**Expert-driven testing and proposed improvements to a bycatch estimator toolkit**

Elizabeth A. Babcock[[1]](#footnote-1) , William J. Harford2, Ana Adao2, Todd Gedamke3,

SUMMARY

*Babcock (2022) has developed a toolkit that uses model-based and design-based procedures in a semi-automated process of estimating total annual bycatch by expanding the data from an observer program to the total effort from logbooks or landings records. As part of on-going efforts to improve the functionality and user-experience of this toolkit, an expert-driven hybrid workshop was conducted July 25th to 27th, 2023 in Miami, Florida, USA. The goal of this workshop was to allow experts familiar with bycatch data and statistical aspects of fisheries bycatch estimation to engage in ‘beta-testing’ of the BycatchEstimator R package. Workshop participants* *recommended changes to the workflow of the R package, allowing data setup, design-based estimators, and model-based estimators to be separate steps to improve user experience and to maintain scientific rigor throughout the analysis. Participants also contributed a comprehensive list of user experience and technical recommendations that will enable this toolkit to become more widely accessible to users and comprise a deeper set of statistical methods and diagnostics for bycatch estimation.*

*KEYWORDS*

*Bycatch, catch statistics, simulation, model testing, LLSIM, blue marlin*

1. **Introduction**

Babcock (2022) has developed a toolkit that uses model-based and design-based procedures in a semi-automated process of estimating total annual bycatch by expanding the data from an observer program to the total effort from logbooks or landings records. In addition to bycatch estimation, this toolkit can also be used to estimate an annual index of abundance, calculated only from the observer data. This toolkit has previously been subject to simulation testing and application to actual data to examine its efficacy and its utility for use for bycatch estimation in ICCAT CPC fisheries (Babcock and Goodyear 2021; Babcock et al. 2022; Babcock et al. 2023).

As part of on-going efforts to improve the functionality and user-experience of this toolkit, an expert-driven hybrid workshop was conducted July 25th to 27th, 2023 in Miami, Florida, USA. The goal of this workshop was to allow experts familiar with bycatch data and statistical aspects of fisheries bycatch estimation to engage in ‘beta-testing’ of the BycatchEstimator R package (Babcock 2022). This report summarizes the set of recommendations provided by participants as it relates to training and user guide materials, technical improvements to the toolkit, and future directions for training and capacity building.

**2. Methods**

***2.1 Bycatch estimator toolkit***

The bycatch estimator toolkit is a statistical framework for bycatch estimation that has been developed as a freely accessible R package under the title BycatchEstimator (Babcock 2022). Model-based bycatch estimation uses generalized linear models (GLM) based on the user’s choice of observation error models (e.g., delta-lognormal, negative binomial) and predictor variables (e.g., year, season, depth). Information criteria (AICc, AIC or BIC) can be used to find the best set of predictor variables, and cross-validation can be used to compare observation error models (e.g., negative binomial, delta lognormal). The selected GLMs are used to predict total bycatch in all logbook trips (or only unsampled trips, if desired) and total bycatch is estimated by summing across trips. The design-based methods include a stratified ratio estimator, and the delta-lognormal estimator of Pennington (1983), and the user may specify the stratification variables for these estimators (e.g., seasons, spatial areas). For the design-based estimators, if any strata have less than a user-specified number of observations (e.g., sets), estimates for those strata are made by pooling across the user’s choice of stratification variables. Pooling can be done either across adjacent years, or by defining a more aggregated variable to be used in pooling (e.g., seasons rather than months), or by pooling across all levels of a variable. A recent exploration of plausible extensions of this toolkit included a geostatistical model, showed that total bycatch estimates were much more precise when spatial and/or spatiotemporal random effects were included in the model (Babcock et al. 2023).

***2.2 Workshop structure and training materials***

The workshop was held at The Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami from July 25th to 27th, 2023. Using a hybrid approach, 12 subject matter experts participated online or in person. Led by Dr. Babcock, the workshop consisted of presentation materials, R-based coding sessions, and group-based discussions. Participants focused on improvement of scientific methodologies for bycatch estimation. During the three-day workshop, a preliminary set of training materials were presented to participants that covered a diverse set of topics, including design-based estimators, model-based estimators, diagnostics, CPUE standardization, and various considerations related to data cleaning and formatting. Training sessions were supplemented by discussions where participants proposed ways in which to improve and advance the use of the software package.

**3. Results**

***3.1 Workflow and user experience improvements***

Participants recommended changes to the workflow of the R package (**Figure 1**), where bycatch estimation should proceed in three steps. Step one ‘bycatch setup’, was proposed to encompass the tasks of data formatting, data summaries, warnings and diagnostics allowing for quality assurance and quality control of data inputs, as well as verification of inputs in a similar manner to that of ‘echo input’ used in stock assessment software Stock Synthesis. This step is intended as a precursor to saving a verified set of inputs for subsequent use in analysis. Step two ‘design-based estimators’, was proposed to isolate functions and analyses based on design-based estimators to compartmentalize this set of methodologies and to ensure ease of use. Step three ‘model-based estimators’ was proposed to isolate functions and analyses based on model-based estimators to compartmentalize this set of methodologies and diagnostics and other related model-fitting facets of bycatch estimation and CPUE standardization.

Workflow improvements were also recommended for outputs of the analysis. Principally, participants requested organization of outputs into a structured system of folders as well as uniformity in naming conventions, all aimed at improving the user experience. Minor recommendations included time-stamped file names, improved section headings and descriptions of automated report contents, and default use of HTML-based outputs (and optional pdf output) to avoid technical complications associated with generating pdf-based work products.

Finally, participants noted the importance of updates to the user guide and training materials to improve user experience. Recommendations tended to address pain points related to bycatch setup and various guidance needed before proceeding to analysis. Technical recommendations are addressed in the next section of the report. Recommendations for bycatch setup and guidance needed before proceeding to analysis were as follows:

* Accompanying the modified workflow, participants requested that additional guidance be provided on checks and data validation, including automated warnings, summary tables, and simple predictive estimates of target catches (e.g., to address scaling challenges).
* Develop additional explanations and examples regarding matching of observation units (e.g., trips or sets) between logbook and observer data. This can be a particularly challenging task in many fisheries.
* Provide improved explanations of observation error structure on catch units (e.g., catch in weight is reported as an integer in negative binomial, so units and scale is important) and expand upon background material regarding theory and practice of selecting sampling distributions. This could also be improved through the addition of visualizations of data distributions to BycatchSetup.
* Add R script examples to demonstrate code syntax for glm equations (e.g., y~Year).
* Provide Excel-based templates by set and by trip, for observer and logbook and supplement with guidance for making decisions about data formatting (e.g., missing data, outliers, missing value filling, etc.).
* Provide guidance on working with rare species and/or low sample sizes (e.g., how to combine years into blocks, etc.).
* Provide advice on how to respond to violations of assumptions (e.g., outliers, dispersion, zero inflation) and NAs or bad data. For example, how to impute missing data in logbooks (which must be done because logbooks produce total effort, missing observer data can just be excluded).
* Provide guidance on choice of sample unit (sets, trips, aggregated) and other sample space issues (e.g., vessels that fish in multiple gears/fisheries with potential overlapping between observer programs).
* Develop improved explanations of pooling for design-based estimators, including how to parameterize pooling, perhaps with examples.
* Provide clarity on the importance of the order of variable specification, especially as it pertains to pooling for design-based estimators.
* Add additional background material regarding theory and practice of information theoretic approaches, including citing background reading, common procedural practices, and on-going scientific debates, as variable selection is semi-automated but requires user interpretation. Further, additional guidance is needed on the use of dredgeCrossValidation, including when it is advisable to revisit variable selection during validation of observation error models.

***3.2 Technical improvements***

The authors have maintained a comprehensive list of technical coding improvements and functionality recommendations provided by workshop participants. For brevity, a summary of these items is provided here. Workshop participants coalesced around the following thematic areas of improvement to the toolkit:

* **Bycatch setup:** prior to proceeding to analysis, participants recommended a suite of diagnostics and summary plots. These summaries include a host of plots (e.g., histograms, box plots, violin plots, bar plots) aimed at understanding data distributions, potential outliers, sample sizes by factor level and/or explanatory variables, comparisons between logbook and observer program data quantity by factor level and/or explanatory variables, and spatial coverage patterns in cases where latitude and longitude are provided. Related summaries are needed in the form of time series plots. Additionally, output should contain automated warnings and other simple diagnostics before proceeding to analysis (e.g., NAs and missing data, factor level sample size comparisons between logbook and observer data sets, data-coverage by strata).
* **Design-based estimators:** participants emphasized the need for comprehensive pre-analysis data validation. This could include a system of checks and warnings as it relates to missing data, sample size considerations, strata coverage, etc. Additionally, outputs of design-based estimators should include a variety of confidence interval types (e.g., 95% and 80%).
* **Model-based estimators:** like the design-based estimators, recommendations for model-based estimators emphasized the need for comprehensive pre-analysis data validation, but additionally emphasized post-analysis diagnostics of model fit. For example, participants recommended all modeling outputs should be retained as an R-based object, allow users to access the finer details of model fits. This could include model-specific diagnostics and parameters (e.g., dispersion parameters from negative binomial and Tweedie). Further, models that are considered to fail should be retained along with warnings and errors (e.g., convergence issues, failure messages), thus providing information about why these models were excluded from the analysis. Additional metrics to be included in model output summaries include R-squared and adjusted R-squared, covariance between explanatory variables, variable coefficient values, and CPUE standardization residuals by factor-level. Finally, predicted catch by logbook entry should be reported, allowing for subsequent aggregate summaries for variables not included in the model, such as ICCAT areas.

**4. Discussion**

Through a combination of training materials and beta-testing, this expert-driven workshop produced a comprehensive set of recommendations for improvements to the bycatch estimator toolkit. Beyond an immediate set of improvements, participants also identified several longer-term proposals. These proposals included inclusion of an even broader set of observation error models (e.g., Poisson, zero-inflated models), as well as requests to explore the feasibility of generalized additive models and to continue work on the feasibility of spatial modeling (e.g., see Babcock et al. 2023). There was also interest in simple estimators based on ratios of target to bycatch species, in instances where logbook effort is unavailable, with suggestion that such approaches could be subject to reliability testing via simulations. Finally, analysis of ‘low data fleets’ could potentially be improved through meta-analysis or joint analysis of bycatch across fleets (e.g., bycatch estimation for billfish and shark species).

Participants also highlighted the need to address uncertainty, in terms of defensible application of statistical methods. Discussions related to statistical methods highlighted challenges in estimating variance related to both inputs and outputs to the analysis. Regarding uncertainty in inputs, it is not yet possible to input variances for total effort; however, this is being developed using Bayesian methods in a related project (Babcock et al. 2018). Regarding uncertainty in outputs (i.e., variance of bycatch estimates), estimates of uncertainty in bycatch can be useful, especially for assessment models that can incorporate input variances for removals.

Given the ability to estimate bycatch, including quantification of confidence intervals, additional challenges were highlighted in the context of using this information in reporting and decision-making. For example, bycatch estimates are obtained for each year in a time series of logbook and observer program data sets. With the annual addition of new data and re-estimation (much like biomass estimation in stock assessment), historical bycatch estimates could potentially change o­­ver time. It remains unclear how this circumstance should be addressed in decision-making. A related recommendation was to develop functionality for conducting ‘retrospective analyses’ to help diagnose the extent to which recent bycatch estimates may be considered unstable, for example, in instances where recent retention bans or changes in gear have taken place.

In concluding their discussions about next steps in advancing this toolkit, participants also discussed expanded training and capacity building opportunities. In addition to engaging with a wider audience for training on the bycatch estimator toolkit, participants suggested that the preliminary training materials be formalized into a set of self-guided tutorials to be made available online. Participants highlighted the need for both expansion of hands-on training, including application to actual data sets, as well as refinement of online reference materials.

**5. Acknowledgements**

This work was conducted within the ICCAT SCRS annual workplan with financial support from the ICCAT Science Budget and the European Union through the EU Grant Agreement No. EMFAF-2021-VC-ICCAT5-IBA-02 - Strengthening the scientific basis for decision-making in ICCAT. E. A. Babcock’s work was also partly supported by NOAA, via the Cooperative Institute for Marine and Atmospheric Studies.

**6. References**

Babcock, E.A., W.J. Harford, T. Gedamke, S. Anderson, C. P. Goodyear. 2023. Simulation-testing model-based and design-based bycatch estimators. Collect. Vol. Sci. Pap. ICCAT, 80(6): 51-79.

Babcock, E.A. 2022. Bycatch Estimator R library. <https://ebabcock.github.io/BycatchEstimator/>

Babcock, E.A., W.J. Harford, T. Gedamke, D. Soto, and C. P. Goodyear. 2022. Efficacy of a bycatch estimation tool. Collect. Vol. Sci. Pap. ICCAT, 79(5): 304-339.

Babcock, E.A. and C. P. Goodyear. 2021. Testing a bycatch estimation tool using simulated blue marlin longline data. ICCAT Collect. Vol. Sci. Pap. ICCAT, 78(5): 179-189.

Babcock, E.A., M. Barnette, J. Bohnsack, J.J. Isely, C. Porch, P. M. Richards, C. Sasso, and X.Zhang. 2018.Integrated Bayesian models to estimate bycatch of sea turtles in the Gulf of Mexico and southeastern U.S. Atlantic coast shrimp otter trawl fishery. NOAA Technical Memorandum NMFS-SEFSC-721. 47 pp.

Pennington, M. 1983. Efficient Estimators of Abundance, for Fish and Plankton Surveys. Biometrics 39 (1): 281–86. <https://www.jstor.org/stable/2530830>.



**Figure 1.** Current workflow of the bycatchEstimator R package (Babcock 2022).

1. Department of Marine Biology and Ecology, Rosenstiel School of Marine, Atmospheric and Earth Science, University of Miami, 4600 Rickenbacker Cswy., Miami, Florida, 33149, USA. ebabcock@miami.edu

   2 Nature Analytics, 551 Lakeshore Rd E, Suite 105, Mississauga, Ontario, L5G 0A8, Canada. bill@natureanalytics.ca

   3 MER Consultants, 5521 SE Nassau Terrace, Stuart, Florida, 34997, USA. todd@merconsultants.org [↑](#footnote-ref-1)