Independent peer review of *A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels*

conducted for the
Center of Independent Experts (CIE)
By

Joseph E. Powers

November 2021

Executive Summary

The report herein represents a review of: *Linden, D.W. 2021. A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels. Working Paper*. That analysis "predicts" the electronic monitoring catch in weight by species of a haul as a function of the Vessel Trip Records (VTR, logbooks) reported catch for a haul. The goal is to utilize VTR data to predict the discarded catch from video monitors of Electronically Monitored (EM) vessels for those EM trips where the video record has not been processed/reviewed. Essentially, the analysis is trying to determine the rate of reviews (and indirectly the cost) that are needed to achieve specified levels of precision and accuracy.

The data used were from a pilot program in which electronic monitoring was voluntarily chosen by vessels in lieu of at-sea-observers for the years 2017-2020. According to information provided in the webinar, video cameras were placed on each vessel to monitor activities. For this analysis there were data from 2,593 hauls across 801 sub-trips by 23 vessels. More than half of the effort came from bottom otter trawls (59%), a quarter from sink gillnets (25%) and the remaining effort was split between bottom longlines (8%) and handlines (8%).

Three models were evaluated: 1) the delta-model; 2) Zero-Inflated Poisson (ZIP); and 3) Poisson. The models utilized species intercepts, year effects and vessel effects. Each model predicts discarded catch based on the VTR data for each EM observed trip. So, the prediction is the catch on an EM vessel even though the portions of the EM video data have not been reviewed and not utilized. Results indicate that sufficient R^2 values were achieved with this data set with the two Poisson-related models performing better. Further analysis indicated that these two models could achieve a 30% CV of catch per trip for video review rates of 35% or higher. The analysis answered the basic question: yes, one can get usable precision and accuracy in predictions of a trip's EM observed vessel discarded catch by species by only looking at a portion of the EM data by relating it to VTR data from the controlled experiment that existed 2017-2020.

The Linden analysis is an initial step in the evaluation of possible EM data collection programs. Further analyses are needed to address the different objectives of discard estimation for assessment, within-season monitoring, post-quota monitoring and compliance. Subtle nuances in the implementation of an EM program are likely to alter the modeling approaches that could be used. To that end recommendations are made to: evaluate the cumulative catch by species for the 2017-2020 data set with alternative models and EM video review rates; differences in discarding behavior for under-sized vs post-quota discards, effects of growth and density within a year of a species; and the overall mix of EM vessels, at-sea-observed vessels and VTR only vessels.

Background

Estimation of total catch in a commercial fishery is important for stock assessment and effective management decision-making, as dictated by the Magnuson-Stevens Act. A primary challenge of catch estimation lies in the largely unknown portion of catch that is discarded at sea rather than retained. Catch can be discarded for multiple reasons related to regulatory compliance (e.g., species and/or size retention prohibited) and non-compliance (e.g., illegal actions incentivized by quota limitations). While the retained portion is recorded by multiple information streams, discard estimation typically relies on a sampling of trips that carry a human at-sea observer. Low sampling rates may still achieve adequate

precision (e.g., coefficients of variation < 30%) but rely on a random selection to ensure a representative sample, as with any survey design. In the absence of representative samples, estimators of discarded catch will be biased.

Amendment 23 of the Northeast Multispecies (i.e., groundfish) Fishery Management Plan (FMP) seeks to improve the reliability and accountability of catch reporting in the commercial groundfish industry. Current monitoring approaches fail to adequately ensure compliance with regulations and may incentivize behavior that degrades the accuracy of catch estimation. There were differences in vessel fishing activity related to the presence/absence of a human observer on a given trip (i.e., observer bias). Electronic monitoring (EM) systems provide opportunities for unbiased estimation of catch by combining 100% video surveillance of vessel fishing effort with a random selection of video reviewing, removing the opportunity for observer bias. Pilot programs in the Northeast have suggested that consistent catch handling and self-reporting by vessels can allow for low (~20%) video review rates, enabling a framework of monitoring that achieves the intended goals and objectives of Amendment 23.

To that end, an analysis was conducted that utilized both Vessel Trip Reports (VTRs) with EM video review to generate accurate estimation of groundfish discards on EM vessels. The analysis used several modeling approaches. Because of the cost of reviewing all of the EM video records, a key question is: can sufficient precision and accuracy be achieved by only reviewing a portion of the available EM records? And, if so, how large would that portion have to be? The analysis being reviewed here addresses the options for modeling approaches and the extent of EM reviews that are needed.

Description of the Individual Reviewer's Role in the Review Activities

The activities for this review consisted of an evaluation of the report: *Linden, D.W. 2021. A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels. Working Paper*, utilizing the supporting documents (Appendix 1) and a one-hour webinar on Sept 9, 2021 introducing the subject to the three independent reviewers (myself and two others). The webinar provided the reviewers opportunity to ask basic questions about the structure of the "experiment" from which the EM and VTR data came and some basic modeling questions.

Each individual reviewer then individually provided a review (this report is my review) which addresses the terms of reference listed in Appendix 3 in the following section.

Summary of Findings for each Term of Reference

The data used were from a pilot program in which electronic monitoring was voluntarily chosen by vessels in lieu of at-sea-observers for the years 2017-2020. According to information provided in the webinar, video cameras were placed on each vessel to monitor general activities. More specifically, after a haul, the crew would place fish to be discarded in a bin or bucket and then individually place a fish on a measuring board which was recorded by video monitoring. Upon the end of the trip the video records were reviewed and from those reviews the size and species of each fish were recorded (12 species were defined). No information was provided in the report to indicate the amount of time that was required for the crew to process the discarded catch, nor for the process of reviewing the videos of the individual

hauls. Amendment 23 provides some information on the video review process. Clearly, it is extensive and the implication is that it is expensive. Thus, there is motivation for sub-sampling the videos.

For this analysis there were data from 2,593 hauls across 801 sub-trips by 23 vessels. More than half of the effort came from bottom otter trawls (59%), a quarter from sink gillnets (25%) and the remaining effort was split between bottom longlines (8%) and handlines (8%). Effort from participants was variable across the 4 fishing seasons (1 May to 30 April), and review rates decreased from 100% to 50% of trips in 2019, resulting in the most sub-trips and hauls coming from the 2018 season (Table 1).

Table 1: Effort from groundfish EM vessels that was fully video reviewed during fishing years 2017–2020, with resulting total quantities (mt) of discards reported by video review and Vessel Trip Reports.

Fishing Yr	#vessels	#subtrips	# hauls	EM mt	VTR mt
2017	11	173	565	14.35	11.67
2018	15	298	865	12.29	10.15
2019	15	143	486	7.19	6.82
2020	14	187	677	11.11	10.47

Note: subtrips were defined as portions of a trip corresponding to specific areas; also vessel trip reports are essentially logbooks which are required from every trip, with or without at-sea-observers or electronic monitoring.

Judging from Table 1, I estimate that the amount of discarded catch is averaging around 50 lbs per haul which is divided amongst the 13 species. Thus, I expect that that the average discards of a species per haul probably doesn't get much larger than 10 lbs.

The models presented "predict" the electronic monitoring catch in weight by species of a haul as a function of the VTR reported catch for a haul. The observed dependent variable is the EM reviewed catch. Since EM reviewed catch is assumed to be "truth" (and the report indicated there was evidence of this), there is no reason to predict the reviewed EM catch *per se*. But the models are developed to predict discarded catch when EM-video catch records have been collected but have not been reviewed. A better characterization is that the models predict discarded catch as a function of the VTR reporting. The obvious incentive here is that the EM records exist, so there would be motivation for the VTR records to not deviate too much from EM because at some point that deviation could be found out.

Three models evaluated in the Linden 2021 were: 1) the delta-model (Thorsen version); 2) Zero-Inflated Poisson (ZIP); and 3) Poisson. The models utilized species intercepts, year effects and vessel effects. The model fit results are shown in Table 2.

Given these results (the simplicity of the Poisson and good R^2's), the Poisson was further evaluated with <100% EM-reviews using a) species and vessel factors; and b) species factors alone.

All of the models are predicting catch in weight. This causes some concern in that the Poisson and ZIP models are defined for integer observations and the pounds caught were converted to integers. However, in the webinar I asked "why are fish discarded"? the answer was that primarily they were undersized fish. If this is the case, then one would expect that the sizes of fish of a species that were undersized would not differ very much within a species. Thus, a poundage of the fish would be expected

to be reasonably proportional to numbers. However, if discarding is done for larger fish as when quotas have been reached, then I'd expect this relationship to deteriorate, and the efficacy of the Poisson models may deteriorate as well

Table 2: Measures of model fit, including Bayesian and Leave-one-out (LOO) adjusted R2 for each of the 3 models fit to groundfish discard data from EM vessels during 2017–2020.

Model	R^2 type	Estimate	Error	Q2.5	Q97.5
delta	Bayesian	0.7141	0.0296	0.6551	0.7648
delta	LOO	0.5782	0.0589	0.4462	0.6759
ZIP	Bayesian	0.8728	0.0032	0.8657	0.8783
ZIP	LOO	0.8768	0.0180	0.8359	0.9072
Poisson	Bayesian	0.9121	0.0007	0.9106	0.9134
Poisson	LOO	0.8967	0.0178	0.8570	0.9272

This also brings up the issue (or question) of what are the discard estimates used for. The preamble to the tasking order mentions stock assessment and fisheries management. But the webinar discussion and the gist of the Linden report indicate that the emphasis is on within-season monitoring. I expect that an objective of getting annual (or quarterly) discarded catch by species by size for use in stock assessment would result in a different modeling approach than Linden. Another objective is compliance, whether a vessel's VTR is "reasonably" close to the EM-review. Modeling could be used to define outliers and thus areas for further compliance investigation. So, there are a lot of uses of "discard estimates" and a need for having a better understanding of the likely uses of the estimates.

Another question that arises is why are the R^2 values of all of the models relatively high. It should be recalled that the data are from a fairly controlled "experiment" where the participation was voluntary as was the crew's time/effort to process the discards under the video. Also, my experience looking at observer versus logbook data is that often logbooks are reasonably accurate for the species that are important (barring compliance issues), but less so for species that are not. It is interesting that the "best" model, the Poisson, simply assumes standard Poisson patchiness of the density of the fish. Perhaps this is because the areas visited within a trip are not too dispersed. Thus, local Poisson density is seen.

Given these general comments, I will address the specific terms of reference:

Do the statistical analyses address the objectives of providing reasonably accurate and precise estimation of groundfish discards on EM vessels?
This term of reference is worded vaguely. You don't need any statistical analyses to get very precise and very accurate estimates of discards on EM vessels. You just have to look at the EM data! What is desired here are statistical models that predict with useful precision and accuracy the weight of discarded fish on EM vessels without looking at all the EM data. The author appears to have fulfilled that objective. Without defining "reasonably", the author provided a range of EM-review rates and the resulting CV's, accuracy and precision of vessel related subtrips. The definition of "reasonably" is left to others.

Are the statistical models adequately described?
The author was often cryptic in defining data, symbols and objectives. Nevertheless, I believe that I have interpreted the structure of the models and the data sources adequately enough to

provide useful insight into the process. However, better descriptions would be helpful in communicating to the Council, the SSC and the lay public.

Do the statistical models adequately fit the data?
The R^2 values in Table 2 above indicate the degree of fit. The ZIP and Poisson models appear to have useful fits in predicting discarded catch per haul of the volunteer vessels in this study.

Is the complexity of the statistical models justified?
In order to provide some basis for this answer, I looked at a simple Poisson process with varying mean catch rates per haul (mu) and then calculated the CV per trip for varying EM review rates (R). CV=SQRT[(1-R)/((hauls/trip)*mu)]. I used a constant 5 hauls per trip in this example. These results are compared to Figure 11 of the Linden Report (Figure 1 in this report):

Figure 1: CVs of Poisson Model by MU Value and Year
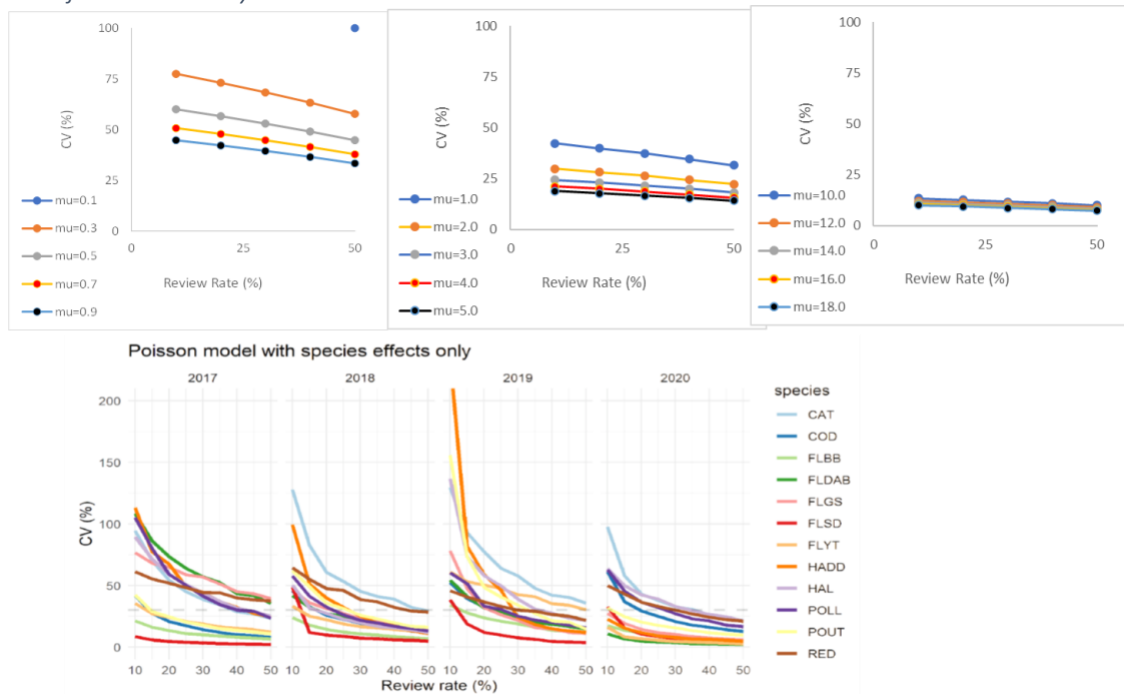
By this comparison I am trying to get at the question of what would the CV's be if no modeling were done and that a simple random sample of trips were used for the estimate and VTR data were ignored. I assumed a perfect Poisson process for each haul, where mu=sigma^2 and sigma/mu =1/sqrt(mu).

The results above indicate that the CVs of Figure 11 generally perform better than the simple Poisson assumption for the likely range of densities, especially for lower review rates. Also, note in the figure above, my simple Poisson assumptions imply that the relationship is near linear over the relevant R range. But the slope of the CV wrt R is more negative with larger R. Whereas, in the Linden relationship it is less negative. This is evidence that justifies the modeling. I suppose that a better stratified random sampling strategy of trips and/or hauls might improve my simple Poisson process. But nevertheless, at this stage the complexity of the statistical modeling is justified.

Comparing Linden's Poisson with species/vessel with Poisson species only (fig 10 vs Fig 11), the latter provides a better CV vs review rate relationship. However, intuitively it seems that including vessel effects would be more realistic. The assumption is that within a strata, the density of a species is the same. Thus, differences in vessel effects are essentially differences in catchability which is a manifestation of nuances of gear, fishing strategies, etc. It is hard for me to argue that these should be ignored.

Are there alternative models that would be preferable in terms of estimation performance and computational efficiency?
See the above section on model complexity.

In terms of estimation performance, there is always the possibility that some alternative model might be defined which could perform better. But at this point for the limited objectives of the Linden Report, there are not obvious better alternatives.

The author notes that the models were fit using Markov-chain Monte Carlo methods with Bayesian R^2 values using default vague prior distributions and 3 chains of 2,000 iterations, which was sufficient for model convergence. Bayesian fitting methods are generally more computationally intensive, but no information was provided about the amount of computational time needed to get results. From an analyst's viewpoint, computational efficiency and less computing time is desired. But it is doubtful in my opinion that computational time will become an issue in any implementation of these models within a decision-making framework.

Have the sources of uncertainty and caveats in the analyses been adequately described? Given the objectives and the performance of presented models, are there additional potential sources of uncertainty that can be quantified and should be incorporated?

The basic approach for all three models is that there is a common density of the individual species that does not change within the year and that changes in individual growth information is not important in the overall analysis. Thus, the modeling is essentially predicting changes in catchability reflected by vessel characteristics and other data. However, fish grow and die during the year. It is unclear how these factors affect the outcomes and whether time-dependent size and density factors could or should be included.

The analysis does not discuss the reason for discards and the use of discard data in the assessment/monitoring/compliance continuum. In the webinar, the information that was given was that most discards were for undersized fish. But other places in the document mention quota monitoring and implied compliance. In my opinion the most appropriate approach for modeling may well change depending on which of the assessment/monitoring/compliance objectives is being addressed. Thus, in some instances (see previous paragraph) within year issues may be important.

I am unaware of the details of how at-sea-observer data are currently being used to estimate discards of unobserved trips in GARFO. I know that for some species that occur in the northwest Atlantic, GLM approaches are used to estimate discards. It is unclear to me if the ATO data might be utilized in conjunction with EMs where a subset of EM and ATO vessel data are sufficiently related.

There are lots of details about how an EM data collection process might work that affect how the data are collected. The most obvious is how much effort/time the crew has to do to process discards and who

pays for that time/effort, and how much time/effort and who pays for EM reviews. Also, if the EM data are to be used for within-season monitoring, how quickly can they be turned around and at what frequency?

I can foresee a number of regulatory issues. For example, the government will be the repository of a lot of unreviewed-EM data. These are data that are available and have been collected at someone's' (whose?) expense. If there is some critical management issue that complete review of the EM data will help to resolve, then the government could become legally vulnerable for not using it and, thus, would be required to pay for further EM reviews. Similarly, an individual vessel might want to petition the government to review its EM trip, which could delay the management process. Fairness and equity would have to be adequately defined in choices of the mix of EM/ATO/VTR only and how those choices are made. These things are, of course, outside of the scope of the present analysis. However, they are real questions that future analyses will probably have to address.

Conclusions and Recommendations

The scope of the Linden document is rather limited: examination of the efficacy of three models in predicting discarded catch of EM-observed vessels on a haul based on VTR data when the EM data have not been reviewed. The implied goal is to avoid the cost of reviewing all of the EM data in future implementations. The scope is also limited in that the data were derived from volunteer cooperating vessels without immediate management consequences. The author has provided a useful template to address that objective. However, this analysis is only a part of the pursuit of larger goals of implementing a data collection program to obtaining precise, accurate discard estimates for use in stock assessment, monitoring and compliance.

The author has answered the basic question: yes, one can get "reasonable" precision and accuracy in predictions of a trip's EM observed vessel discarded catch by species by only looking at a portion of the EM data by relating it to VTR data from the controlled experiment that existed 2017-2020.

However, while scientific review of any analysis is beneficial to the scientific process, in some ways I think this review is premature. Any real-world implementation will subtly and not so subtly change fisher behavior and incentives. Those changes are likely to alter the modeling approaches that are most appropriate. In other words, subtle changes in implementation could mean that the Linden models would no longer be appropriate.

Nevertheless, the approach one takes in these types of regulatory analyses is to try to establish the best that can be achieved with near-perfect conditions (as with the voluntary 2017-2020 data) and then try to quantify the losses to precision and accuracy using various compromises in the implementation. The Linden Report helps to establish the baseline for these comparisons.

There are a number of possible analyses which were not pursued in the document. In fairness they were outside the scope of the given objectives of the report. But they are clearly within the objectives of Amendment 23. Nevertheless, they would have helped the progress in understanding the implications of various monitoring programs. If this review were a typical CIE review where several days were spent with the analyst (even in a webinar, rather than in-person), then a number of these options would have been explored and more progress would have been made.

Therefore, recommended analyses in the future should include:

- Cumulative catch by species within a year for all EM vessels compared with estimates and variances of EM vessels with ranges of review rates (addressing within-year monitoring of catch).

- Discussion/analysis on the scaling up of these results to the total fleet, addressing things like how many EM vessels are needed before overall annual catch precision is "acceptable". And the balance of using EM reviewed vessels, EM vessels which are not reviewed but predicted based on VTR, at-sea-observed vessels with VTR and no EM and no at-sea-observed vessels with VTR only.

- Discussion/analysis on modeling approaches when discards that are being estimated are those that are discarded after a TAC has been reach or for other than minimum size regulatory purposes.

- Discussion/analysis of different modeling approaches that might be used for 1) stock assessment; 2) within season monitoring; 3) post closure monitoring; 4) compliance of individual vessels.

- Explore the use of size categories and within-year time strata in the models.

Appendix 1: Bibliography of materials provided for the review

The document to be reviewed

Linden, D.W. 2021. A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels. Working Paper

Background documents provided

Gelman, Andrew, Ben Goodrich, Jonah Gabry, and Aki Vehtari. 2019. "R-Squared for Bayesian Regression Models." The American Statistician 73: 307–9.

New England Fishery Management Council. 2021. Northeast Multispecies Fishery Management Plan – Amendment 23. Including a Final Environmental Impact Statement and Initial Regulatory Flexibility Analysis. August 2021 Vols 1 and 2. https://www.nefmc.org/library/amendment-23

Thorson, J. T. (2018). Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. Canadian Journal of Fisheries and Aquatic Sciences, 75(9), 1369-1382.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. "Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC." Statistics and Computing 27 (5): 1413–32

Venables, W. N., & Dichmont, C. M. (2004). GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. Fisheries research, 70(2-3), 319-337

Appendix 2: CIE Performance Work Statement

Performance Work Statement
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review

A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels

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[1]. Further information on the Center for Independent Experts (CIE) program may be obtained from www.ciereviews.org.

Scope

Estimation of total catch in a commercial fishery is important for stock assessment and effective management decision-making, as dictated by the Magnuson-Stevens Act. A primary challenge of catch estimation lies in the largely unknown portion of catch that is discarded at sea rather than retained. Catch can be discarded for multiple reasons related to regulatory compliance (e.g., species and/or size retention prohibited) and non-compliance (e.g., illegal actions incentivized by quota limitations). While the retained portion is recorded by multiple information streams, discard estimation typically relies on a sampling of trips that carry a human at-sea observer. Low sampling rates may still achieve adequate precision (e.g., coefficients of variation < 30%) but rely on a random selection to ensure a representative sample, as with any survey design. In the absence of representative samples, estimators of discarded catch will be biased.

Amendment 23 of the Northeast Multispecies (i.e., groundfish) Fishery Management Plan (FMP) seeks to improve the reliability and accountability of catch reporting in the commercial groundfish industry.

---

[1] https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2005/m05-03.pdf

Justification for the amendment came from evidence that current monitoring approaches fail to adequately ensure compliance with regulations and may incentivize behavior that degrades the accuracy of catch estimation; evidence included differences in vessel fishing activity related to the presence/absence of a human observer on a given trip (i.e., observer bias). The proposed solutions involve increased and/or enhanced monitoring that maximizes the value of collected information while minimizing costs to the industry and NMFS. Electronic monitoring (EM) systems provide opportunities for unbiased estimation of catch by combining 100% video surveillance of vessel fishing effort with a random selection of video reviewing, removing the opportunity for observer bias. Pilot programs in the Northeast have suggested that consistent catch handling and self-reporting by vessels can allow for low (~20%) video review rates, enabling a framework of monitoring that achieves the intended goals and objectives of Amendment 23.

Here, we illustrate a modeling approach that leverages Vessel Trip Reports (VTRs) with EM video review to generate accurate estimation of groundfish discards on EM vessels. We fit and compared multiple generalized linear mixed and delta/hurdle models to estimate the relationship between EM-reviewed discards and VTR-reported discards, $E(EM) \sim f(VTR + \ldots)$. The models accommodated nuances of the data (e.g., zero inflation) and variation according to groundfish species, vessel/trip attributes, and changes in data availability during the fishing season. We used data from 2017–2021 on EM trips that were fully reviewed including 31 vessels, >1,100 sub-trips, and >3,500 hauls. The data were subset to explore review rates ranging towards 1% and to identify thresholds of reasonable precision, which depended on the model structure used. This modeling framework will be used for groundfish EM vessels during FY2021 and beyond.

Given the implications of this new monitoring approach, it is important that the methods represent the best available science and are statistically sound. Therefore, the CIE reviewers will conduct a peer review of the statistical modeling based on the Terms of Reference (TORs) referenced below. Given the public interest, it will be important for NMFS to have a transparent and independent review process of the model used in this assessment.

The specified format and contents of the individual peer review reports are found in Annex 1. The Terms of Reference (TORs) of the peer review are listed in Annex 2.

Requirements
NMFS requires three reviewers to conduct an impartial and independent peer review in accordance with this Performance Work Statement (PWS), OMB Guidelines, and the ToRs below. The reviewers shall have working knowledge and recent experience in **statistical modeling, with applications to fisheries management and/or quantitative ecology**. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Tasks for reviewers
Each CIE reviewers shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

1. Pre-review Background Documents: Review the following background materials and reports prior to the review:

Thorson, J. T. (2018). Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. *Canadian Journal of Fisheries and Aquatic Sciences*, 75(9), 1369-1382.

Venables, W. N., & Dichmont, C. M. (2004). GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research. *Fisheries research*, 70(2-3), 319-337.

Approximately, two weeks before the peer review, the NMFS Project Contacts will send by electronic mail or make available at an FTP site to the CIE reviewer all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contacts will consult with the CIE on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review.

2. Webinar: Additionally, approximately two weeks prior to the peer review, the CIE reviewers will participate in a webinar with the NMFS Project Contacts and other staff to address any clarifications that the reviewers may have regarding the ToRs or the review process. The NMFS Project Contacts will provide the information for the arrangements for this webinar.

3. Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and ToRs, and shall not serve in any other role unless specified herein. Modifications to the PWS and ToRs cannot be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor.

4. Contract Deliverables: Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the PWS. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Place of Performance
Each CIE reviewer shall conduct an independent peer review as a desk review; therefore, no travel is required.

Period of Performance
The period of performance shall be from the time of award through October 2021. Each reviewer's duties shall not exceed 10 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.

Schedule	Milestones and Deliverables
Within two weeks of award	Contractor selects and confirms reviewers
Two weeks prior to the review	Contractor provides the pre-review documents to the reviewers.

Within two weeks prior to the review	Reviewers participate in Webinar.
August 2021	Each reviewer conducts an independent peer review as a desk review
Within two weeks after review	Contractor receives draft reports
Within two weeks of receiving draft reports	Contractor submits final reports to the Government

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

Since this is a desk review travel is neither required nor authorized for this contract.

Restricted or Limited Use of Data

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

Project Contacts

Dr. Daniel Linden (lead contact)
NOAA/NMFS/GARFO, Analysis and Program Support Division
55 Great Republic Road
Gloucester, MA 01930
daniel.linden@noaa.gov
978-281-9265

Dr. J. Michael Lanning
NOAA/NMFS/GARFO, Analysis and Program Support Division
55 Great Republic Road
Gloucester, MA 01930
j.michael.lanning@noaa.gov
978-281-9308

Annex 1: 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 main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role 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.

3. The reviewer report shall include the following appendices:
a. Appendix 1: Bibliography of materials provided for review
b. Appendix 2: A copy of the CIE Performance Work Statement

Annex 2: Terms of Reference for the Peer Review

The reviewers will provide a scientific peer-review of the following document:

Linden, D.W. 2021. A predictive model of discarded catch that leverages self-reporting and electronic monitoring on commercial fishing vessels. *Working Paper*

The reviewers will provide input on the following questions:

1. Do the statistical analyses address the objectives of providing reasonably accurate and precise estimation of groundfish discards on EM vessels? a. Are the statistical models adequately described?

b. Do the statistical models adequately fit the data?

c. Is the complexity of the statistical models justified?

d. Are there alternative models that would be preferable in terms of estimation performance and computational efficiency?

2. Have the sources of uncertainty and caveats in the analyses been adequately described? Given the objectives and the performance of presented models, are there additional potential sources of uncertainty that can be quantified and should be incorporated?