INTEGRATION OF CPUE DATA INTO ASSESSMENTS USING THE GENERALISED YIELD MODEL

G.P. Kirkwood
Renewable Resources Assessment Group
Department of Environmental Science and Technology
Imperial College, RSM Building
Prince Consort Road
London SW7 2BP, United Kingdom
Email – g.kirkwood@ic.ac.uk

A.J. Constable
Australian Antarctic Division
Channel Highway, Kingston 7050
Tasmania, Australia

Abstract

For the last three years, CCAMLR’s Working Group on Fish Stock Assessment (WG-FSA) has accorded high priority to the development of methods of integrating different indicators of stock status into assessments using the Generalised Yield Model (GYM). In this paper, we propose a method, based on the use of importance sampling, for incorporating information on trends in standardised CPUEs into GYM assessments. The use of the method is illustrated using data for Patagonian toothfish (Dissostichus eleginoides) in Subarea 48.3. The method requires only very small adjustments to the computer program implementing the GYM assessments.

Résumé

Ces trois dernières années, le groupe de travail de la CCAMLR chargé de l’évaluation des stocks de poissons (WG-FSA) a accordé une haute priorité à la mise en place de méthodes d’insertion de différents indicateurs de l’état du stock dans les évaluations effectuées au moyen du Modèle de rendement généralisé (GYM). Dans ce document, nous proposons une méthode fondée sur l’utilisation de l’échantillonnage de l’importance pour intégrer les informations sur les tendances des CPUE normalisées dans les évaluations effectuées par le GYM. L’utilisation de cette méthode est illustrée au moyen des données sur la légine australe (Dissostichus eleginoides) dans la sous-zone 48.3. La méthode ne nécessite que des ajustements minimes du logiciel permettant de réaliser les évaluations par le GYM.

Resumen

Durante los últimos tres años, el grupo de trabajo sobre la evaluación de las poblaciones de peces (WG-FSA) de la CCRVMA ha dado prioridad urgente al desarrollo de métodos para la integración de varios indicadores del estado de las poblaciones de peces a las evaluaciones que utilizan el modelo general de rendimiento (GYM). Este estudio propone un método basado en el muestreo de importancia para incorporar la información sobre las tendencias de los datos normalizados de CPUE a las evaluaciones realizadas con el modelo general de rendimiento (GYM). Se aplicó el método a los datos pertinentes al bacalao de profundidad (Dissostichus eleginoides) en la Subárea 48.3 para ilustrar su utilización. La aplicación del método requiere solamente de pequeños ajustes al programa de informática utilizado en las evaluaciones llevadas a cabo con el modelo GYM.

Резюме

На протяжении последних трех лет Рабочая группа АНТКОМа по оценке рыбных запасов (WG-FSA) придает большое значение разработке методов включения в оценки различных индикаторов состояния запаса, используя общепринятую модель вылова (GY-модель). В данной статье предлагается метод включения информации о тенденциях стандартизованного CPUE в оценки по GY-модели. Метод основан на выборке по важности, и в качестве иллюстрации его работы
INTRODUCTION

For several years CCAMLR’s Working Group on Fish Stock Assessment (WG-FSA) has had available to it two indicators of stock status for Patagonian toothfish (Dissostichus eleginoides) in Subarea 48.3, but it is not yet possible to make direct assessments of the abundance of fish in the population (SC-CAMLR, 1995). The first of these indicators is a time series of estimates of biomass arising from use of the Generalised Yield Model (GYM) and driven by estimates of recruitment strength derived from trawl surveys (Constable and de la Mare, 1996). The second is a time series of longline CPUEs, standardised using Generalised Linear Models (GLMs). In 1997 and 1998 WG-FSA observed that the standardised CPUE series had been declining in the most recent years, and that the trend in biomass predicted by the GYM had shown a smaller decline than that indicated by the standardised CPUEs. On these grounds, it recommended that the catch limits should be less than those calculated using the GYM, but it was unable to advise on the extent of the reduction that would be appropriate (SC-CAMLR, 1997, paragraph 4.169; SC-CAMLR, 1998, paragraph 4.117). In 1999, WG-FSA noted that the standardised CPUE had increased since the 1997/98 season, and it recommended that the catch limit should be equal to that indicated by the GYM analysis (SC-CAMLR, 1999, paragraphs 4.141 and 4.142).

Recognising the recurring difficulty in quantitatively reconciling standardised CPUE trends with those trends predicted by the GYM analyses, in each of the last three years WG-FSA has accorded high priority to the development of methods for integrating different indicators of stock status into assessments (e.g. SC-CAMLR, 1999, paragraph 4.145).

Assessments using the GYM

The GYM is a generalised version of the krill yield model used in CCAMLR (Constable and de la Mare, 1996). The evolution of these approaches is described by Constable et al. (2000). In summary, the GYM is a general age-structured fish stock projection model for assessing the long-term annual yield that satisfies objectives for the maintenance of the stock spawning biomass in accordance with CCAMLR criteria. These specify a bound on the probability that the spawning biomass may become depleted to below some specified level over a specified period and set a further constraint on the long-term status of the stock relative to the median pre-exploitation spawning biomass. The model provides a flexible method for assessing the influence of different patterns of growth, natural mortality, spawning and fishing on estimates of yield and yield per recruit. It can also be used to evaluate stochastic stock trajectories under a specified catch regime. The procedure numerically integrates a set of differential equations that incorporate functions that specify growth, mortality, age-dependent selectivity and seasonal patterns in fishing mortality derived from existing data.

In the assessment of toothfish, projections include the time series of survey estimates of recruitment of age 4 fish into the population (coupled with the uncertainty in these estimates) as well as the time series of historical catches. Recruitments in future years and in determining the age composition of the population at the beginning of a projection trial are drawn from a lognormal distribution estimated from the time series of recruitments. Uncertainties in the demographic parameters and in the estimates of recruitment are dealt with in the projections by drawing the estimates for an individual trial from the distributions provided to the model as input data. Consequently, each trial will have a different initial population size and will have different trajectories through the historical period as well as into the future.

The long-term annual yield is determined according to the status of the stock during projections into the future, the number of years for such projections being specified in the CCAMLR criteria. The abundance of the stock in a given year for each trial is provided in output tables. In the absence of other information, each trajectory is plausible given the data available and the specified uncertainties. Summaries of stock status over the historical catch period as well as into the future have been presented in CCAMLR as median abundance in each year with 95% confidence.
Integration of CPUE data into assessments using the GYM

The GYM assumes that the historical recruitment distributions are independently lognormally distributed, with means and variances equal to those estimated from length–density analyses of trawl survey length–frequency data. By sampling repeatedly from these recruitment distributions, and applying historical catches and a potential future catch, each run of the model produces $N$ sets of simulated projections of fishable biomass \(B_{y+1}, y = 1, \ldots, Y+P, n = 1, \ldots, N\). Here, simulation year $Y$ is the latest year for which we have data from the fishery, and simulation years $Y+1$ to $Y+P$ represent projections into the future. The decision rules for analysis of the output of the GYM require monitoring of the distributions of the fishable biomass in the final year and of the proportion of times the spawning stock biomass (SSB) fell below 20% of its pre-exploitation level. For simplicity of presentation, we concentrate here on a single statistic related to the status of the stock: the ratio of the spawning stock biomass in the final year of a projection to the median unexploited spawning stock biomass for that projection. A description of application of the method to the full assessment for *D. eleginoides* in Subarea 48.3 is given in SC-CAMLR (2000).

In the current GYM, the posterior distribution of $B_{y+1}$ values across simulations is determined simply by calculating a frequency distribution of values of \(B_{y+1,n}, n = 1, 2, \ldots, N\), according equal weight to each of the $N$ values. Now suppose that, from the GLM fitted to the CPUE data, we also have a series of standardised CPUEs \(l_y, y = 1, \ldots, Y\). Again for simplicity of formulation, here we assume that there is a standardised CPUE for each year of historical data. In practice, there are gaps, but adjustment for this is simple. The essential idea of the new method is to give different weights to each of the $N$ trajectories of \(B_{y+1}\), where the weights reflect how well each of the simulated historical fishable biomass trajectories match the trends in the standardised CPUE data.

The method proposed in Gasiukov (1999) set these weights to one or zero respectively, depending on whether or not the (scaled) simulated trajectory passed through the 95% confidence limits for the standardised CPUEs. We propose that the weights assigned to each trajectory should be proportional to the statistical likelihood of the standardised CPUE data given the simulated trajectory. Heuristically, this would appear to be a sensible approach, as simulated trajectories more closely matching the CPUE trends would then be accorded greater weight, while no simulated trajectories would be ignored completely. Fortunately, this approach is not only heuristically sensible, it also has a completely sound formal statistical basis, since the approach is entirely equivalent to using importance sampling, with the likelihood of the CPUE data for each simulation acting as the importance function (McAllister et al., 1994).

Taking Account of a Time Series of CPUE

The time series of catch per unit effort (CPUE) provides an indication of the trajectory of fishable biomass of toothfish in the South Georgia region. This time series could provide a means of determining which of the replicate trials in the GYM are the most plausible for this fish stock and therefore which could be given more weight in the overall assessment than other trials which do not follow the trends shown in the CPUE. A first step towards this was contained in Gasiukov (1999). The method proposed in that paper essentially restricted the use of simulated trajectories in GYM decision rules to those which also passed through the 95% confidence limits of the annual standardised CPUEs (after rescaling). In an example based on the 1998 assessment results, where the standardised CPUE trend suggested a greater decline than that in the median biomass from the GYM, Gasiukov (1999) showed that application of this method led to lower recommended catch limits. However, the strictness of the criterion used to accept or reject GYM trajectories meant that only a very small proportion of these trajectories could be used in the assessments, and WG-FSA had no time to consider this method in detail.

The aim of this paper is to present an alternative and less statistically ad hoc method which uses all the GYM-simulated trajectories. This method had been suggested during discussions of Gasiukov (1999); see SC-CAMLR (1999), paragraph 3.144. The method itself is described in the next section, after which a series of goodness-of-fit diagnostics are suggested and an example analysis is given.

Incorporating CPUE Data into GYM Assessments using Importance Sampling

The GYM assumes that the historical recruitment distributions are independently lognormally distributed, with means and variances equal to those estimated from length–density analyses of trawl survey length–frequency data. By sampling repeatedly from these recruitment distributions, and applying...
For the $i$th simulated trajectory, we have fishable biomasses $\{B_{i,y}, y = 1, \ldots, Y\}$ and corresponding standardised CPUEs $\{I_{i,y}, y = 1, \ldots, Y\}$. If $q$ is the catchability coefficient, then assuming that the CPUEs are lognormally distributed, the negative log-likelihood of the CPUEs with respect to the simulated biomass trajectory is given up to a constant by

$$-\ln L(q, \sigma^2 | B) = \frac{Y}{2} \ln \sigma^2 + \frac{1}{2 \sigma^2} \sum_{y=1}^{Y} \left( \ln \frac{I_{y}}{qB_{y,n}} \right)^2$$

(1)

Differentiating with respect to $q$ demonstrates that

$$\hat{q} = \frac{1}{Y} \sum_{y=1}^{Y} \ln \left( \frac{I_{y}}{B_{y,n}} \right)$$

(2)

and differentiating with respect to $\sigma^2$ implies that

$$\hat{\sigma}^2 = \frac{1}{Y} \sum_{y=1}^{Y} \left( \ln \frac{I_{y}}{qB_{y}} \right)^2$$

(3)

Consequently, when calculating the weighted frequency distributions, the weights used for simulation $n$ should be proportional to

$$\exp \left[ -\frac{Y}{2} \ln(\hat{\sigma}^2) - \frac{1}{2 \hat{\sigma}^2} \sum_{y=1}^{Y} \left( \ln \frac{I_{y}}{qB_{y,n}} \right)^2 \right]$$

(4)

Apart from calculating weighted rather than unweighted frequency distributions, results of an analysis of the GYM outputs necessary for the production of estimates of long-term yields are otherwise entirely unchanged.

Diagnostics

As always, it is also important to examine appropriate diagnostics of goodness of fit of the model. The simplest of these, and the easiest to interpret, is a plot of the percentiles (e.g. 2.5, 50 and 97.5%) of the importance-weighted posteriors of annual fishable biomasses and of the standardised CPUEs rescaled by dividing by the posterior median $\hat{q}$. If more than a small proportion of the rescaled standardised CPUE values fall outside the 95% posterior credibility intervals, then this provides evidence of a lack of fit (McAllister and Kirchner, 2001).

A more formal test of goodness of fit (Gelman et al., 1995) is based on the model deviance. The test statistic here is

$$D = \sum_{y=1}^{Y} \left( \frac{I_{y} - \hat{q}B_{y}}{\sigma^2_B} \right)^2$$

(5)

where $\{B_{y}\}$ is the trajectory of biomasses accorded the greatest importance weight, and $\hat{q}$ and $\hat{\sigma}^2$ are the corresponding estimates of $q$ and $\sigma^2$ from equations 2 and 3. In equation 5, the denominator is estimated variance of the predicted CPUE. The highest weighted trajectory was chosen for use in this calculation, rather than the sequence of medians of the posteriors for the annual fishable biomasses, as the latter sequence does not represent an actual simulated trajectory.

Under the assumptions of the model and the likelihood in equation 1, the statistic $D$ is distributed as chi-squared with $Y-1$ degrees of freedom. Unusually in this case, the appropriate test is two-sided, with departures at either end indicating a failure of either model assumptions or mis-specification of the likelihood function.

It is also necessary to check that there is not a small subset of the trajectories that is accorded very high weight relative to the remainder. This is simply done (McAllister and Ianelli, 1997) by examining the frequency distribution of importance weights and ensuring that the maximum importance weight is not too large (e.g. less than 1% if all weights have been scaled to sum to 1).

Finally, not so much as a diagnostic but rather to examine the effect of the importance weighting, plots of (for example) the 50 highest-weighted and 50 lowest-weighted trajectories of fishable biomass along with the appropriately rescaled standardised CPUEs allow an appreciation to be gained of which trajectories fit the rescaled CPUEs well and which fit badly (and are therefore down-weighted).

Example Application

The new method was used in the assessment undertaken for $D. eleginoides$ in Subarea 48.3, as described in SC-CAMLR (2000), paragraphs 4.148 to 4.152. Here, we illustrate the method using
one set of simulated projections from the GYM (examining the implications of a long-term catch of 4 500 tonnes) and the GLM-standardised CPUE series (SC-CAMLR, 2000, paragraphs 4.109 to 4.117). Input parameters for the GYM are given in SC-CAMLR (2000), Tables 33 and 34. A total of 1 001 trajectories was simulated.

The stock status indicator examined was the spawning stock depletion, which is calculated for each projection as the ratio of the SSB in the final year to the median of the unexploited SSB (Constable et al., 2000). Note that because of this definition, spawning stock depletion ratios can therefore be greater than 1. Figure 1 shows two posterior frequency distributions for the spawning stock depletion, one calculated using the new method with importance weighting, and the other calculated using equal weighting.

Somewhat surprisingly, given views expressed in earlier years about apparent discrepancies between the trends in standardised CPUE and GYM projections, the differences between the two frequency distributions is less than might have been expected. Furthermore, what differences there are imply that the mean spawning stock depletion is slightly higher under importance weighting (i.e. when trends in CPUE are taken into account) than under equal weighting (when they are not).

The reason for this is easily seen when the sets of 50 highest-weighted and 50 lowest-weighted trajectories given in Figures 2(a) and (b) are examined. As would be expected, the highest-weighted trajectories generally follow the trend of the standardised CPUEs, and indicate that the stock has been reduced from its unexploited levels by fishing. In contrast, the lowest-weighted trajectories fail to match the standardised CPUE trends at all, and crucially they generally indicate higher stock sizes at the end of the fishing period than before it. These trajectories would therefore correspond generally to larger spawning stock depletion ratios. As these are down-weighted by importance sampling, but given equal weight by the alternative method, the reason for the observed difference between the posterior frequency distributions is now clear.

While the 50 highest-weighted trajectories generally follow a similar trend to the standardised CPUEs, there are obvious qualitative differences between the trends. Whether or not these differences are significant is examined in Figure 3, in which are plotted percentiles of the importance-weighted posteriors for annual fishable biomass and the rescaled standardised CPUEs. As can be seen, most of the rescaled CPUEs fall within the 95% posterior credibility intervals for fishable biomass, but in two years (1996 and 1997), they fall just outside. The second proposed goodness-of-fit statistic, based on the model deviance, was calculated to be 13.66, very close to the mean of a $\chi^2_{12}$ random variable, and thus providing no evidence of lack of fit.
Figure 2a: Trajectories of fishable biomass accorded the 50 highest importance weights, and scaled standardised CPUE (solid line). Scaling using the average catchability coefficient for the 50 highest-weighting trajectories.

Figure 2b: Trajectories of fishable biomass accorded the 50 lowest importance weights, and scaled standardised CPUE (solid line). Scaling using the average catchability coefficient for the 50 highest-weighting trajectories.
Figure 3: Percentiles of importance-weighted posterior distributions of annual fishable biomasses and standardised CPUEs rescaled by dividing by the median of the importance-weighted marginal posterior for the catchability coefficient.

Figure 4: Frequency distribution of importance weights.

Finally, Figure 4 shows the frequency distribution of importance weights (scaled so that they sum to 1). The largest weight, 0.19%, is much less than the limit of 1% suggested by McAllister and Ianelli (1997); in fact the ratio of the largest weight to the smallest is only 5, a very small value. This certainly indicates that no small group of trajectories has been accorded high importance weighting relative to the rest.

DISCUSSION

The approach described here for incorporating standardised CPUE trends into GYM assessments of long-term annual yield appears to resolve the problems experienced by WG-FSA over the last three years in assessing the stocks of *D. eleginoides* in Subarea 48.3. Fortunately, it also requires only very small modifications to the computer programs implementing the GYM analyses and it
was used successfully at the 2000 meeting of WG-FSA (SC-CAMLR, 2000). It should also be noted that this approach can easily be extended to allow incorporation of indicators of stock status other than standardised CPUE into GYM assessments if desired. All that is required is that the likelihood of the data, given the simulated biomass trajectory (or any other related simulated trajectory), can be calculated.

While the effect of using importance weighting was only small in the _D. eleginoides_ assessment in Subarea 48.3, it is clear that the method helps to reduce the weight of apparently unrealistic projections in the analyses. The adjustment to be made to the indicators of stock status depends on the variation in potential trajectories of fishable biomass. As the time series of recruitments increases in length along with the time series of standardised CPUEs, we would expect that greater filtering of the trials would be possible and better estimates of stock status will result.

Overall, the diagnostics examined provided few indications of significant lack of fit. Nevertheless, the qualitative differences in CPUE and fishable biomass trends seen in Figure 2 are still the cause of some concern. These differences could result from problems in the underlying assumptions or parameterisation of the GYM, or in the standardised CPUE series.

With respect to the latter, the unbalanced nature of the CPUE data set on which the GLM standardisation was carried out has been noted many times, and SC-CAMLR (2000) noted in paragraph 4.111 that 'doubt must still remain about how well the relative levels of standardised CPUEs in early and later seasons have been estimated'. Unfortunately, there is no clear way of resolving this problem.

With respect to the former, the current implementation of the GYM assumes a consistent selectivity-at-length pattern across years. A notable feature in recent seasons has been a tendency by the some of the fleet to fish in shallower waters, taking more smaller fish and fewer larger fish than in the deeper waters fished in earlier seasons. The GLM CPUE analysis found that depth was a significant factor and this has been taken into account in the standardisation. Adjustments to the GYM to allow for annually variable selectivity patterns are planned, but they have not yet been completed. It is possible that, with them, some of the qualitative differences in trends may be resolved.

CONCLUSIONS

Apart from the changing selectivity of the fishery, all available data are now included in the assessment of long-term annual yield of _D. eleginoides_ in Subarea 48.3. The management procedure for this fishery now includes regular surveys of recruitment-aged fish, CPUE data from the fishery, an assessment method and a set of decision rules to specify catch limits in the coming seasons. An important further step in the assessment process will be to evaluate the management procedure as a whole to determine how well the procedure might achieve CCAMLR objectives under different scenarios for dynamics of the population of _D. eleginoides_ in the region and the interaction of the fishery with the stock (Constable et al., 2000) and to compare and evaluate the utility of other available assessment methods.

ACKNOWLEDGEMENTS

The authors are grateful to Murdoch McAllister for his advice in setting up this problem in terms of an application of importance sampling, and for providing guidance on appropriate diagnostics. We also thank Tony Smith and an anonymous referee for their comments on the manuscript. Geoff Kirkwood thanks Pavel Gasiukov for his forbearance in tolerating his confusion between importance sampling and application of the sampling/importance resampling (SIR) algorithm in an earlier version of this paper.

REFERENCES


Liste des figures

Figure 1: Distributions de fréquences a posteriori du rapport entre la biomasse du stock reproducteur de 1999 et la biomasse du stock reproducteur médian d’avant l’exploitation en utilisant des pondérations selon l’importance ou des pondérations égales.

Figure 2a: Trajectoires de la biomasse exploitable pour les 50 pondérations selon l’importance les plus élevées, et la CPUE normalisée étaillonnée (trait continu). L’étalonnage repose sur le coefficient moyen de capturabilité pour les 50 trajectoires pondérées les plus élevées.

Figure 2b: Trajectoires de la biomasse exploitable pour les 50 pondérations selon l’importance les moins élevées, et la CPUE normalisée étalonnée (trait continu). L’étalonnage repose sur le coefficient moyen de capturabilité pour les 50 trajectoires pondérées les moins élevées.

Figure 3: Percentiles des distributions a posteriori pondérées selon l’importance des biomasses annuelles exploitées et des CPUE normalisées ré-étalonnées en divisant par la médiane de la marge a posteriori pondérée selon l’importance pour le coefficient de capturabilité.

Figure 4: Distribution de fréquence des pondérations de l’importance.

Список рисунков

Рис. 1: Апостериорное частотное распределение отношения биомассы нерестового запаса в 1999 г. к медианной предэксплуатационной биомассе нерестового запаса, используя веса важности и равные веса.

Рис. 2а: Траектории промысловой биомассы с 50 наивысшими весами важности и пересчитанный стандартизированный CPUE (сплошная линия). При пересчете использовались средние коэффициенты уловистости для 50 траекторий с наибольшим весом.
Figura 1: Distribuciones resultantes de las frecuencias de la razón entre la biomasa del stock de desove en 1999 y la mediana de la biomasa del stock de desove antes de la explotación, utilizando ponderaciones de importancia o equivalentes.

Figura 2a: Trayectorias de la biomasa explotable con las 50 ponderaciones más importantes, y CPUE normalizado a escala (línea sólida). Para aplicar la escala se utilizó el promedio del coeficiente de capturabilidad de las trayectorias con las 50 ponderaciones más altas.

Figura 2b: Trayectorias de la biomasa explotable con las 50 ponderaciones de menor importancia, y CPUE normalizado a escala (línea sólida). Para aplicar la escala se utilizó el promedio del coeficiente de capturabilidad de las trayectorias con las 50 ponderaciones más altas.

Figura 3: Percentiles de las distribuciones de las biomasa anuales explotables y del CPUE normalizado luego de aplicar ponderaciones de importancia, ajustadas a escala dividiéndolas por la mediana del coeficiente de capturabilidad resultante de la ponderación de importancia.

Figura 4: Distribución de la frecuencia de las ponderaciones de importancia.