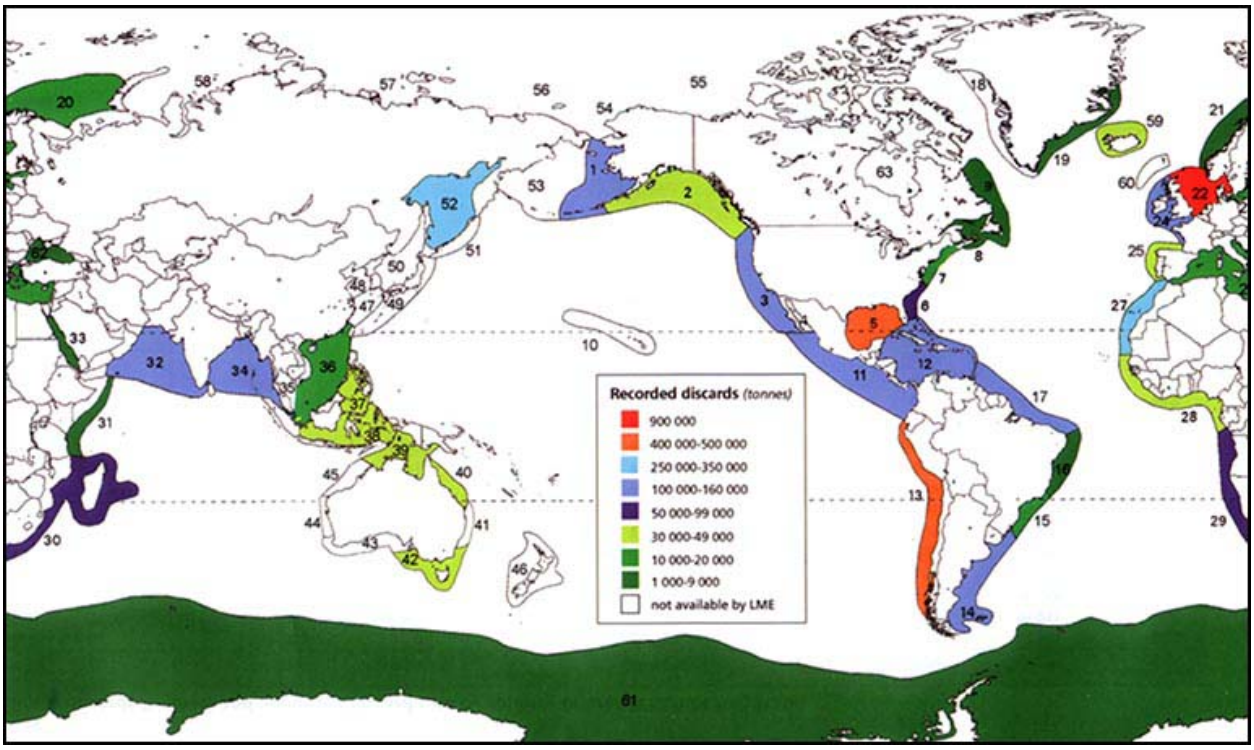
# Report on the review of Cumulative Discard Methodology

By

**Robin Cook**




*Source: FAO*

Prepared for
**Center for Independent Experts**
**Independent System for Peer Review**

# Contents

# Executive summary

i. In 2010 the cumulative discard methodology was reviewed and recommended a further review when more data for analysis had accrued. This review fulfils that recommendation and took place in Gloucester, MA from the 7-9th November 2016.

ii. The cumulative method uses a separate ratio estimator that expresses species discards as a ratio of the total biomass kept of all species to raise sampled discards to fishery level. This is a robust estimator that is known to perform well at least in mixed demersal fisheries.

iii. Since the purpose of the estimator is to track the accumulation of discarded catch during the fishing year, it should perform well provided the observer coverage is adequate. However, since the estimator expresses a ratio of accumulated discard and kept catch, it is not representative of the instantaneous rate of discarding at any point in time. If applied retro-actively it may introduce bias to the historical estimates of discards.

iv. A software package developed in R, DiscaRd, has been developed to explore alternative stratification schemes and transition rate methods that may improve the discard estimates. This uses bootstrap methodology to quantify uncertainty around the discard estimates. It is an extremely useful tool.

v. The uncertainty calculated from the bootstrap, and the CVs calculated using the Cochrane formula are conditioned on the assumption that the discard samples are measured without error. This means that the error distribution of the discards is a minimum estimate.

vi. The bootstrapped discard distributions provide a basis for comparing the performance of different stratification scenarios, but because they only account for sampling error they probably are not sufficiently comprehensive to measure the probability of premature fishery closure or exceeding the catch cap.

vii. The longfin squid fishery analysis suggests that a trimester based stratification would improve the butterfish discard estimates. For the other fisheries it was less clear that adopting higher or lower levels of stratification offered significant advantages. These analyses did, however, provide a basis for further analysis of promising stratification schemes.

viii. Because the methodology is applied from the start of the fishing year, early season estimates of discards suffer from very low observer coverage. To overcome this problem, a transition rate is calculated which uses samples from the previous year. Two principal transition rate methods were investigated, the “five trip” rule and the moving window. If, as is likely, the true discard rate is time varying, the five trip rule will be less biased but prone to noise while the converse will be true for the moving window method. It was not possible to identify which approach is superior without a simulation experiment where the true discard rates are known.

ix. Estimates of cumulative discards were sensitive to the choice of transition rate method but problems with the transition rate method will not affect the end-year estimate of accumulated discards because samples from the previous year fall out of the calculation.

x. The ratio estimator is likely to work best for mixed demersal fisheries rather than high volume pelagic fisheries where discard percentages are generally small. It would be worth

considering whether application of the current methodology and management tools for protecting haddock and river herring and shad (RHS) could be improved or replaced with a more robust approach that took into account the somewhat different capture processes in these fisheries.

xi. It is possible that some of the problems of the current methodology related to the transition rate and the estimation of uncertainty could be overcome with the use of time series models. The “memory” in these model would avoid the need to calculate a transition rate. It is recommended that such an approach be explored.

xii. It may be beneficial to perform a more detailed exploratory analysis of discard ratios to identify potential differences by stratum before running the full simulations. This would help in delimiting the more promising stratifications and whether sample sizes could adequately support higher levels of stratification. This could usefully be done as a joint exercise with the NEFSC.

# Background

1. Many fisheries take a bycatch of fish that are not the intended target because they may be too small, comprise the wrong species, have no commercial value or the vessel may not have sufficient catch quota to land the fish. These fish are often discarded and generally have low post catch survival which affects the overall mortality rate on the population. In a number of New England fisheries discards are monitored both to be able to perform more accurate stock assessments and so that fishery managers can control the catch within agreed limits. This may be to ensure fishing mortality rates and stock biomass conform to relevant reference points or that vulnerable species are protected. The approach is to use data collected from observer trips to estimate the current cumulative discard catch from the start of the fishing year. Cumulative discard estimates are then used for in-season management adjustments and for computing annual catches to monitor compliance with Annual Catch Limits.
2. In 2010 the methodology used to track the discard catch for management purposes was reviewed and it was suggested that the methodology should be reviewed again when more observer data had been accumulated. The current review requested by the Greater Atlantic Regional Fisheries Office (GARFO) fulfils that recommendation and took place in Gloucester, MA from the 7-9 November 2016. Draft working papers were received on the 26$^{th}$ October and finalized papers were available on the 4$^{th}$ November. These were reviewed prior to the meeting. In addition, reports for the 2010 review were examined. During the meeting, the reviewer participated in the discussions and commented on various aspects of the analyses presented. On the final day of the meeting, preliminary findings were discussed with the analysts including options for further development of the method.

# Summary of findings

3. This review is confined to the statistical methods used to estimate discards given the samples the available. It assumes that observer coverage is, in general, adequate, representative and unbiased. The latter are extremely important issues that merit careful review in their own right to ensure that the overall performance of the monitoring system is sound, but they were not subjects for this review.
4. Comments on each TOR are given below. However, it is perhaps useful to provide a general review of the methodology first in order to put in context more specific comments.

## Current methodology

5. At the heart of the methodology is a ratio estimator, r, that is used to scale the discard sample from observer trips to the total fishery. The estimator is the ratio of the sampled discards (for a given species) to the total sampled catch of all species kept (the "kept all" or $k_{all}$). Applying this ratio to the *kept all* ($K_{all}$) for all trips gives an estimate of the total discards by species. The estimator has been shown to perform well relative to other estimators (Stratoudakis *et al.*, 1999) and was favoured in the 2010 review.
6. The ratio is calculated on a cumulative basis from the start of the fishing year ($t_0$) so that for the current point in time ($t_c$) all the samples between $t_0$ and $t_c$ are summed before the r value for the year-to-date is calculated. Such an estimator is appropriate for making an estimate of total <u>accumulated</u> discards at any point in time up to the end of the fishing year. However, it is

important to bear in mind that the mean r calculated at $t_c$ may not be representative of the discard ratio at some earlier point in the fishing year if the discard rate is a time varying quantity. This is important if the average discard ratio at time $t_c$ is used retro-actively to re-calculate discards at an earlier time period and may result in biased estimates of historical discards.

7. An advantage of the cumulative discard estimator is that, in principle, the precision of the estimate will improve over the course of the fishing year because of the increasing number of samples available. As currently applied, the estimator is reset at the start of each new fishing year and this presents the problem that at this point there are few, if any, samples from the current year to perform the calculation. In order to overcome this difficulty, a "transition rate" is calculated using an *ad hoc* rule where a weighted average of the mean r for the previous year and the current year is used until at least 5 observer trips have been sampled. Hence the preceding year mean discard ratio is, in effect, being used as a "prior" for the start of the current year. It makes sense to smooth the initial estimates of r for the current year in order to mitigate the problem of small sample sizes, but there is a trade-off between bias and precision. Although it has the virtue of simplicity, the weighting procedure presently used is unlikely to strike an optimum balance of these effects. This is because the weighting is a hard wired formula that does not explicitly consider potential bias in the mean r for the previous year or the precision of the r value calculated for the current year.
8. The working papers presented at the meeting investigated an alternative "moving window" approach to smoothing the early values of r to obtain a transition rate. Here the r estimate at any point in the fishing year uses samples for a full year. At the start of the year r will be calculated almost entirely from samples in the previous year. As new samples are added during the current year, samples from the previous year are dropped so that by the end of the year only samples from the current year enter the estimate. Typically, this will result in much stronger smoothing than the "five trip" rule with a greater risk of bias if the true discard ratio is time varying. In most of the simulations presented at the meeting, the cumulative discard estimate rarely approached the catch cap in the early part of the year so the problem of the estimate triggering inappropriate management action is likely to be low. However, the cumulative discard estimates in this early period may give a misleading impression of the rate at which the catch cap is being approached which might lead to poor operational decisions. A similar problem will exist with the five trip rule except that the discard estimate after five trips is likely to be dominated by noise rather than systematic bias.
9. Within a fishery discard rates are likely to vary by sector as defined by for example, gear, area fished or season. The application of the current methodology is to account as far as possible for such differences by stratifying the observer samples and then calculating the mean r values by stratum to obtain discard estimates, which can then be summed to fishery level. Such stratification should increase the precision of the discard estimates by accounting for between-stratum variance. However, a high level of stratification will result in fewer samples per stratum so that the within-stratum variance can be very large. The choice of stratification is therefore a trade-off between within-stratum variance (mostly sampling level) and between-stratum variance (process variability). The additional complication in this instance is that catch caps are defined in terms of management units so a minimum stratification requires adequate

stratification to support those management units and is likely to require a minimum consequent level of observer coverage.

10. As applied, the present procedure for calculating the ratio estimator is based entirely on the design (stratification) matrix which relies directly on the observer samples. Such an approach does not attempt to model the discard process itself. Modelling the process may help to improve precision and predictability of the discards and is discussed further in paragraphs 32-38 where a time series model is outlined.

## DiscaRd package

11. At the meeting a number of alternative stratification schemes were investigated and their performance measured by the overall CV on the cumulative discards. These analyses were made using an R package "DiscaRd" developed for the purpose (WP#1). This is an extremely useful tool for extracting observer and trip data from respective databases and structuring them for routine calculations to simulate the cumulative discard methodology. The package allows the user to define strata and then bootstrap samples to obtain asymptotic estimates of variance for the quantities of interest. It will also calculate the standard Cochran based variance estimate for the ratio estimator presented in WP#1.
12. At the meeting there was some discussion of the way in which the bootstrap was configured particularly in relation to exchangeability of the samples. Both reviewers were concerned that the true discard ratio is likely to be time dependent, and hence drawing the samples at random may violate this dependency. After discussion with the analysts it was clear that in each draw of the bootstrap the time sequence of the samples was retained and treated in the correct temporal sequence in subsequent analyses. This is particularly important since it means that the results presented at the meeting will not be invalidated by any temporal effects that are otherwise not explicitly taken into account in the stratification.

## TOR 1

*For each fishery subject to in-season discard monitoring utilizing the cumulative discard method, summarize the variability in discard rate by measurable strata: fishery, gear, area, season, volume of catch, etc.*

13. For this TOR I interpret "discard rate" to mean the cumulative discard estimate as it develops over the year. Working papers 2-5 summarise analyses for four fisheries: Atlantic sea scallop, long fin squid, Atlantic herring and mackerel, and groundfish. These papers consider various stratifications and transition rate options (paragraphs 8 and 9) using the bootstrap methodology described in WP#1 and discussed in paragraph 11 and 12. For each scenario a "focal year" was chosen to hindcast the development of discards over the year given a complete year of real observer data. The choice of year depended on the availability of samples given current management of the fisheries. Without detailed knowledge of the fisheries and the observer programmes these choices appeared to me to be appropriate.
14. Variability was characterised by the estimated distribution of cumulative discards from the bootstrapped samples and shown as percentiles in plots of the cumulative discards over the course of the year. The analytic (Cochran CV was also calculated either as an unweighted estimate or weighted by the stratum discards. The latter estimate is perhaps most useful in identifying which strata merit the greatest sampling effort in order to improve precision.
15. It is important to be clear that these analyses identify variability in the discard estimates conditioned on the assumption that the discards as sampled on each observer trip are

themselves measured without error. The implication is that with 100% coverage the discards would be known exactly. This is doubtful given the practical difficulties of operations at sea that are likely to give rise to significant measurement error and possible bias. While such variability is beyond the scope of this review, it is important to understand the limitations of the calculated variance, especially if managers choose a particular percentile as a basis for an intervention. For example, if the point in time where the upper 95$^{th}$ percentile crosses the catch cap is chosen as a risk averse strategy when the fishery is closed to protect a bycatch species, then the existing assumption will underestimate the true risk and management action will occur too late.

## TOR 2

*Identify more optimal applications of the current cumulative method for in-season estimation of discards in comparison to existing cumulative discard methodology and stratification schemes. Alternatives identified will include*

*a. Existing cumulative discard methodology and stratification scheme as a baseline*
*b. Pooling data across current stratifications to increase information and precision. As an example, pooling across sectors and gears.*
*c. Including seasonality as a stratification*
*d. Allocate/restrict sampling requirements to those strata which in aggregate constitute a target fraction of total stock-specific discards. (i.e, excluding or minimizing sampling for strata with negligible discard totals).*

*Sea scallop fishery*

16. Seasonal patterns in discarding of yellowtail and windowpane flounder showed some differences between seasons but this was not consistent across years and was not therefore adopted as a stratification factor. While I agree there is no consistency in seasonal patterns, there do nevertheless appear to be differences by season within years which suggests that the discard ratio is time varying and may need to be accounted for. A modelling approach to this issue is discussed in paragraphs 32-38.
17. Other stratifications that collapsed the baseline fleet/gear/area stratification into gear only or area/gear performed much the same as the baseline when evaluated on the basis of the estimated discards and CV. The most parsimonious stratification would therefore use only gear as a factor.

*Longfin squid fishery*

18. Exploratory analyses showed convincingly that discard ratios of butterfish differed by trimester with trimester 2 showing much lower ratios. Failing to account for this temporal effect resulted in bias. Other factors such as vessel length, volume of catch and proportion of longfin squid in the landings were also investigated. However, while some effects were apparent, these were not consistent. It appears, therefore that trimester is the main stratification that needs to be accounted for.

*Atlantic herring and mackerel*

19. A number of potential stratification criteria were investigated that included gear, area, vessel category and landings category. For haddock discards observer coverage meant that with high levels of stratification the available sample size within a stratum was too small and discard ratios could not be estimated with any precision. As a result, while fine grain stratification might more

realistically separate true differences in discard ratios, sampling coverage was too low to exploit this effect. In general, therefore, there was little indication improvements over the baseline from a higher level of stratification. However, it did appear that a simple area stratification using statistical area 522 and "other Georges Bank", improved the precision of the estimates.

20. For the river herring and shad (RHS) catch cap the small number of samples available for analysis meant that no useful improvement in precision could be detected in using a higher level of stratification.

*Groundfish*

21. Stratification criteria included vessel length category, combining sectors and broad stock area. Due to the complexity of the groundfish fishery the results show no consistency in the CV across stocks, though it did appear the vessel length category tends to improve precision while combining sectors was generally worse. The analyses presented were essentially exploratory and provide a basis for further work rather than indicating definitive conclusions.
22. In this fishery discards are counted against the catch. Bootstrapped estimates of the total catch are therefore dominated by the actual landings when the discard ratio is low. Hence in Table 3 in WP#5 the 95% CI on the catch is estimated to be very small. This raises the important question of the purpose for which the estimates are being made. If, as in this case, the discard rate is low then only a very crude estimate of discards may suffice and stratification to improve precision could be a relatively minor issue for the purposes of management of the total catch.

## TOR 3

*Methods identified in TOR 2 will be compared using the following metrics*

*a. Precision of the discard estimates for a given level of observer coverage*
*b. Consistency of discard estimates calculated over the course of the fishing year.*
*c. Precision and consistency of the CV discard metric for a given level of observer coverage*
*d. Sensitivity to missing or erroneous data.*

23. The analytical CV calculated from equations 1-4 in WP#1 is one way of characterising the precision of the estimated discards. As indicated earlier it assumes that the discard samples are measured without error and will therefore be a minimum estimate of the true CV. My understanding of the results presented was that the reported CVs in the working papers were based on this estimator and were not themselves derived from the bootstrap. It might be useful to calculate the analytical CV within each iteration of the bootstrap to obtain a distribution of the estimated CV to determine how well this estimator performs. It would provide a statistical basis for identifying whether alternative stratifications really do provide a measurable improvement. It also could be compared to the CV obtained empirically from the bootstrapped samples to assess consistency. A similar bootstrapped calculation could be done for the estimate of the number of samples to achieve a target CV to evaluate how well such a prediction can be made.
24. The analyses presented were based on real samples from previous years and therefore the observer coverage is predetermined by the actual coverage in the year concerned. The performance of different stratifications under different levels of coverage was not therefore investigated. As useful diagnostic presented shows the calculated CV under different levels % coverage that can be compared to the CV observed for the actual % coverage. In addition, this CV can be calculated weighted by the stratum discards. Disparity between these alternative

estimates indicates the strata where increased coverage will most reduce the CV of the discard estimate and would be useful for in-year decision making on observer effort allocation.

25. It is not very clear how TOR 3d is best investigated. Qualitatively it is obvious that missing data equates to lower observer coverage and this will inflate the CV. However, data can be "missing" in many different ways (e.g. by gear, area, time step, etc.) and the extent to which these adversely affect the overall CV will be case specific. It would be necessary to describe likely missing data scenarios before embarking on a comprehensive analysis of the problem. A similar comment can be made about "erroneous data". It would be necessary to characterise the nature and scale of the errors before any analysis was pursued.

## TOR 4

*Examine methods for including data from past years to improve predicting the in-season estimation of discards.*

26. Three possible recipes for using prior information were considered. These included the five trip rule, the moving window (described above), and in WPs #3 and #4 a variant that included prior samples until the number of new samples gave a target CV. All of these approaches make the pragmatic assumption that historical estimates of the discard ratio are the best predictor of the current ratio in the absence of better information. Paragraphs 7 and 8 above discuss the issues with such estimators.
27. Since the analyses presented were based on real samples where the "true" values of the discard ratio are in reality unknown, comparison of the different transition rate assumptions will only reveal their different effects rather than identifying unambiguously which performs best. The analyses do show that in-year estimates are sensitive to the transition rate assumption, though this problem resolves itself as the transition rate value falls out of the calculation towards the end of the year. It does mean, however, that early season estimates may be heavily dependent on the transition rate assumption and this will have implications for management if catch caps are approached rapidly.
28. Full testing of the transition rate rules would require a simulation study where the underlying truth was known. This in turn would mean trying to capture the characteristics of the fishery in an operating model which would be used to generate samples on which the transition rate could be tested. As the real world is itself unknown, it would be necessary to devise a number of candidate operating models to show that a particular transition rate assumption performed robustly when faced with alternative worlds.

## TOR 5

*Use archived data to simulate in-season behavior (with various time steps and discarding patterns) and recommend a preferred method for each fishery with consideration of the following:*

*a. Feasibility, particularly the implications of stratum size and within-year pattern of precision.*
*b. The probability and timing of premature closure (i.e. false positive).*
*c. The probability and magnitude of exceeding a cap (i.e. e. false negative).*

29. Most of the issues related to TOR 5a are discussed in the sections above. As regards 5b and 5c this was addressed in the analyses by considering the estimated percentiles of cumulative discards derived from the bootstrap. The point at which a chosen percentile intersects the catch cap is a potential indicator of a false negative or positive. These percentiles are likely to be an indicator of the precision of the discard estimate. This is useful when investigating alternative stratification schemes or observer coverage because the change in precision will at least indicate

the direction of improvement. Whether these percentiles offer an adequate means of assessing false negatives or positives is more questionable since what is implied is an absolute, as opposed to a relative, estimate of precision.

30. As discussed earlier these percentiles do not account for measurement error in the discard samples and this may be very large in some instances. If measures to satisfy TOR 5a and 5b are required, then some estimate of the discard measurement error is needed.
31. There is also the question of the consequences of the management response if a particular percentile is chosen to trigger a management response. If a fishery is closed due to a false positive, this will benefit the stock but impose a cost on the industry. Similarly, if the fishery is not closed due to a false negative there will be a cost to the stock. Choosing an appropriate percentile for management purposes would need to consider loss of utility rather than simply estimating probabilities of exceeding the catch cap.

## Time series modelling of the discard ratio

32. The current cumulative method uses the sampling design to estimate the discard ratio. Samples within each stratum are treated as independent and used to calculate a mean and variance. Such an approach is vulnerable to small sample sizes and this is most acute at the start of the fishing year when few samples have been collected in the observer programmes. In trying to overcome this difficulty, a number of transition rate methods have been used to improve the estimates. The underlying assumption in these methods is that samples from the past provide a useful predictor of the current discard ratio.
33. There are good reasons to suppose that the discard rate varies with time and that this pattern may not repeat itself each year. For example, in many groundfish stocks where discards comprise small fish of the target species, the discard rate can be heavily dependent on the size of incoming year classes where large cohorts lead to high discard rates. Year class strength is highly variable so annual discard rates are likely to differ. Furthermore, where discards by species are expressed as a ratio of the total kept species there will be multiple year class effects influencing the estimator so there is no reason to expect it to be constant over the year. In spite of effects such as this, one might expect discard rates over short time periods (e.g. days or weeks) to be similar since the impact of biological phenomena such as recruitment, growth and migration will be gradual.
34. The complexity of biological effects makes it hard to predict with any precision how discard rates will change over time and modelling such processes presents a formidable challenge. It might therefore be useful to consider using time series models where the discard ratio is modelled as a random walk through time. This is illustrated here assuming a single stratum but could easily be generalised to multiple strata. Suppose then that for a particular species the discard ratio at time t, $r_t$, is updated by:

$$(1) \qquad r_t = r_{t-1}\exp\left(\varepsilon_{r,t}\right), \varepsilon_{r,t} \sim N(0, \sigma_r^2)$$

Where $\sigma_r^2$ is the variance of a white noise process. Note that r here is not the cumulative discard ratio as described in the working papers, but the instantaneous value at any point in time. Since the total kept catch in the sample, $k_{all}$, is known, the true sampled discard, d, is given by:

$$(2) \qquad d_t = r_t k_{all,t}$$

The observed discards from a sample, $\widehat{d_t}$, can then be expressed as:

(3) $\widehat{d_t} = d_t \exp(\varepsilon_{d,t})\,;\ \varepsilon_{d,t} \sim N(0, \sigma_s^2 + \sigma_m^2)$

Where $\sigma_s$ is the standard deviation of the sampling error and $\sigma_m$ is the standard deviation of the measurement error. Given a series of observations, $\widehat{d_t}$, it should be possible to estimate $r_t$ and then apply the ratio to $K_{all}$ from all recorded trips to get the stratum level discards, $D_t$, at any time interval. This could be done using a variety of approaches, though a Bayesian model using OpenBUGS (Lunn et al. 2009) would be an obvious way forward. Having estimated $D_t$ it is then simple to sum the D values over time to calculate the cumulative discard over any chosen fishing year. MCMC sampling would allow the estimation of the credible intervals of this or any other derived quantity and would circumvent the need to bootstrap the data.

35. There is the question of a prior on $r_0$, the discard ratio at the start of the time series. It is not obvious what this ratio would be since it can take any value (0,∞). However, if the proportion of fish discarded from the total sample is p, then r can be expressed as:

    (4) $r = \frac{p}{(1-p)}$

    And since p is defined on the interval (0,1) a beta distribution configured on the basis of historical proportion discarded could be used.

36. The potential advantages of such an approach are:
    a. The model has a memory so there is no need to model a transition rate. The transition rate is effectively expressed in equation (1).
    b. Modelling $r_t$ as time dependent means that bias as a result of retro-actively applying a mean cumulative discard ratio would be avoided and seasonal effects could easily be investigated without using externally imposed time strata with fixed boundaries.
    c. Provided there is sufficient data (the discard samples, $\hat{d}_t$ ) it should be possible to estimate both process error $\sigma_r$ and measurement error $\sigma_m$ ($\sigma_s$ can be calculated directly from the Cochranformula and adjusting for a log scale). Accounting for both sampling and measurement error may provide more realistic estimates of the uncertainty in the overall discard values.
    d. There is no need to restart the "clock" at the start of each fishing year. The stream of discard sample observations could be used over any time period. Longer time series would improve estimates of $\sigma_r$.
    e. Where samples are missing at a point in time, the model will automatically fill in a value using the projection equation. The precision of filled in values will, of course, be dependent on the quality of the samples either side of the missing value and the width of the time interval.
37. If data are sparse and/or of poor quality, it may not be possible to estimate both process and measurement error. In these circumstances specifying process error variance might be the best option as it amounts to a smoothing parameter. It might be possible to identify a value for process error using cross-validation.
38. The above model could be generalised to accommodate multiple strata. Just as $r_t$ is predictor of $r_{t+1}$, there may be correlation between adjacent strata which can be exploited to improve the estimates. Thus if $\mathbf{r_t}$ is a vector of stratum r values at time t, and $\mathbf{\Sigma_r}$ is a matrix expressing the between stratum covariance, one could consider the multivariate normal model:

(5) $log(\boldsymbol{r_t}) \sim MVN(\log(\boldsymbol{r_{t-1}}), \boldsymbol{\Sigma_r})$

as an alternative to equation (1).

# Conclusions and recommendations

39. The principal thrust of the analyses discussed at the review was to identify better stratification schemes and transition rate methods for the fisheries. This was based on the separate ratio estimator identified as the preferred method at the 2010 review. The terms of reference were highly specific and a thorough analysis using the best science available was applied within the limits of the TORs set.
40. The DiscaRd package provides a very valuable tool for investigating alternative transition rate and stratification scenarios which uses bootstrapping to estimate the posterior distribution of cumulative discards. These distributions account for sampling error only and will provide minimum estimates of the precision of the cumulative discards.
41. With the exception of the longfin squid fishery where trimester stratification was demonstrably better, the analyses tended not to show, definitively, improved stratification schemes. They did reveal possible further avenues of analysis, but it was not possible to firmly choose revised stratifications.
42. The bootstrapped discard distributions provide a basis for comparing the performance of different stratification scenarios, but because they only account for sampling error they probably are not sufficiently comprehensive to measure the probability of premature fishery closure or exceeding the catch cap.
43. Estimates of cumulative discards were sensitive to the choice of transition rate method during the period when prior samples dominated the calculation. Other than to note this sensitivity it is not really possible to demonstrate which method is best since the underlying “truth” is unknown. A simulation study with known underlying discard rates would be needed to test these methods.
44. The methodology applied in all the fisheries considered during the review was based on the same underlying estimator. The separate ratio estimator is probably best suited to fisheries using non-selective gear such as bottom trawls where a mixture of species is taken with a routine bycatch component. Pelagic fisheries that target single species in high volume using purse seines or mid-water trawls present a somewhat different problem since any bycatch is often a very small percentage of the total catch which makes sampling difficult and the large raising factors on the sample can seriously inflate errors. It would be worth considering whether application of the current methodology and management tools to protecting haddock and RHS could be improved or replaced with a more robust approach that took into account the somewhat different capture processes in these fisheries.
45. It is possible that some of the problems of the current methodology related to the transition rate and the estimation of uncertainty could be overcome with the use of time series models. The “memory” in these models would avoid the need to calculate a transition rate. Modelling the discard ratio as time dependent means that bias as a result of retro-actively applying a mean cumulative discard ratio would be avoided and seasonal effects could easily be investigated without using externally imposed time strata with hard boundaries. I would recommend that such an approach is explored.

46. Much of the analysis presented relied on the overall performance as measured by the CV or bootstrapped discard distribution to compare stratification schemes. As was done for the longfin squid fishery, it may be beneficial to perform a more detailed exploratory analysis of discard ratio to identify potential differences by stratum before running the full simulation. This would help in delimiting the more promising stratifications and whether sample sizes could adequately support higher levels of stratification. This could usefully be done as a joint exercise with the NEFSC.

## Comments on the review process

47. The review was conducted in a very positive atmosphere with excellent co-operation from the analysts, GARFO staff, and members of the public who participated in the meeting. The working papers provided documented a substantial amount of work, were well presented and received in a timely manner.
48. I felt the wording of the terms of reference reflected what was expected from the analysts rather than the reviewers which made it harder to respond to each term of reference individually as required by the CIE contract. Wording directed more clearly at what should be reviewed might help in the future. For example, “Review methods to estimate cumulative discards…” or “Review the choice of stratification and metrics to compare performance…” would help clarify what was expected.

## References

Stratoudakis, Y., Fryer, R. J., Cook, R. M., and Pierce, G. J. 1999. Fish discarded from Scottish demersal vessels: Estimators of total discards and annual estimates for targeted gadoids. ICES Journal of Marine Science, 56: 592–605.

Lunn, D., Spiegelhalter, D., Thomas, A. and Best, N. (2009) The BUGS project: Evolution, critique and future directions (with discussion), *Statistics in Medicine* **28**: 3049-3082.

# Appendix 1: Bibliography of materials provided for review

## Working papers

WP#1. Daniel W. Linden, Benjamin Galuardi, and Brant M. McAfee. Methods for examining in-season behavior of the cumulative discard estimation in the Greater Atlantic Region.

WP#2. Benjamin Galuardi. Cumulative discard methodology review for catch cap monitoring in the Atlantic sea scallop (Placopecten magellanicus).

WP#3. Jerome M. Hermsen. Cumulative discard methodology review for butterfish (Peprilus triacanthus) discards in the longfin squid (Doryteuthis (Amerigo) pealeii) fishery.

WP#4. Brant M. McAfee. Cumulative discard methodology review for catch cap monitoring in the Atlantic herring (Clupea harengus) and Atlantic mackerel (Scomber scombrus) fisheries.

WP#5. Daniel E. Caless Cumulative discard methodology review for groundfish discards in the Northeast United States groundfish fishery.

## Presentations

J. Michael Lanning. Cumulative Discard Methodology Review.

Brant McAfee. Cumulative Discard Methodology Review for Haddock and River Herring/Shad Catch Caps in the Atlantic Herring and Mackerel Fisheries.

Daniel W. Linden. Methods for examining in-season behavior of the cumulative discard estimation.

Benjamin Galuardi. Cumulative Discard Methodology in the Atlantic Sea Scallop (Placopecten magellanicus) Fishery.

Dan Caless. Cumulative Discard Methodology Review, Groundfish.

Jerome M. Hermsen. Cumulative discard methodology review for butterfish discards in the longfin squid fishery.

## 2010 review reports

R. O'Boyle. SSC Review of the Northeast Region Discard Estimation Methods for Groundfish Quota Monitoring and Annual Catch Limits.

James Bence. SSC Review of the Northeast Region Discard Estimation Methods for Groundfish Quota Monitoring and Annual Catch Limits.

Working papers associated with the 2010 SSC review were available at http://nefsc.noaa.gov/groundfish/discard/

# Appendix 2: Statement of Work

**National Oceanic and Atmospheric Administration (NOAA)**

**National Marine Fisheries Service (NMFS)**

**Center for Independent Experts (CIE) Program**

**External Independent Peer Review**

***Review of Cumulative Discard Methodology***

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards.

(http://www.cio.noaa.gov/services_programs/pdfs/OMB_Peer_Review_Bulletin_m05-03.pdf). Further information on the CIE program may be obtained from www.ciereviews.org.

Scope

The Greater Atlantic Regional Fisheries Office (GARFO) requests a review of the Cumulative Discard Methodology currently used to monitor fishery discards throughout the year. Cumulative discard estimates are used for in-season management adjustments and for computing annual catches to monitor compliance with Annual Catch Limits.

Requirements

NMFS requires three reviewers to conduct an impartial and independent peer review in accordance with the SOW, OMB Guidelines, and the TORs below. Reviewers shall have working knowledge and recent experience in:

- Advanced statistical sampling theory and survey design.
- Application of real time-estimation methods for decision-making and imputation.
- Risk analysis as applied to natural resource management
- Basic understanding of fishery monitoring.

Tasks for Reviewers

- Review background materials and reports prior to the review meeting
- Attend and participate in the panel review meeting ○ The meeting will consist of presentations by NOAA and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
    - The reviewers shall conduct an independent peer review in accordance with the requirements specified in this SOW, OMB guidelines, and TORs, in adherence with the required formatting and content guidelines; reviewers are not required to reach a consensus
    - Each reviewer may assist the Chair of the meeting with contributions to the summary report, if required by the TORs
- Deliver their reports to the Government according to the specified milestone dates

Foreign National Security Clearance

When reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for reviewers who are non-US citizens. For this reason, the reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/ and http://deemedexports.noaa.gov/compliance_access_control_procedures/noaa-foreign-nationalregistration-system.html. The contractor is required to use all appropriate methods to safeguard Personally Identifiable Information (PII).

Place of Performance

The place of performance shall be at the contractor's facilities and at the in-person review meeting in Gloucester, MA.

Period of Performance

The period of performance shall be from the time of award through December 31, 2016. Each reviewer's duties shall not exceed 14 days to complete all required tasks.

**Schedule of Milestones and Deliverables:** The contractor shall complete the tasks and deliverables in accordance with the following schedule.

| | |
|---|---|
| No later than September 22, 2016 | Contractor selects and confirms reviewers |
| No later than October 25, 2016 | Contractor provides the pre-review documents to reviewers. |
| **November 7 –9, 2016** | Panel review meeting and independent peer review in Gloucester, MA |
| November 23, 2016 | Contractor receives draft reports |
| December 7, 2016 | Contractor submits final reports to the Government |

* The Summary Report will not be submitted, reviewed, or approved by the CIE.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel

All travel expenses shall be reimbursable in accordance with Federal Travel Regulations

(http://www.gsa.gov/portal/content/104790). International travel is authorized for this contract.

Travel is not to exceed $15,000.

Restricted or Limited Use of Data

The contractors may be required to sign and adhere to a non-disclosure agreement.

**Principal Review Meeting Contact(s):**

Dr. J. Michael Lanning
Greater Atlantic Regional Fisheries Office
55 Great Republic Drive, Gloucester, MA 01930-2276
J.Michael.Lanning@noaa.gov (Phone: 978-281-9308) (FAX: 978-281-9333)

Hannah Goodale, Assistant Regional Administrator
Greater Atlantic Regional Fisheries Office
55 Great Republic Drive, Gloucester, MA 01930
Hannah.F.Goodale@noaa.gov Phone: 978-281-9101

**Terms of Reference for the Peer Review**

*Version: February 10, 2016*

Review of Cumulative Discard Methodology

1. For each fishery subject to in-season discard monitoring utilizing the cumulative discard method, summarize the variability in discard rate by measurable strata: fishery, gear, area, season, volume of catch, etc.

2. Identify more optimal applications of the current cumulative method for in-season estimation of discards in comparison to existing cumulative discard methodology and stratification schemes.

Alternatives identified will include:

   a. Existing cumulative discard methodology and stratification scheme as a baseline
   b. Pooling data across current stratifications to increase information and precision. As an example, pooling across sectors and gears. c. Including seasonality as a stratification
   d. Allocate/restrict sampling requirements to those strata which in aggregate constitute a target fraction of total stock-specific discards. (i.e, excluding or minimizing sampling for strata with negligible discard totals)

3. Methods identified in TOR 2 will be compared using the following metrics:
   a. Precision of the discard estimates for a given level of observer coverage
   b. Consistency of discard estimates calculated over the course of the fishing year.
   c. Precision and consistency of the CV discard metric for a given level of observer coverage
   d. Sensitivity to missing or erroneous data.

4. Examine methods for including data from past years to improve predicting the in-season estimation of discards.

5. Use archived data to simulate in-season behavior (with various time steps and discarding patterns) and recommend a preferred method for each fishery with consideration of the following:
   a. Feasibility, particularly the implications of stratum size and within-year pattern of precision.
   b. The probability and timing of premature closure (i.e. false positive).
   c. The probability and magnitude of exceeding a cap (i.e. e. false negative).

**Draft AGENDA**

*Version February 10, 2016*

**Cumulative Discard Methodology Review**

**GARFO**

**November 7 – 9, 2016**

- Welcome and Meeting Expectations
- Review of the cumulative discard method
- Discussion of pros and cons of the cumulative discard method current implementation
- Review of the modeling techniques to simulate in-season behavior with archived data
- By fishery utilizing the cumulative discard method:
  - Summary of the variability of the in-season discard estimation
  - Discussion of alternative applications of the cumulative method
  - Comparison of proposed alternate applications results
- Meeting Summary and Deliverables

**Peer Review Report Requirements**

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether or not the science reviewed is the best scientific information available.

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs.

   a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the summary report that they believe might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each TOR, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of this Statement of Work
   Appendix 3: Panel membership or other pertinent information from the panel review meeting.

# Appendix 3: Panel membership and other pertinent information from the panel review meeting.

Review panel:
Robin Cook, CIE
Shijie Zhou, CIE

GARFO Analysts:
Dan Caless
Benjamin Galuardi
Jerome Hermsen
Daniel Linden
Brant McAfee

GARFO staff:
J. Michael Lanning
Hannah Goodale

Members of the public including industry representatives, NGOs, NEFMC staff, NEFSC scientists participated in the meeting.