanding of the relative merits of the different datasets to the assessment of stock status will be achieved with the development of age–length keys for disaggregating the length data into age classes and the consequent introduction of a fully age-based model. Until then, the tensions between the datasets are unlikely to be resolved.

Acknowledgements

Thanks are due to Tim Lamb and Troy Robertson for processing the data, and to Dirk Welsford and John McKinlay for helpful comments and suggestions during the preparation of this work. Also the comments of members of the Australian Fisheries Management Authority's sub-Antarctic Resource Assessment Group are appreciated. We owe a great deal of thanks to Alistair Dunn for providing his time and expertise to help establish these assessments. We also thank members of WG-FSA and WG-SAM for freely giving their advice on this assessment. Finally, the comments of Dr Marinelle Basson, Dr Richard Hillary, Alistair Dunn, and an anonymous referee on an earlier draft of this paper, were very constructive and resulted in considerable improvement in the revision and presentation of this work.

References

Agnew, D.J., J. Moir Clark, P.A. McCarthy, M. Unwin, M. Ward, L. Jones, G. Breedt, S. Du Plessis, J. Van Heerdon and G. Moreno. 2006. A study of Patagonian toothfish (*Dissostichus eleginoides*) post-tagging survivorship in Subarea 48.3. *CCAMLR Science*, 13: 279–289.

Agnew, D.J., R. Hillary and J. Pearce. 2007. Preliminary assessment of the South Georgia toothfish stock, 2007. Document *WG-FSA-07/29*. CCAMLR, Hobart, Australia.

Bull, B., R.I.C.C. Francis, A. Dunn, A. McKenzie, D.J. Gilbert and M.H. Smith. 2005. CASAL (C++ algorithmic stock assessment laboratory): CASAL User Manual v2.07-2005/06/23. *NIWA Technical Report*, 126: 268 pp.

Butterworth, D.S., J.N. Ianelli and R. Hilborn. 2003. A statistical model for stock assessment of southern bluefin tuna with temporal changes in selectivity. *Afr. J. Mar. Sci.*, 25: 331–362.

Candy, S.G. 2002. Empirical binomial sampling plans: model calibration and testing using Williams' method III for generalized linear models with overdispersion. *JABES*, 7 (3): 373–388.

Candy, S.G. 2004. Modelling catch and effort data using generalised linear models, the Tweedie distribution, random vessel effects and random stratum-by-year effects. *CCAMLR Science*, 11: 59–80.

Candy, S.G. 2008. Estimation of effective sample size for catch-at-age and catch-at-length data using simulated data from the Dirichlet-multinomial distribution. *CCAMLR Science*, this volume.

Candy, S.G., A.J. Constable, T. Lamb and R. Williams. 2007. A von Bertalanffy growth model for toothfish at Heard Island fitted to length-at-age data and compared to observed growth from mark–recapture studies. *CCAMLR Science*, 14: 43–66.

Constable, A. and W.K. de la Mare. 1996. A generalised yield model for evaluating yield and the long-term status of fish stocks under conditions of uncertainty. *CCAMLR Science*, 3: 31–54.

Constable, A.J., S.G. Candy and I. Ball. 2006a. An investigation of integrated stock assessment methods for the Patagonian toothfish (*Dissostichus eleginoides*) in Division 58.5.2 using CASAL. Document *WG-FSA-SAM-06/14.* CCAMLR, Hobart, Australia.

Constable, A.J., S.G. Candy and I. Ball. 2006b. An integrated stock assessment for the Patagonian toothfish (*Dissostichus eleginoides*) in Division 58.5.2 using CASAL. Document *WG-FSA-06/64.* CCAMLR, Hobart, Australia.

de la Mare, W.K. 1994. Estimating krill recruitment and its variability. *CCAMLR Science,* 1: 55–69.

de la Mare, W.K., A. Constable, E. van Wijk, T. Lamb, D. Hayes and B. Ronai. 2002. *CMIX: User's Manual and Specifications*. Australian Antarctic Division, Kingston, Australia: 50 pp.

Dunn, A. and S.M. Hanchet. 2006. Assessment models for Antarctic toothfish (*Dissostichus mawsoni*) in the Ross Sea including data from the 2005/06 season. Document *WG-FSA-06/60*. CCAMLR, Hobart, Australia.

Dunn, A. and S.M. Hanchet. 2007. Assessment models for Antarctic toothfish (*Dissostichus mawsoni*) in the Ross Sea including data from the 2006/07 season. Document *WG-FSA-07/37*. CCAMLR, Hobart, Australia.

Dunn, A., D.J. Gilbert, S.M. Hanchet and B. Bull. 2004. Development of an Antarctic toothfish (*Dissostichus mawsoni*) stock model for CCAMLR Subarea 88.1 for the years 1997/98 to 2003/04. Document *WG-FSA-04/36*. CCAMLR, Hobart, Australia.

Dunn, A., D.J. Gilbert and S.M. Hanchet. 2005. A single-area stock assessment model of Antarctic toothfish (*Dissostichus mawsoni*) in the Ross Sea for the 2004/05 season. Document *WG-FSA-05/33*. CCAMLR, Hobart, Australia.

Dunn, A., S.M. Hanchet, S.L. Ballara and M.P. Francis. 2007. Preliminary investigations of an assessment for skates and rays in the Ross Sea. Document *WG-SAM-07/4*. CCAMLR, Hobart, Australia.

Gon, O. and P.C. Heemstra (Eds). 1990. *Fishes of the Southern Ocean*. J.L.B. Smith Institute of Ichthyology, Grahamstown, South Africa: 462 pp.

Hillary, R.M., G.P. Kirkwood and D.J. Agnew. 2005. An assessment of toothfish in Subarea 48.3 using CASAL. Document *WG-FSA-05/16*. CCAMLR, Hobart, Australia.

Hillary, R.M., G.P. Kirkwood and D.J. Agnew. 2006. An assessment of toothfish in Subarea 48.3 using CASAL. *CCAMLR Science*, 13: 65–95.

Kirkwood, G.P. and M.H. Walker. 1984. A new method for estimating tag shedding rates with application to data for the Australia salmon, *Arripis trutta esper* Whitely. *Aust. J. Mar. Freshw. Res.*, 35 (5): 601–606.

Lawes Agricultural Trust. 2002. *The Guide to GenStat Release 6.1*. Part 2: Statistics. VSN International, Oxford, UK.

McAllister, M.K. and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Can. J. Fish. Aquat. Sci.*, 54 (2): 284–300.

Maunder, M.N. 2003. Paradigm shifts in fisheries stock assessment: from integrated analysis to Bayesian analysis and back again. *NRM*, 16 (4): 465–475.

Nowara, G.B. and T. Lamb. 2007. Report on a random stratified trawl survey to estimate distribution and abundance of *Dissostichus eleginoides* and *Champsocephalus gunnari* conducted in the Heard Island Region (Division 58.5.2), May–June 2007. Document *WG-FSA-07/46*. CCAMLR, Hobart, Australia.

R Development Core Team. 2006. *R: a Language and Environment for Statistical Computing*. Version 2.4.0. R Foundation for Statistical Computing, Vienna, Austria: www.R-project.org.

SC-CAMLR. 1996. Report of the Working Group on Fish Stock Assessment. In: *Report of the Fifteenth Meeting of the Scientific Committee (SC-CAMLR-XV)*, Annex 5. CCAMLR, Hobart, Australia: 301–447.

SC-CAMLR. 2006. Report of the Working Group on Fish Stock Assessment. In: *Report of the Twenty-fifth Meeting of the Scientific Committee (SC-CAMLR-XXV)*, Annex 5. CCAMLR, Hobart, Australia: 259–478.

SC-CAMLR. 2007a. Report of the Working Group on Fish Stock Assessment. In: *Report of the Twenty-sixth Meeting of the Scientific Committee (SC-CAMLR-XXVI)*, Annex 5. CCAMLR, Hobart, Australia: 251–406.

SC-CAMLR. 2007b. Report of the Working Group on Statistics, Assessments and Modelling. In: *Report of the Twenty-sixth Meeting of the Scientific Committee (SC-CAMLR-XXVI)*, Annex 7. CCAMLR, Hobart, Australia: 511–553.

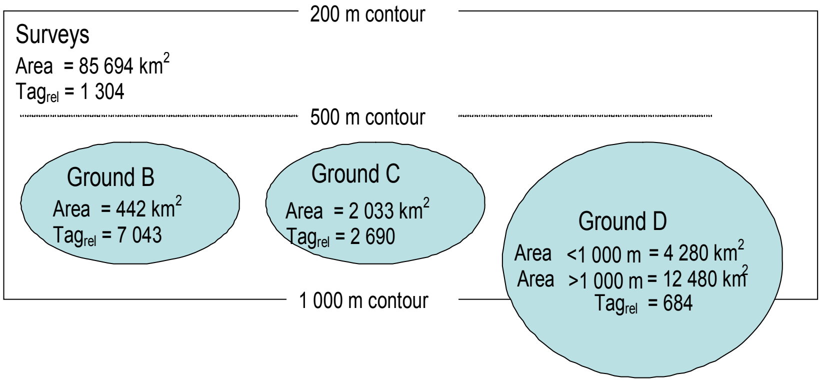
Tuck, G.N., W.K. de la Mare, W.S. Hearn, R. Williams, A.D.M. Smith, X. He and A. Constable. 2003. An exact time of release and recapture stock assessment model with an application to Macquarie Island Patagonian toothfish (*Dissostichus eleginoides*). *Fish. Res*., 63: 179–191.

Welsford, D.C. and G.B. Nowara. 2007. Summary of holdings of Patagonian toothfish (*Dissostichus eleginoides*) otoliths and size-at-age estimates from Heard and McDonald Islands (Division 58.5.2). Document *WG-FSA-07/45*. CCAMLR, Hobart, Australia.

Welsford, D.C., A.J. Constable, T. Lamb and T. Robertson. 2006a. Preliminary assessment of Patagonian toothfish, *Dissostichus eleginoides*, in the vicinity of Heard Island and McDonald Islands (Division 58.5.2), based on a survey in May–June 2006 using the Generalised Yield Model. Document *WG-FSA-06/45*. CCAMLR, Hobart, Australia.

Welsford, D.C., A.J. Constable and G.B. Nowara. 2006b. Review of the use of survey data and length-at-age models in the assessment of *Dissostichus eleginoides* in the vicinity of Heard Island and McDonald Islands (Division 58.5.2). Document *WG-FSA-06/44*. CCAMLR, Hobart, Australia.

Welsford, D.C., T. Lamb and G.B. Nowara. 2007. Overview and update of Australia's scientific tagging program in the Patagonian toothfish (*Dissostichus eleginoides*) fishery in the vicinity of Heard and McDonald Islands (Division 58.5.2). Document *WG-FSA-07/48.* CCAMLR, Hobart, Australia.



Area of origin	Recaptures		
Ground B	1 377	2	4
Ground C	3	264	8
Ground D	5	3	18

Figure 1: Schematic showing the general relationships between the survey area (between the 200 and 1 000 m contours) and the fishing grounds. The approximate areas of these grounds are given, along with the number of tags released in the fishing grounds which were not recovered within 60 days or within the same population/fishery model period. The number of recaptures in each fishing ground (columns under bubbles) is given where recaptures have been separated into the grounds of origin.

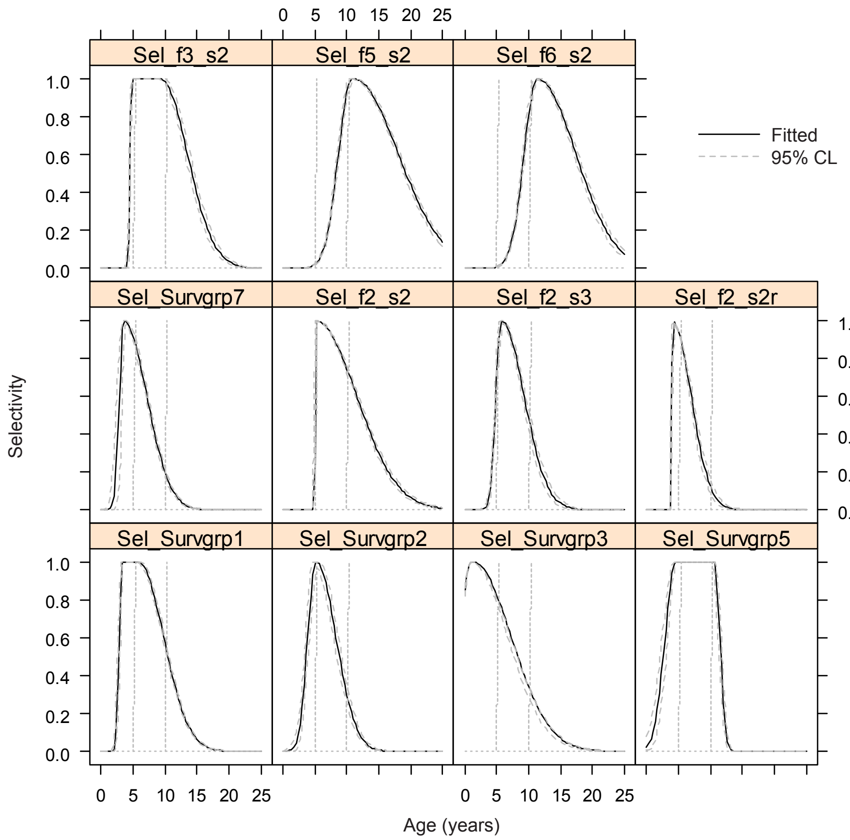


Figure 2a: Model *a2-tags-calpe* fitted double-normal plateau (DNP) and double-normal (DN) fishing selectivity curves showing 95% confidence bounds obtained from the multivariate normal (MVN) sample. Panel headings: Survgrp1 (survey years 2001, 2002 and 2004 to 2007), Survgrp2 (survey year 1999), Survgrp3 (survey year 1990), Survgrp5 (survey year 1993), Survgrp7 (survey year 2003), f2_s2, f2_s3 (Trawl Ground B, seasons 1 and 2, season 3), f2_s2r (Trawl Ground B, 2006, 2007 all seasons), f3_s2 (Trawl Ground C, all seasons), f5_s2 (Longline Ground C, season 2), f6_s2 (Longline Ground D, season 2). Reference lines are shown at ages 5 and 10.

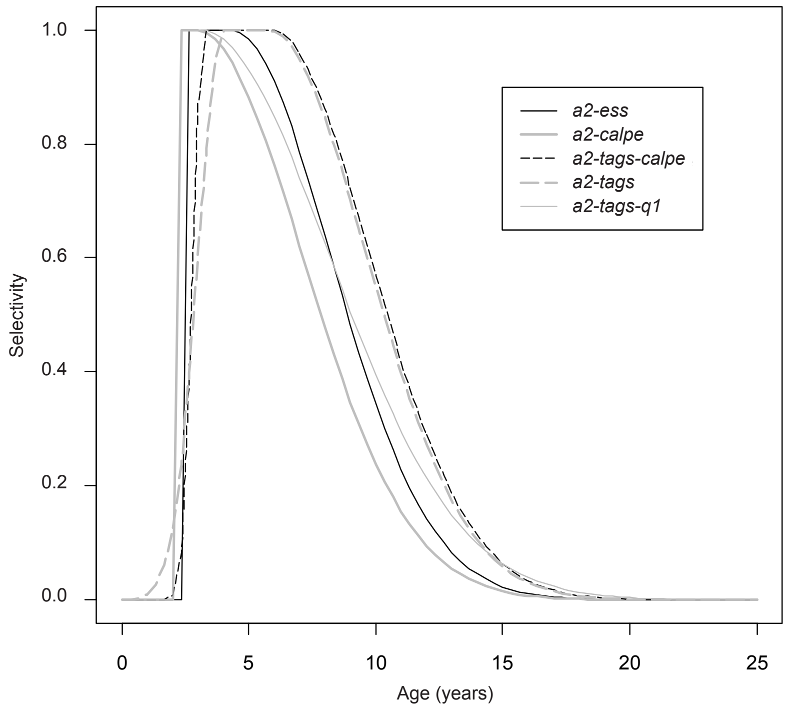


Figure 2b: Fitted double-normal plateau (DNP) fishing selectivity curves for Survgrp1 (survey years 2001, 2002 and 2004 to 2007) for each fitted population/fishery model.

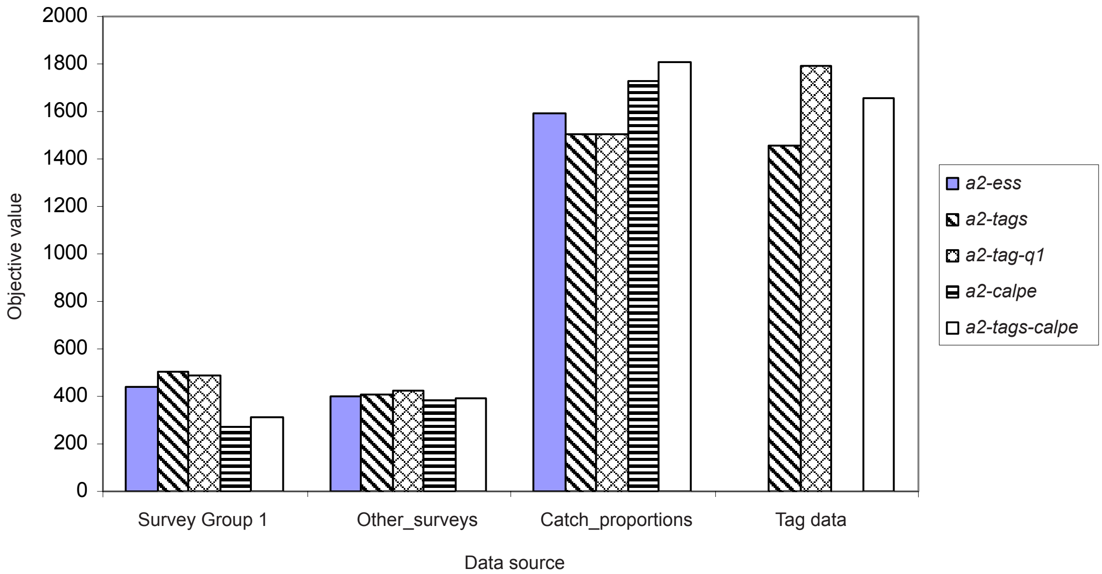


Figure 3: Contribution to the objective function minimised by CASAL for different relative weighting of datasets or allowing catchability q to be estimated for Survey Group 1 (cf. *a2-tags* versus *a2-tags-q1*). The contributions of the CPUE data were too small to represent on this plot.

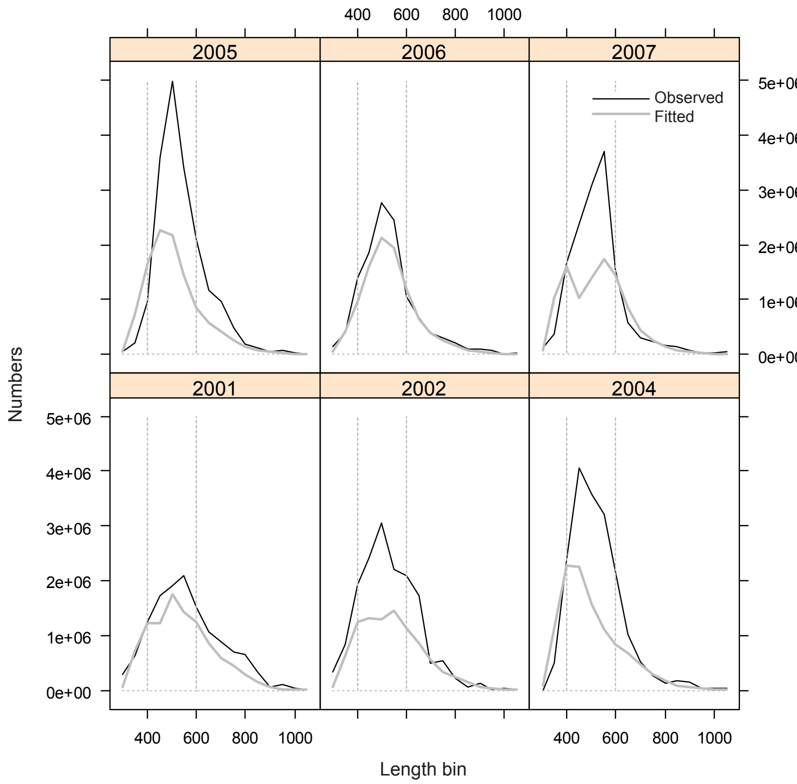


Figure 4a: Model *a2-tags-calpe* fits to Survey Group 1 abundance data with fitted values shown in a row and column trellis graph with reference lines at 400 and 600 mm.

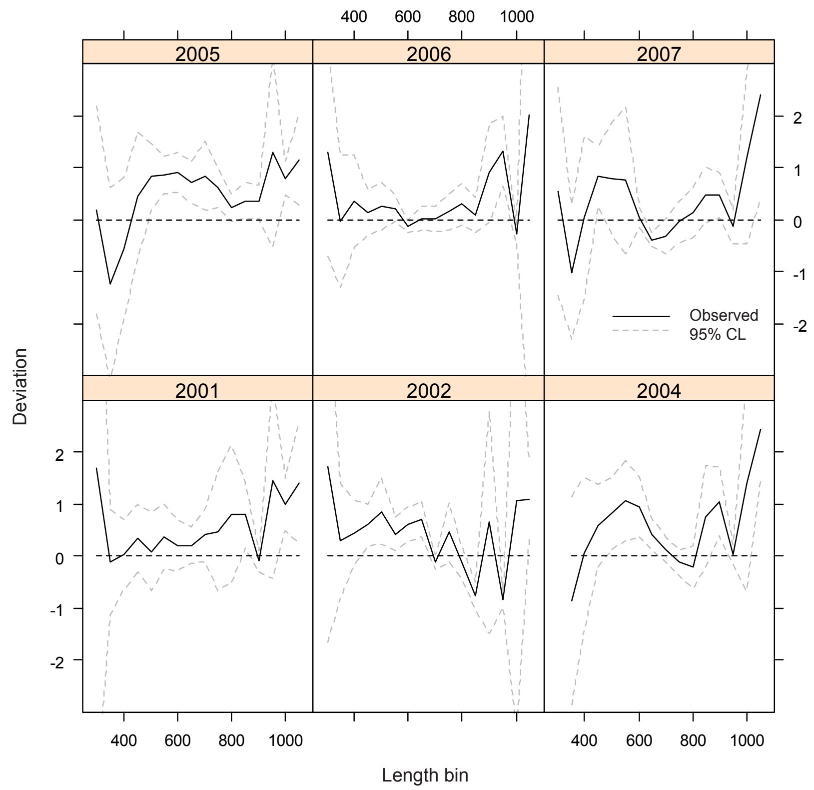


Figure 4b: Observed minus fitted log abundance (deviation) for Survey Group 1 data with model *a2-tags-calpe* fitted values and 95% confidence bounds.

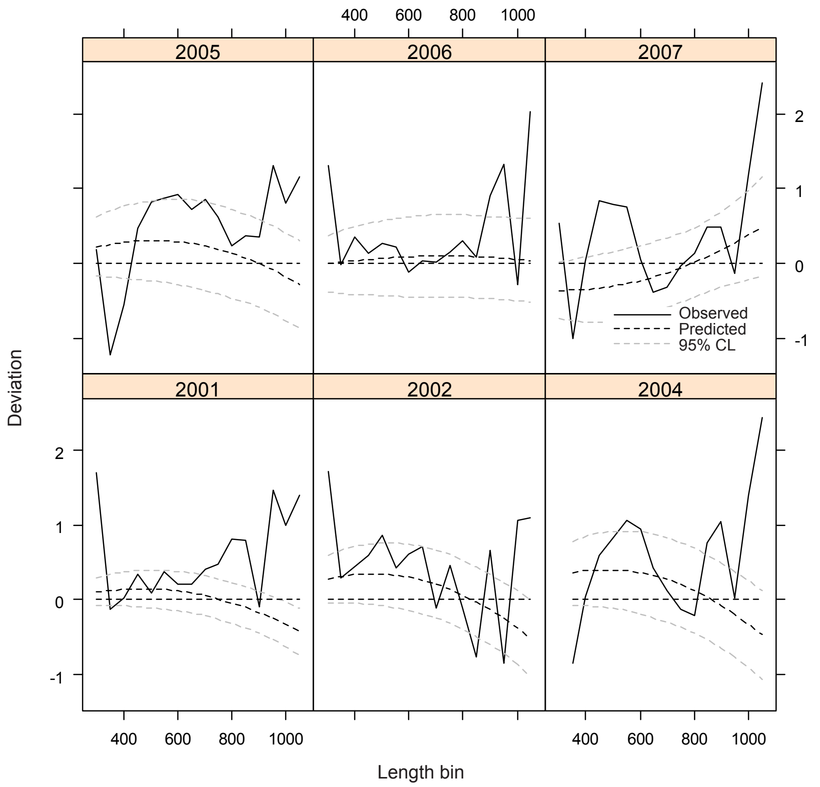


Figure 4c: Observed minus fitted log abundance (i.e. deviation) for Survey Group 1 data with model *a2-tags-calpe* fitted values showing fitted systematic lack-of-fit (SLOF) smooth trends and their 95% confidence bounds.

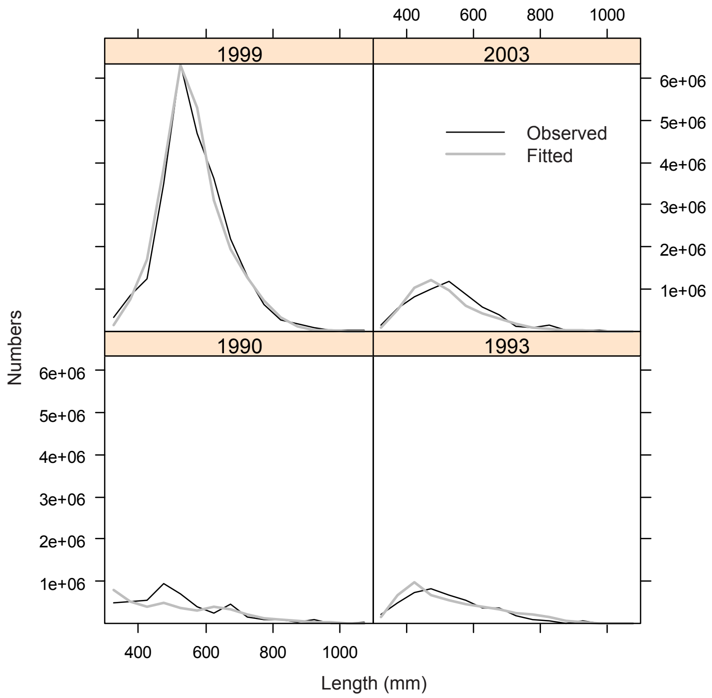


Figure 5: Model *a2-tags-calpe* fits to Survey Groups 3, 5, 2 and 7 data (see Table 2) – comparison of observed (black line) and expected (grey line) numbers-at-length for Survey Groups 3 (1990), 5 (1993), 2 (1999) and 7 (2003).

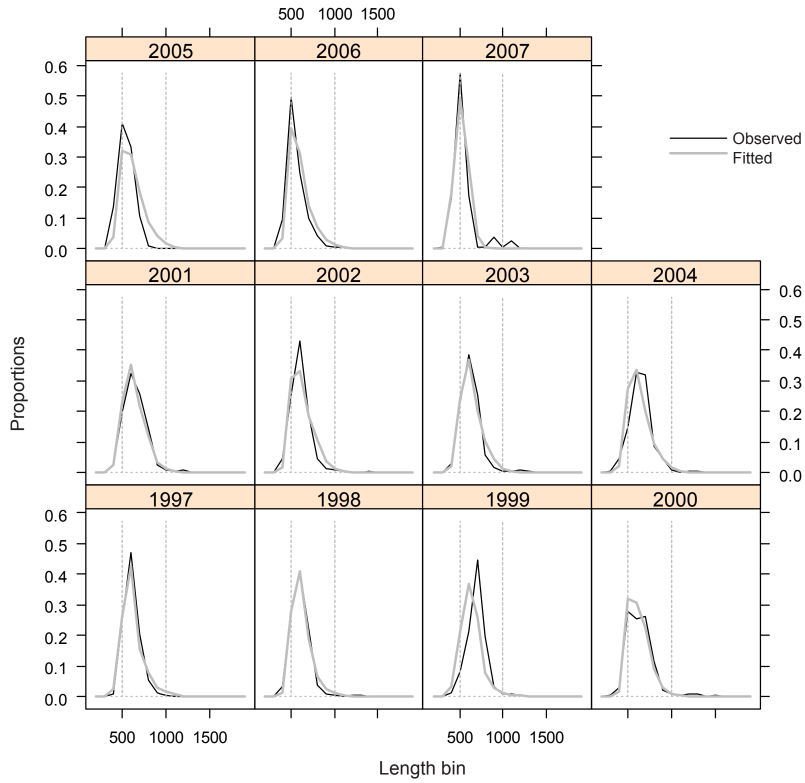


Figure 6a: Model *a2-tags-calpe* fits to catch-at-length proportions for Trawl Ground B, season 2 (fishery f2_s2) with reference lines at 500 and 1 000 mm.

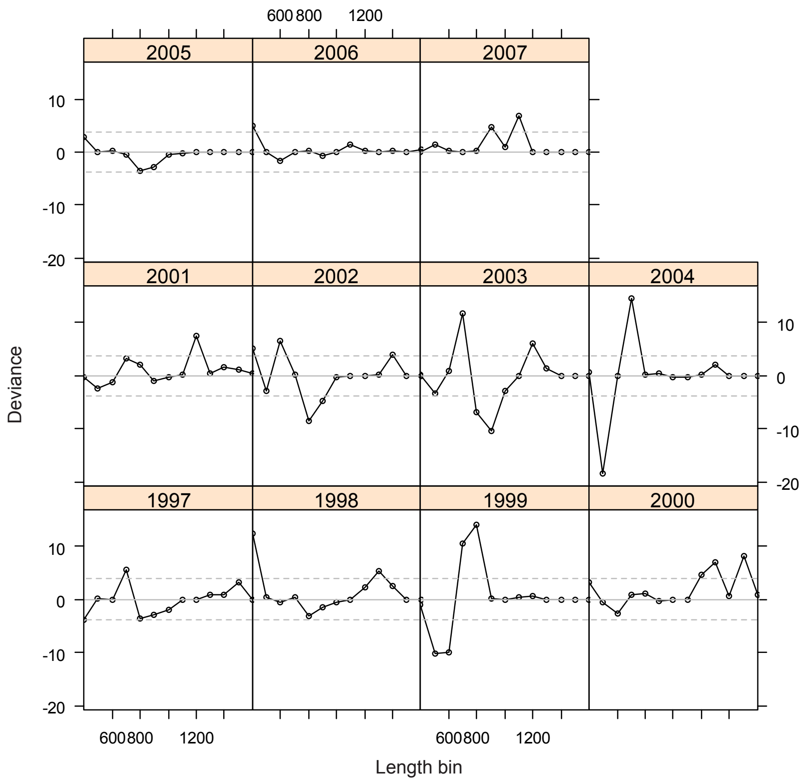


Figure 6b: Deviance residuals from systematic lack-of-fit (SLOF) model with predictions from model *a2-tags-calpe* for catch-at-length proportions for Trawl Ground B, season 2 (fishery f2_s2) with approximate 95% confidence bounds shown by the dashed lines.

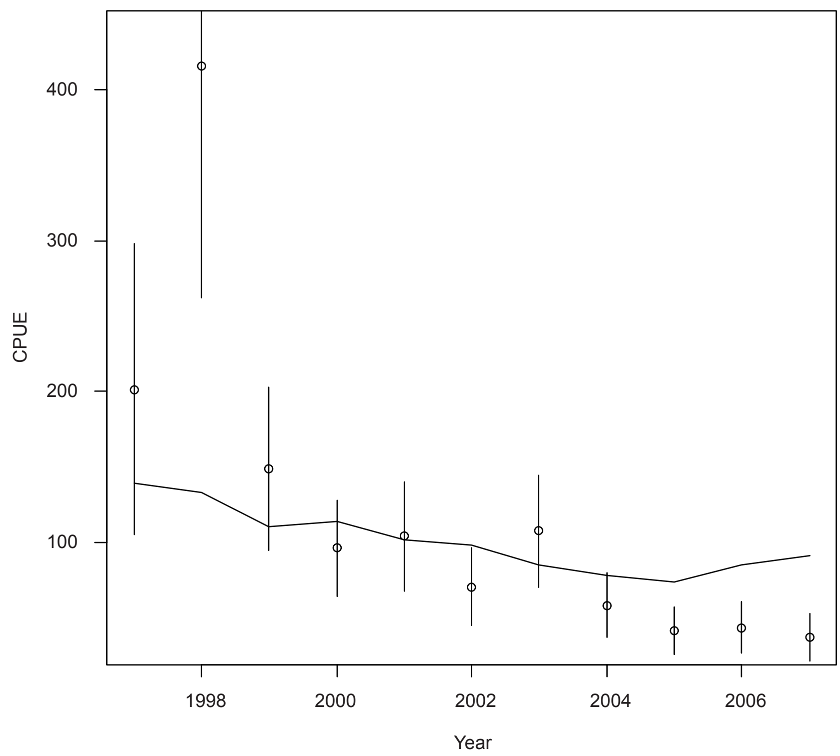


Figure 7: Estimated CPUE series from the generalised linear mixed model (GLMM) for Trawl Ground B (f2) (circles) with bars corresponding to ±1 standard error of the estimate and the model *a2-tags-calpe* fitted series (line).

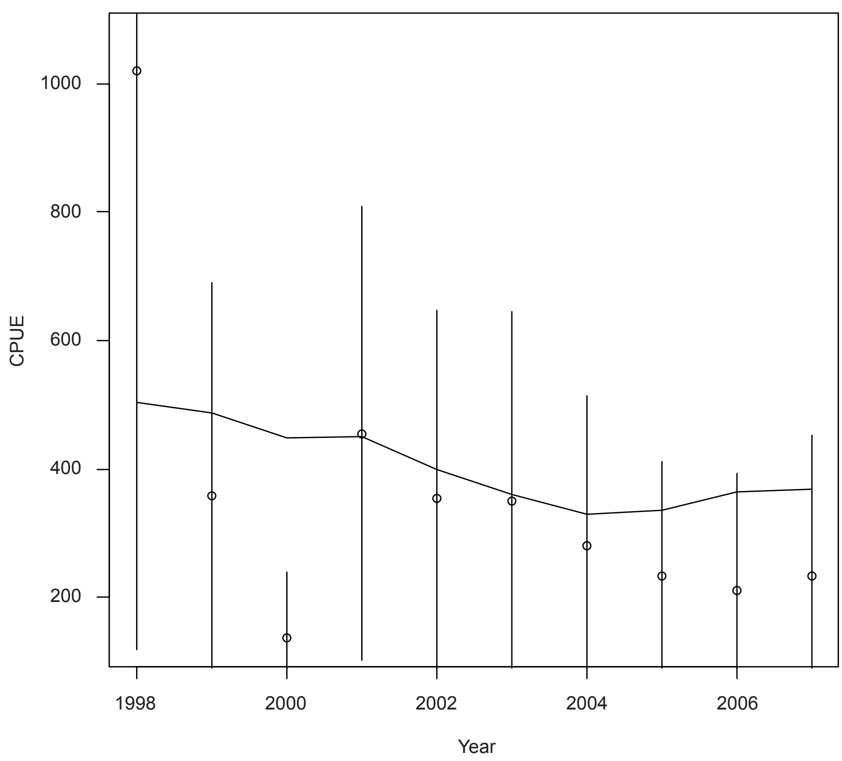


Figure 8: Estimated CPUE series from the generalised linear mixed model (GLMM) for Trawl Ground C (f3) (circles) with bars corresponding to ±1 standard error of the estimate and the model *a2-tags-calpe* fitted series (line).

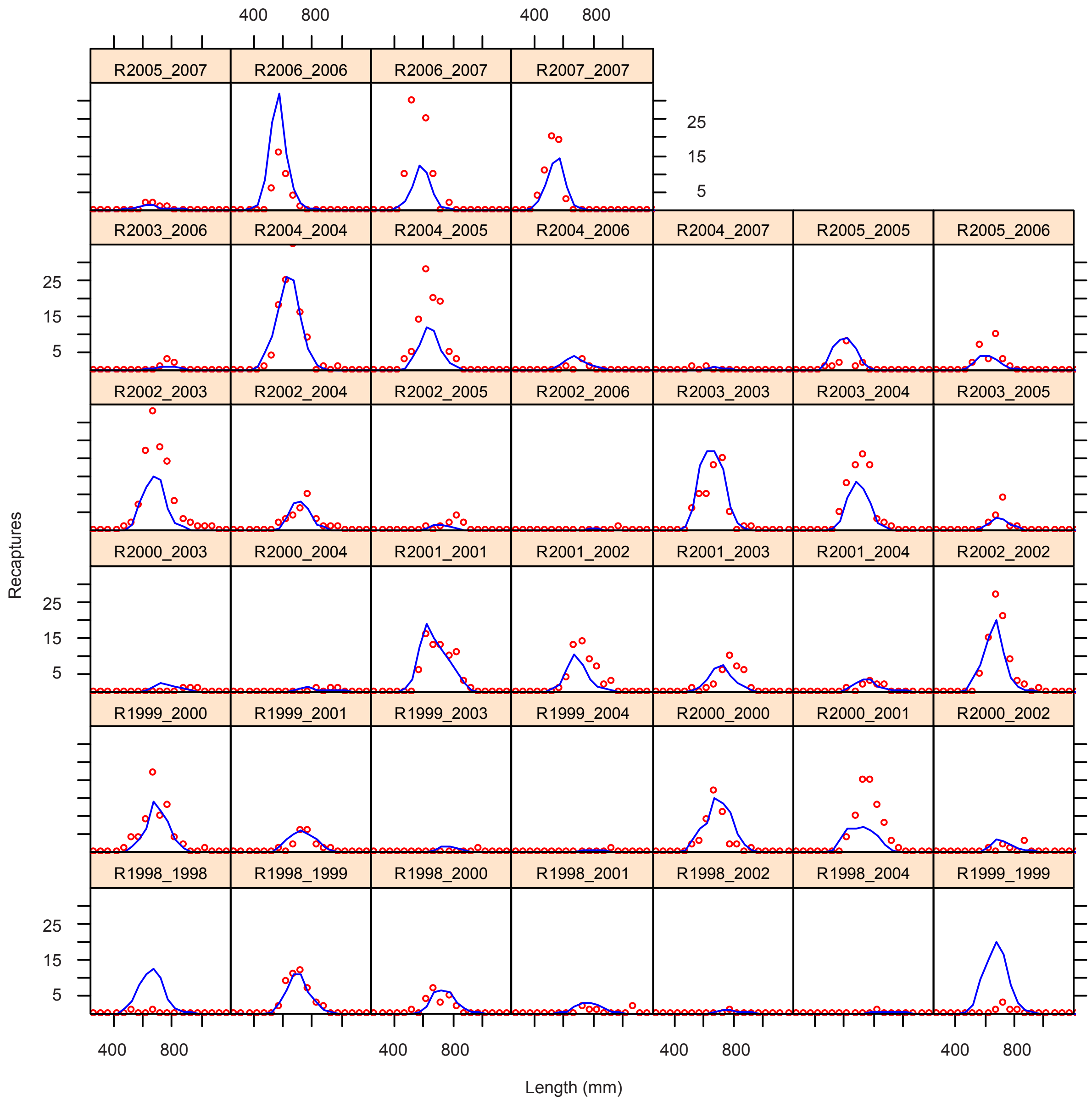


Figure 9: Model *a2-tags-calpe* (see Table 3) – comparison of observed and predicted number of recaptures by 50 mm length bin for releases and recaptures in Trawl Ground B (f2) where predictions were obtained from the fit of population/fishery model *a2-tags-calpe*. For the panel titles, the year directly after 'R' is the release while the following labelled year is the year of recaptures in f2. Within a row of panels, the consecutive recapture years (only present if there was at least one recapture) are read from left to right.

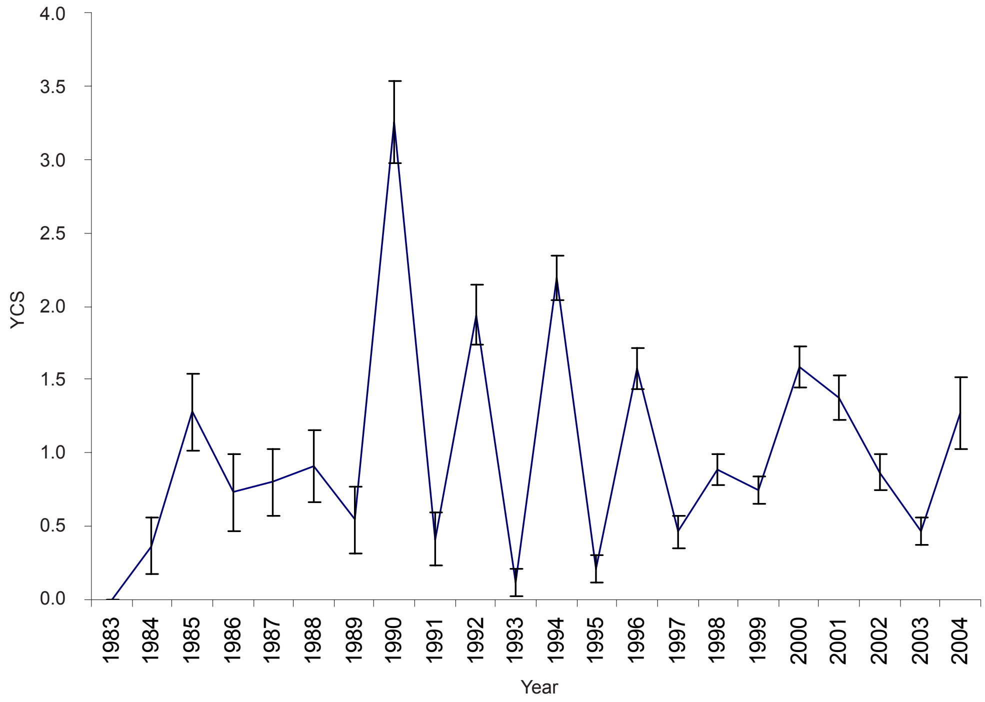


Figure 10: Year-class strength (YCS) estimates (±SE) for *a2-tags-calpe* model.

Liste des tableaux

Liste des figures

normal multivarié (MVN). En-têtes : Survgrp1 (années de campagnes 2001, 2002 et 2004 à 2007), Survgrp2 (année de campagne 1999), Survgrp3 (année de campagne 1990), Survgrp5 (année de campagne 1993), Survgrp7 (année de campagne 2003), f2_s2, f2_s3 (lieu de chalutage B, saisons 1 et 2 et saison 3), f2_s2r (lieu de chalutage B 2006, 2007 toutes saisons), f3_s2 (lieu de chalutage C, toutes saisons), f5_s2 (lieu de pêche à la palangre C, saison 2), f6_s2 (lieu de pêche à la palangre D, saison 2). Des traits de référence sont indiqués aux âges 5 et 10.

Figure 2b: Courbes ajustées normales doubles avec plateau (DNP) de sélectivité de la pêche pour le Survgrp1 (années de campagnes 2001, 2002 et 2004 à 2007) pour chaque modèle de population/pêcherie ajusté.

Figure 3: Contribution à la fonction objective réduite par CASAL pour différentes pondérations relatives des jeux de données, permettant d'estimer la capturabilité *q* pour le groupe de campagnes d'évaluation 1 (cf. *a2-tags* vs *a2-tags-q1*). La contribution des données de CPUE était trop faible pour pouvoir être représentée sur ce schéma.

Figure 4a: Ajustements du modèle *a2-tags-calpe* aux données d'abondance du groupe de campagnes d'évaluation 1 avec les valeurs ajustées données dans un graphique de treillis avec rangées et colonnes et des lignes de référence à 400 et 600 mm.

Figure 4b: Abondance logarithmique observée des données du groupe de campagnes d'évaluation 1 moins abondance logarithmique ajustée (déviation) de ces données avec valeurs ajustées par le modèle *a2-tags-calpe* et intervalle de confiance à 95%.

Figure 4c: Abondance logarithmique observée des données du groupe de campagnes d'évaluation 1 moins abondance logarithmique ajustée (déviation) de ces données avec valeurs ajustées par le modèle *a2-tags-calpe* montrant les tendances lissées ajustées du défaut systématique d'ajustement (SLOF) et l'intervalle de confiance à 95%.

Figure 5: Ajustements du modèle *a2-tags-calpe* aux données des groupes de campagnes 3, 5, 2 et 7 (voir tableau 2) – comparaison du nombre par longueur observé (trait noir) et prévu (trait gris) pour les groupes de campagnes 3 (1990), 5 (1993), 2 (1999) et 7 (2003).

Figure 6a: Ajustements du modèle *a2-tags-calpe* aux proportions de la capture selon la longueur pour le lieu de chalutage B, saison 2 (pêcherie f2_s2) avec lignes de référence à 500 et 1 000 mm.

Figure 6b: Résidus de la déviance du modèle de défaut systématique d'ajustement (SLOF) avec prédictions du modèle *a2-tags-calpe* pour les proportions de capture selon la longueur pour le lieu de chalutage B, la saison 2 (pêcherie f2_s2) avec limites approximatives de l'intervalle de confiance à 95% représentées par les lignes en tirets.

Figure 7: Série de CPUE estimée à partir du modèle mixte linéaire généralisé (GLMM) du lieu de chalutage B (f2) (cercles), les barres correspondant à ± 1 erreur standard de l'estimation et série ajustée (ligne) du modèle *a2-tags-calpe*.

Figure 8: Série de CPUE estimée à partir du modèle mixte linéaire généralisé (GLMM) du lieu de chalutage C (f3) (cercles), les barres correspondant à ± 1 erreur standard de l'estimation et série ajustée (ligne) du modèle *a2-tags-calpe*.

Figure 9: Modèle *a2-tags-calpe* (voir tableau 3) – comparaison entre le nombre observé et le nombre prévu de recaptures par lots de 50 mm de longueur pour le marquage et la recapture dans le lieu de chalutage B (f2) lorsque les prévisions proviennent de l'ajustement du modèle *a2-tags-calpe* de population/pêcherie. Dans les en-têtes, l'année suivant la lettre 'R' est celle du marquage alors que la suivante est celle de la recapture en f2. Dans une rangée de cases, les années consécutives de recapture (données uniquement lorsqu'il y a eu au moins une recapture) se lisent de gauche à droite.

Figure 10: Estimations (±SE) d'abondance des classes d'âges (YCS) pour le modèle *a2-tags-calpe*.

Список таблиц

Табл. 1: Общее изъятие (т) на Участке 58.5.2 по промыслам и сезонам. В оценку не включены небольшие уловы, полученные вне пределов основных участков. В результате общее изъятие здесь может отличаться от общего изъятия на Участке 58.5.2.

Список рисунков

Lista de las tablas

Lista de las figuras

Figura 3: Contribución a la función objetivo minimizada con CASAL para las distintas ponderaciones relativas de los conjuntos de datos, permitiendo la estimación de la capturabilidad *q* para el Grupo de estudio 1 (cf. *a2-tags* versus *a*2-*tags-q1*). Las contribuciones de los datos de la CPUE fueron demasiado pequeñas como para ser representadas en este gráfico.

Figura 4a: Ajustes del modelo *a2-tags-calpe* a los datos de abundancia del Grupo de estudio 1; los valores ajustados se muestran en un gráfico trellis con paneles dispuestos en filas y columnas, con líneas de referencia en 400 y 600 mm.

Figura 4b: Abundancia observada menos log abundancia ajustada (desviación) para los datos del Grupo de estudio 1 con los valores ajustados del modelo *a2-tags-calpe* e intervalos de confianza del 95%.

Figura 4c: Abundancia observada menos log abundancia ajustada (desviación) para los datos del Grupo de estudio 1 con los valores ajustados del modelo *a2-tags-calpe* mostrando tendencias suavizadas de las fallas sistemáticas del ajuste (SLOF) con su intervalo de confianza del 95%.

Figura 5: Ajustes del modelo *a2-tags-calpe* a los datos de los Grupos de estudio 3, 5, 2 y 7 (ver tabla 2) – comparación del número por talla observado (línea negra) y esperado (línea gris) para los Grupos de estudio 3 (1990), 5 (1993), 2 (1999) y 7 (2003).

Figura 6a: Ajustes del modelo *a2-tags-calpe* a las proporciones de tallas de la captura para el Caladero B de la pesca de arrastre, temporada 2 (pesquería f2_s2), con líneas de referencia en 500 y 1 000 mm.

Figura 6b: Residuales de la desviación de la simulación de la falla sistemática del ajuste (SLOF) con predicciones del modelo *a2-tags-calpe* para la proporción de tallas de la captura en el Caladero B de la pesca de arrastre, en la temporada 2 (pesquería f2_s2) con intervalos de confianza aproximados de 95%.

Figura 7: Serie estimada de la CPUE con el modelo lineal mixto generalizado (GLMM) para el Caladero B de la pesca de arrastre (f2) (círculos) con barras de ±1 SE de la estimación, y serie ajustada del modelo *a2-tags-calpe* (curva).

Figura 8: Serie estimada de la CPUE con el modelo lineal mixto generalizado (GLMM) para el Caladero C de la pesca de arrastre (f3) (círculos) con barras de ±1 SE de la estimación, y serie ajustada del modelo *a2-tags-calpe* (curva).

Figura 9: Modelo *a2-tags-calpe* (ver tabla 3) – comparación del número de recapturas observadas y pronosticadas, por intervalo de talla de 50 mm, para los ejemplares liberados y recapturados en el Caladero B de la pesca de arrastre (f2), donde las predicciones fueron obtenidas del ajuste del modelo de población/pesquería *a2-tags-calpe*. En las etiquetas de los paneles, el año después de 'R' corresponde al año de liberación mientras que el año siguiente es el año de recaptura en f2. Dentro de una fila de paneles, los años consecutivos de recaptura (sólo si hubo por lo menos una recaptura) se leen de izquierda a derecha.

Figura 10: Estimaciones de la abundancia de la clase anual (YCS) (±SE) para el modelo *a2-tags-calpe*.

APPENDIX 1

Estimating Parameters Used with Tagging Data

Tag shedding rate and tag detection probability

From the start of the mark–recapture program in 1997, all fish tagged have been tagged with two dart tags always placed together, generally behind the dorsal fin. In addition for fish larger than 400 mm in total length a TIRIS© (Texas Instruments) electronic tag (i.e. pit tag) was also inserted. Tag shedding rates for the electronic tags can reasonably be assumed to be zero while the dart tag shedding rate for a single tag was estimated from the number of fish recaptured with a single tag compared to the number with both tags intact for the given number of days at liberty using the method of Kirkwood and Walker (1984). This method assumed a constant instantaneous tag shedding rate so that after D days at liberty expressed as a fraction of the year (i.e. days/365), assuming that the tag shedding rate was not elevated straight after release and the tags were shed independently, the probability, $P_b(D)$, that both tags have been shed is $P_b(D) = P_r(D)P_l(D)$, the product of the probability that the right and left tags have each been shed where $P_r(D) = P_l(D) = 1 - \exp\{-\exp(\alpha_0 + \ln(D))\}$. The estimate of α_0 using maximum likelihood and the data for 2 411 double-tagged released fish that were observed with either one or two tags at recapture was –2.4526. Therefore, for 1, 2 and 3 years after release $P_b(D)$ is estimated to be 0.0068, 0.0250 and 0.0518 respectively. The assumption of a constant instantaneous tag shedding rate can be relaxed by allowing the slope of the term ln(D) to vary from 1 to give $P_r(D) = P_l(D) = 1 - \exp\{-\exp(\alpha_0 + \alpha_1 \ln(D) + \ln(D))\}$. The likelihood ratio test suggests rejection of the hypothesis $H_0 : \alpha_1 \equiv 0$ ($X^2 = 90, \chi_1^2$, $P < 0.001$) with $\hat{\alpha}_1 = -0.626$ and $\hat{\alpha}_0 = -2.5219$ so that $P_b(D)$ for 1, 2 and 3 years is then given by 0.00596, 0.00977 and 0.01301 respectively. CASAL allows a tag shedding rate for corresponding proportions of the year defined by the seasons to be specified so that the number of tagged fish for a tagging partition is given by $n'_{ij} = n_{ij} \exp\left(-t_j l_i\right)$ where n_{ij} is the number of tagged fish in tag partition i at the start of period j within the current year, n'_{ij} is the number at the end of the period, t_j is the proportion of the annual tag shedding rate occurring for the period and l_j is the annual tag shedding rate. Note that since the t_j sum to 1 for each year, if only a single dart tag was used in tagging and the Kirkwood and Walker (1984) model is used (i.e. $\alpha_1 \equiv 0$) to predict single tag shedding, then $n'_{ij} = n_{ij}\left\{1 - P_r\left(t_j\right)\right\}$ which gives $l_i = \exp(\alpha_0)$. However, this simple equivalence between CASAL's method of accounting for tag shedding and the Kirkwood and Walker (1984) model does not hold for double-tagged fish since n'_{ij} cannot be made equivalent to $n_{ij}\{1 - P_b(t_j)\}$ by equating l_i and $\exp(\alpha_0)$.

Due to both the inability to deal with double-tagging via the tag shedding rate in CASAL and the fact that the effect of tag shedding on the estimated total number of fish that retain at least one tag and those that have lost both dart tags for a given fishery and recapture year depends on the type of fishing method used to obtain the recapture, it was necessary to incorporate the tag shedding rate as part of the detection probability in the recapture data. This is because TIRIS detectors were not available on longline vessels, but the visual detection rate is relatively high compared to trawl-caught fish (it is assumed that only one dart tag is required for 100% visual detection of longline-caught tagged fish), whereas the TIRIS detectors were installed on the trawl vessels. However, for the following reasons the TIRIS detector does not always deliver a 100% detection rate. Firstly, the TIRIS detector may not be operational for every haul, and secondly the TIRIS tag can fail to be charged by the TIRIS detector if shielded by stainless steel lips on the conveyor chute. CASAL allows detection probability to vary by recapture fishery but not by year of recapture. Note that it was assumed that the visual detection rate is the same when both dart tags are retained as that for a single dart tag retained.

For the tag shedding rate defined for releases it was assumed that the rate was zero for all years of release. For longline-caught tagged fish it was assumed that the detection probability is less that 1 only if some fish have shed both dart tags. The detection probability was set at $1 - P_b(1) = 1 - 0.0068 = 0.9932$, which assumes the proportion of tagged fish shedding both dart tags is constant irrespective of time at liberty which is assumed to be 1 year. This assumption is necessary in order to incorporate tag shedding via the 'detection probability' option in CASAL, however, since the average number of days at liberty for recaptured fish at first recapture, where this period is greater than 60 days, was 373 days, this approximation should not introduce serious bias into estimation of the model parameters. If the tag shedding model, including estimates of both α_0 and α_1, is used, the detection rate changed only slightly to 0.994. For trawl-caught tagged fish an average detection probability across all hauls within the fishery for which there was at least one recapture was calculated. To calculate this average it was necessary to assume that all tagged fish have been given all three tags (i.e. the electronic and two dart tags) but, since there were relatively few fish tagged that did not receive the electronic tag because they were under 400 mm in length, this assumption should not significantly affect estimation. The haul-level detection probability was determined as the probability of the overall trawl-specific visual detection probability, P_{vT}, if the TIRIS detector was not operational during the processing of the haul, and a combined visual and electronic detection probability, P_{veT}, when the TIRIS detector was operational for the haul. If, for the given cruise, the haul occurred in the total number (i.e. totalled across length bins and hauls) of recaptures that were detected only by the TIRIS (electronic) detector, only by visual detection of a dart tag, or detected by both methods are given by n_e, n_v and n_{ve} respectively, then the equivalent estimate of the actual number of recaptured fish, both detected and undetected, described by Tuck et al. (2003), is given by $\hat{n}_r = (n_{ve} + n_e)(n_{ve} + n_v)/n_{ve}$ so that $P_{veT} = n_t / \hat{n}_r$ where $n_t = (n_{ve} + n_e + n_v)$ for $n_{ve} > 0$ and $P_{veT} = 1$ for $n_{ve} = 0$. The cruise-level estimate of P_{vT} was obtained as $P_{vT} = 1 - n_e/n_a$ where n_a is the total number of detected recaptures that were available to the TIRIS detector (i.e. the detector was operational for the haul in which they were recaptured) for the cruise and these values were averaged over all cruises for cruises for which $n_e < n_a$ to give $\overline{P}_{vT}$. It was necessary to average cruise-level P_{vT} because of cases where $n_e = n_a$ so that no information is available for that cruise on P_{vT}. The value of $\overline{P}_{vT}$ obtained was 0.969 and note that $\overline{P}_{vT}$ accounts for missed detections due to both crew or observers not seeing dart tags when at least one was present and missed visual identification of a tagged fish because it has shed both dart tags. This can be seen by noting that $n_e/n_a = \{n_a - (n_{ve} + n_v - [n_t - n_a])\}/n_a$ where $[n_t - n_a]$ is an adjustment to n_v to remove visual detections that occurred when the detector was not operational so that the proportion of TIRIS-only detections of all detections when TIRIS was operational corresponds to the visual non-detection proportion for both of the above reasons. The overall detection probability for trawl catches is then given by $P_T = P_{veT}\delta + (1-\delta)\overline{P}_{vT}$ where δ is an indicator variable that takes the value 1 if the TIRIS detector was operational while the haul catch was processed and zero otherwise. For hauls within a cruise the only variability in P_T is that due to variability in δ. For each fishery the mean value of P_T was obtained and used as the 'detection_probability' for the recapture fishery. These values are given in the text.

Overdispersion

CASAL allows an externally estimated dispersion parameter to be defined for the binomial likelihood for the recapture numbers by length bin conditional on the number of scanned fish for the corresponding length bins for each recapture fishery. A single dispersion parameter across all length bins and recapture years is allowed. If the dispersion parameter is given by ϕ, then the variance for number of recaptures is $Var\left(p_{ij} \middle| \hat{p}_{ij}, N_{ij}\right) = \frac{\phi}{N_{ij}} \hat{p}_{ij}\left(1 - \hat{p}_{ij}\right)$ where N_{ij} is the number of scanned fish in the length bin and p_{ij} and $\hat{p}_{ij}$ are the observed and predicted proportions of N_{ij} that are recaptures respectively for haul i and length bin j. To estimate ϕ prior to the fit of the model using CASAL, the total numbers of recaptures and scanned fish for each haul (i.e. summed across length bins) were obtained as $\sum_j N_{ij} p_{ij}$ and $N_i' = \sum_j N_{ij}$ respectively. The number of recaptures was scaled to allow incorporation of between-haul heterogeneity in detection probability into the estimate of ϕ so that the scaled number of recaptures was given by $n_i' = \sum_j N_{ij} p_{ij} / P_i$ where P_i is $P_{T,i}$ for hauls from trawlers and is constant and equal to 0.993 for longliners. The n_i' were then fitted as a binomial generalised linear model conditional on the N_i' with systematic terms of main effects of fishery and season and the interaction of fishery and season. The estimate of ϕ was obtained as the residual deviance divided by its degrees of freedom. The residual deviance, degrees of freedom and estimate of ϕ obtained were 6 360, 5 626 and 1.13 respectively. This single estimate of the dispersion parameter was used for all recapture fisheries.

APPENDIX 2

Calculating Process Error Variance for Lognormal Data taking into account CVs of Annual Estimates

As an example of how process error variance was estimated in the iterative CASAL/process error estimation procedure consider the CPUE series estimates. If the series estimate for year y (input observation in CASAL) is given by $\tilde{c}_y$ with corresponding estimated CV of $\tilde{\sigma}_y$, and the predicted value from an iteration of the CASAL fit is given by $\hat{\tilde{c}}_y$, then a simple linear SLOF model involves fitting a simple linear regression of $\ln\left(\tilde{c}_y\right)-\ln\left(\hat{\tilde{c}}_y\right)$ on year number with weights given by $1/\tilde{\sigma}_y^2$. The CV for the process error (i.e. the standard deviation on the log scale) is estimated as

$$\hat{\sigma}_{PE}=\left[\left(\omega^2-1\right)\sum_{y=1}^{n}\tilde{\sigma}_y^2/n\right]^{\frac{1}{2}}$$

where ω^2 is the residual mean square from the fit of the regression. If ω^2 was found to be less than 1, then $\hat{\sigma}_{PE}$ was set to zero. To validate the result derived above for calculating process error CV using the estimated CPUE series for Trawl Ground B, the predictions for this series obtained from a population/fishery model fit were used to compare the above weighted regression method with a mixed model analysis whereby the same linear regression was fitted in each case. For the mixed model, the $\tilde{\sigma}_y^2$ were pre-specified as known 'units-level' variances and years were included as random effects (i.e. as a factor) as well as a fixed linear term. The year-level random effect variance estimate (i.e. residual maximum likelihood (REML) estimate), should be similar to $\hat{\sigma}_{PE}^2$. To avoid values of ω^2 less than 1, the actual values of $\tilde{\sigma}_y^2$ were scaled by dividing by either 3, 5 or 50. The corresponding estimates of $\hat{\sigma}_{PE}^2$ for the weighted regression versus REML estimate for scale factor of 3 were 0.077 and 0.074 respectively. For scale factors of 5 and 50 the corresponding values were 0.112 versus 0.106 and 0.126 versus 0.119 respectively. When the scale factor was set to 1, the weighted regression gave a negative value for $\hat{\sigma}_{PE}^2$ while the REML estimate was very close to zero. These calculations were made using the GenStat software package (Lawes Agricultural Trust, 2002).

The same approach was used for estimating process error for the single-year surveys and the multi-year survey taking into account the different SLOF models fitted in each case corresponding to the data being in the form of length-binned estimates (i.e. also by years for the multi-year survey).