# *ObsCovgTools*: Assessing observer coverage needed to document and estimate rare event bycatch

K. Alexandra Curtis[a,b*] and James V. Carretta[b]

[a] Contractor with Ocean Associates, Inc., for Marine Mammal and Turtle Division, Southwest Fisheries Science Center, National Marine Fisheries Service, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA; Email: alex.curtis@noaa.gov

[b] Marine Mammal and Turtle Division, Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 8901 La Jolla Shores Drive, La Jolla, CA 92037, USA; Email: jim.carretta@noaa.gov

* Corresponding author

E-mail: alex.curtis@noaa.gov

Phone: +1-858-546-5625

Fax: +1-858-546-7003

# ABSTRACT

Observer program design and evaluation often overlook the challenges of documenting rare-event bycatch. To support and facilitate consideration of threatened, endangered, and protected species bycatch in evaluating observer programs and assessing fisheries impacts, we developed a software tool to assess observer coverage with respect to several objectives for documenting or estimating rare-event bycatch. The *ObsCovgTools* package for the *R* programming language, also available as an online application, predicts observer coverage performance for a given total fishery effort in relation to three metrics: (1) the conditional probability of observing any bycatch given that bycatch occurred in the fishery and the probability of any bycatch in the total fishery effort, (2) the upper confidence limit for total bycatch when none is observed, and (3) precision (coefficient of variation) of the bycatch estimate. We describe the tool; explore how specific observer coverage targets for these metrics vary with total effort, BPUE, and dispersion index; and apply it to evaluate observer coverage in the California drift gillnet fishery. Our results underscore the importance of considering effort as well as percentage in assessing how well an observer program documents bycatch. We caution that rare species interactions may not be documented in many observer programs, and should be anticipated through a complementary risk assessment approach. The tool's modular design and open source programming approach encourage adaptation and augmentation to address additional objectives or complexities in sampling design or estimation.

**Keywords:**

Bycatch estimation; fishery observer program; limit reference point; precision; risk assessment; software

## 1. Introduction

Effective bycatch management requires that bycatch be detected, identified, and monitored through a fisheries monitoring system to inform assessment, prioritization, and management action (Crowder and Murawski 1998; Hall and Mainprize 2005; Kirby and Ward 2014). At-sea observer programs remain the gold standard for obtaining independent, accurate, and verifiable scientific data on fishing operations, particularly for less frequent bycatch species that are discarded, including endangered or threatened species and other sensitive marine fauna (e.g., Karp and McElderry 1999, Davies and Reynolds 2002; Pérez Roda et al., 2019). Observer programs are most efficient and effective when they are designed to address clearly defined, specific goals and objectives (Davies and Reynolds 2002; Parkes and Kaiser 2004; Kirby and Ward 2014). In the case of bycatch monitoring, objectives should include characterizing bycatch composition and magnitude with sufficient accuracy and precision to support assessment and management (Crowder and Murawski 1998). Addressing these objectives requires designing observer programs with appropriate observer coverage for a given total effort (Hall 1999; Babcock et al., 2003; NMFS 2004; Parkes and Kaiser 2004).

The primary focus in evaluating observer coverage levels has been on the second objective – providing data to estimate bycatch magnitude with sufficient accuracy and precision, with an emphasis on precision (e.g., Lennert-Cody 2001; Bravington et al., 2003; Smith and Baird 2005; Lawson 2006). Bycatch estimates that involve rare events (the norm for threatened, endangered, and protected species) and low observer coverage levels suffer from estimation bias and imprecision (Amandè et al., 2012; Carretta and Moore 2014; Martin et al., 2015). The problem of achieving reasonable precision for rare-event bycatch has also motivated considerable effort towards developing model-based approaches to estimating bycatch that harness information in covariates of bycatch rate to reduce uncertainty (Dixon et al., 2005) and propose different parameterizations or mixed distributions to handle large numbers of zeroes (see Minami et al., 2007 for an overview of these, including hurdle and zero-inflated models). Development of model-based bycatch estimates may inform observer sampling design by guiding stratification to direct increased sampling towards those portions of the fishery where interactions are most likely, particularly for a high-priority species of conservation concern (e.g., Federal Register 2013; Carretta et al., 2017).

Far less attention has been directed towards evaluating observer programs with respect to the first objective of characterizing bycatch composition accurately (see Lyssikatos and Garrison 2018 for a well-executed example). Species that have not been documented in a fishery, whether in a given year or over the history of the fishery, are often assumed not to have been subject to any interactions during that time. Estimates drawing on multiple years of observer data, such as model-based estimates, may alleviate this problem where bycatch has been observed one or more times in the past (Carretta and Moore 2014), but fisheries with extremely low observer coverage may miss rare-event fisheries interactions with species of conservation concern for years or even decades. A thorough bycatch risk assessment should be conducted for every fishery to identify species that are at risk of interacting with it  in terms of geographic, depth, and habitat overlap and susceptibility to gear type and identify populations on which fisheries interactions may have a non-negligible impact (Hobday et al., 2011).

Many fisheries observer programs face challenges that call for customized simulation approaches to sampling design, such as within-haul subsampling and variability (Karp and McElderry 1999) or the need to balance limited resources among a large number of fisheries (Wigley et al., 2007). Often, human resources or technical capacity may be lacking for customized analyses, a sampling problem may be simple enough for standard methods, or managers need more interactive information regarding how observer program performance for a given fishery is expected to vary with coverage. Rather than relying exclusively on *de novo* simulation or statistical approaches for each design question, standardized approaches should be made readily available as a starting point in assessing observer coverage needs.

Here, we present a user-friendly *R* package and shiny web application to facilitate design and evaluation of observer programs with respect to monitoring composition and magnitude of rare-event bycatch of sensitive species (Curtis 2019; Curtis and Coleman 2019). We describe the *ObsCovgTools* package, use it to explore how specific bycatch-oriented observer coverage targets vary with fishery size, bycatch rate, and dispersion index (variance to mean ratio), and demonstrate its application in a case study drawn from U.S. fisheries management.

## 2. Methods

### 2.1 Software description

The *ObsCovgTools* package for the *R* statistical programming language allows evaluation of observer coverage for a given total effort based on three metrics: (1) the probabilities of observing any bycatch and of any bycatch occurring in the total fishery effort, (2) the upper confidence limit of total bycatch that may have occurred without any being observed, and (3) the expected precision of bycatch estimates. For each metric, the user specifies total effort in the fishery and expected dispersion index (variance to mean ratio) of bycatch. For the observation probability and precision metrics, mean bycatch per unit effort (BPUE) is also user-specified. The user can also specify a desired benchmark, i.e., the desired probability of observing bycatch if it occurs, the maximum allowable upper confidence limit for bycatch when none is observed, or the desired estimation precision. For each metric, the package returns a plot of how it varies with observer coverage, and the minimum coverage corresponding to the specified benchmark, if applicable. Effort can be considered in arbitrary units and time frames relevant to management, e.g., trips or sets per year or per five years.

The package utilizes Poisson and negative binomial distributions, parameterized by user-specified BPUE and dispersion index, to simulate statistical properties of bycatch. Among the range of statistical distributions used in bycatch estimation and modeling, the Poisson and negative binomial distributions are widespread and easily parameterized given typically available summaries of bycatch data or estimates, and thus well-suited as a starting point for a user-friendly, general-purpose tool. Hurdle distributions, such as delta-lognormal, make the assumption that all zeros are structural, a poor approximation of a process producing rare events. Moreover, change in effort does not align consistently with change in any hurdle model parameter (Ancelet et al., 2010). Zero-inflated distributions require a zero-inflation parameter to differentiate structural from sampling zeros, and may not be identifiable for rare-event processes (Minami et al., 2007). The *ObsCovgTools* package can nonetheless be adapted to zero-inflated data by omitting effort believed to be associated with structural zeros. Lastly, it is worth noting that given fixed BPUE and dispersion index, negative binomial and quasi-Poisson distributions are one and the same.

For the first metric, probability of observing bycatch if it occurs, the package calculates the probability of observing any bycatch in a given amount of effort $n$ as

$$p_B = 1 - (p_0)^n \tag{1}$$

where $p_0$ is the probability of zero bycatch in a unit of effort. $p_0$ is given by the probability mass function (PMF) for the Poisson (dispersion index = 1) or negative binomial (dispersion index >1) specified by the user-input BPUE ($r$) and dispersion index ($d$) at n=1:

$$p_0 = \begin{cases} e^{-r}, d = 1 \\ d^{-r/(d-1)}, d > 1 \end{cases} \tag{2}$$

where the negative binomial PMF at 0 for n=1 is algebraically rearranged from a more typical formulation $(k/(k+r))^k$ (Hilborn and Mangel 1997) after substituting $k = r/(d\text{-}1)$. Since rare event bycatch may or may not occur in the total fishery effort in a particular time period, the package calculates the probability of observing any bycatch as conditional on the probability of bycatch occurring in the total effort, thereby incorporating both process and observation error.

For the second metric, the upper confidence limit of total bycatch when no bycatch was observed in $n$ effort, first the one-tailed upper confidence limit of $r$ at confidence level 1-α is calculated by setting $\alpha = (p_0)^n$ from Equation (2) and solving for $r$, which rearranges to

$$r = \begin{cases} -\frac{1}{n}\log\alpha, d = 1 \\ \frac{-(d-1)\log\alpha}{n\log d}, d > 1 \end{cases} \tag{3}$$

To calculate the upper confidence limit for total bycatch in the fishery ($B_{1\text{-}\alpha}$), $r$ is multiplied by total effort $N$ and the square root of the finite population correction from Cochran (1977):

$$B_{1-\alpha} = N\, r\, \sqrt{\frac{N-n}{N-1}} \tag{4}$$

For the third metric, precision of bycatch estimates, the package follows the standard approach of simulating the response of root mean square estimation error to observer coverage (e.g., Lennert-Cody 2001; Wigley et al., 2007). The package simulates bycatch in a given amount of observed effort $n$ as $n$ random draws without replacement from a random sample of size $N$ from the probability density function (PDF) for the Poisson or negative binomial specified by the user-input $r$ and $d$. The package runs a default of 1000 simulations per observer coverage level. Bycatch estimation CV at each observer coverage level is calculated as the mean square error of estimated vs true realized BPUE in each simulation, divided by the nominal BPUE of the originating PDF.

The package is also served as a web application at https://kacurtis.shinyapps.io/obscov/ (Curtis and Coleman 2019) built using the *shiny* package for R (Chang et al., 2019).

The package is written in *R* (R Core Team 2018) and was developed with the aid of the *devtools* package (Wickham et al., 2018). It employs several other useful add-on packages, including *dplyr* (Wickham et al., 2019), *tibble* (Müller and Wickham, 2019), and *Runuran*, which includes efficient random number generation for Poisson and negative binomial distributions (Leydold and Hörmann 2019).

## 2.2 Workflow

For the first metric, probability of observing any bycatch,  calling function *plot_probposobs()* with user-specified total fishery effort, BPUE, dispersion index, and (optionally) the desired target probability returns a plot of probability of observing bycatch — and of bycatch occurring in the total effort — versus observer coverage (Fig. 1). If output is assigned to an object, the function returns a list containing the results for minimum observer coverage needed to meet the specified probability of observing bycatch, if applicable, and the probability of bycatch occurring in the total fishery effort. The default benchmark is 95% probability of observing bycatch given that it occurred, corresponding to a high confidence of detecting any impact on the species in question.

```
> plot_probposobs(te = 10000, bpue = 0.0005, d = 2)
The probability that any bycatch occurs in the given total effort is 96.9%.
Minimum observer coverage to achieve at least 95% probability of observing
bycatch when total bycatch is positive is 73% (7300 trips or sets).
```

For the second metric, upper confidence limit of bycatch given none observed, calling function *plot_uclnegobs()* with user-specified total effort and expected dispersion index returns a plot of the one-tailed upper confidence limit versus observer coverage when no bycatch is observed (Fig. 2). The function defaults to a confidence level of 95%, but different level can be specified. Upper confidence limits for *d+/-1* are also plotted. A target upper confidence limit can be specified, e.g. aligning with a limit reference point for the species of interest

```
> plot_uclnegobs(te = 10000, d = 2, targetucl = 2, fixedoc = 20)
Minimum observer coverage to ensure that the upper confidence limit of 2 is
not exceeded when no bycatch is observed is 84.8% (8470 trips or sets).
Upper confidence limit for bycatch given none observed in 20% (2000 trips or
sets) coverage is 19.4.
```

For the third metric, bycatch estimation CV, the user first calls function *sim_cv_obscov()* to simulate bycatch CV based on user-specified observer coverage, given total effort, mean BPUE, and dispersion index. The returned list can then be provided to function *plot_cv_obscov()* to plot estimation CV versus observer coverage, with the option of specifying a target estimation CV (Fig. 3). If assigned to an object, the function returns a list containing the minimum observer coverage needed to meet the specified estimation CV. The default target CV is 0.3, which is the upper end of the recommended range for U.S. bycatch monitoring (NMFS 2004) and aligns with a key management strategy evaluation for limit reference points under the U.S. Marine Mammal Protection Act (Wade 1998). Minimum observer coverage is interpolated from the simulation results for the nearest two observer coverage levels.

```
> simlist <- sim_cv_obscov(te = 10000, bpue = 0.0005, d = 2)
> plot_cv_obscov(simlist)
Minimum observer coverage to achieve CV ≤ 0.3 is 82% (8200 trips or sets).
```

The package also includes a function (*run_shiny()*) to run the web application from within R.

## 3. How do observer coverage needs vary with effort, BPUE, and dispersion index?

We explored how observer coverage targets corresponding to specific objectives for each metric – 95% probability of observing bycatch, 95% upper confidence limit set to an arbitrary magnitude of 2, and estimation CV of 0.3 – vary with total effort, BPUE, and dispersion index by allowing one parameter to vary at a time. We explored the following three parameter spaces for each of the three objectives:

(1) Total effort varied from 500 to 20,000, with BPUE = 0.001 and $d = 2$;
(2) BPUE varied from 0.0001 to 0.0101, with total effort = 10,000 and $d = 2$ (this was omitted for the upper confidence limit objective since it solves for BPUE); and
(3) $d$ varied from 1 to 5, with total effort = 10,000 and BPUE = 0.001.

The results (Fig. 4) show the importance of the number of positive observed samples as a key consideration determining the performance of a given level of observer coverage with respect to the two objectives considered. Positive observed samples are determined primarily by effort observed and BPUE. As total effort increases, given that all else is equal, the amount of observed effort needed to observe at least one bycatch event with 95% probability rises rapidly and then plateaus, while percent coverage needed drops quickly at first and then more gradually (top left panel). Observed effort needed for the maximum upper confidence limit objective rises linearly with total effort, while percent coverage is constant (top middle panel). The observed effort needed to attain the target level of bycatch estimation precision also responds more gradually than that needed for the first objective (top right panel). This result underscores the importance of observer effort as well as percent coverage in assessing how well an observer program may be documenting bycatch. Holding total effort and dispersion index constant, increasing BPUE results in exponentially decreasing minimum observer coverage needed to achieve either the first or third objective, though the drop is slower for the estimation precision objective (middle panels). Increasing dispersion index corresponds to logarithmically increasing observer coverage levels needed to meet all three objectives (bottom panels).

## 4. Case study: California drift gillnet fishery

We applied the *ObsCovgTools* package to marine mammals, seabirds, and sea turtles caught in the California drift gillnet fishery for swordfish and thresher sharks (CDGN). Effort in the CDGN is currently approximately 80 trips (500 sets) per year. We estimated minimum observer coverage levels needed to meet each of the following specific objectives for each species: (1) 95% probability of observing bycatch when any occurs; (2) a 95% one-tailed upper confidence limit when no bycatch is observed, equal to a limit reference point for each species; and (3) an estimation CV of 0.3. We used the most recent Potential Biological Removal calculated under the U.S. Marine Mammal Protection Act as the limit reference point for each marine mammal species (Carretta et al., 2018). For leatherback turtles (*Dermochelys coriacea*), we used an analogous reference point that has been estimated for the U.S. West Coast EEZ for the Western

Pacific population (Curtis et al., 2015). No limit reference point is available for other sea turtle or bird species, so observer coverage was not assessed in terms of an upper confidence limit target for those species.

To characterize bycatch per unit effort and dispersion index, we used observer data comprising approximately 440 trips (2,500 fishing sets) from 2002 to 2016, a period with relatively consistent fisheries regulations and approximately 20% observer coverage (Carretta and Barlow 2011, Carretta *et al.* 2017). We included all marine mammal, seabird, and turtle species caught in the fishery during this period. The species included cover a range of bycatch event frequencies (BPUE) and dispersion indices. Several of the species included are management priorities due to population status. We considered observer coverage with respect to effort in one year, five years (the time frame over which cumulative bycatch is considered for marine mammal stock assessments in the United States under the Marine Mammal Protection Act, or MMPA), and ten years. The results are presented in Tables 1-3.

If ten years of observer data are used, the scenario more closely approximating the current bycatch estimation approach in the fishery (see below), the current nominal observer coverage rate of 20% is sufficient to reach the first observer coverage objective of observing any bycatch for 12 of 17 marine mammals, seabirds, and sea turtles observed in the fishery from 2002 to 2016. The second objective of limiting the upper confidence limit of when no bycatch is observed is expected to be met for all species except leatherback turtles, which have an extremely low limit reference point. The third objective of an estimation CV of 0.3 is only expected to be met for the single most commonly encountered species. We caution that if a bycatch event is observed, the problem of estimation CV becomes paramount in assessment of bycatch magnitude versus a limit reference point.

Although the observer coverage levels needed to meet two of three objectives may be cost-prohibitive for many of the infrequently encountered species, the tool provides a useful perspective on the performance of the observer program with respect to these protected species. The prevalence of species with a total bycatch of only one or a few animals in the 15-year data set used underscores the importance of taking a broader risk assessment approach that considers which species were not observed in the portion of the fishery covered by independent observers, but may be vulnerable due to overlap with the fishery in space and time.

The recent history of bycatch estimation and management for sperm whales (*Physeter macrocephalus*) caught in the fishery provides an instructive example of the potential influence of number of observed interactions and observer coverage level on fisheries management. Protected species bycatch for the fishery used to be routinely estimated with a mean-per-unit estimator applied to a single year of observer data. In 2010, the fishery had 12% observer coverage (59 observed fishing sets out of an estimated 492 total sets) and 2 observed sperm whale entanglements in a single set (Carretta et al., 2017). The high observed BPUE in 2010 (0.03 sperm whales per set) was 30 times higher than the aggregate observed sperm whale BPUE over the 21-year period 1990 – 2010. The resulting estimate of sperm whale bycatch for 2010, 16 animals, far exceeded a biologically-based annual limit reference point ("potential biological removal" or PBR; Wade 1998) that serves as an MMPA management threshold. The PBR level at that time was 1.5 animals (Carretta et al., 2014). In response, NOAA implemented an emergency rule that mandated 100% observer coverage in deep offshore waters where the probability of sperm whale bycatch is highest (NOAA 2013, 2014). The temporary rule also included a trigger that would have terminated the fishing season if a single sperm whale was observed entangled.

Subsequently, methodological improvements to estimating rare-event bycatch were implemented in this fishery, replacing single-year mean-per-unit estimates with model-based estimates that included 26 years of data (Carretta et al., 2017). A revised bycatch estimate for sperm whales for 2010 using the new methodology was equal to the number of observed entanglements that year (two), while bycatch estimates for years previously estimated to have zero bycatch were revised to be positive. Although the value added with the model-based approach should not be discounted (Dixon et al., 2005), increasing the sample size on which the bycatch estimate is based is of primary importance to reducing volatility in year-to-year estimates of bycatch, reducing estimation bias and increasing precision (Amandè et al., 2012; Carretta and Moore, 2014), and ultimately, reducing potential 'management overreaction' to short-term fishery monitoring. Where observer coverage cannot be increased, sample size may still be vastly improved by incorporating multiple years of observer data in bycatch estimates, as long as these are reasonably consistent with respect to fisheries, environmental, and biological variables.

## 5. Discussion

The *ObsCovgTools* package for *R* and its companion web application provide a new tool for evaluating observer coverage, with a focus on the challenges of documenting and estimating rare-event bycatch, and for quantifying coverage needed to reach fishery management objectives for monitoring bycatch of threatened, endangered, and protected species.

The package's utilities for assessing the potential for and possible magnitude of unobserved rare event bycatch should be coupled with a risk assessment approach (e.g., Zhou and Griffiths 2008; Hobday et al., 2011; Brown et al., 2015) to assess potential impacts on vulnerable species (e.g., with small population sizes or low productivity) that have not been observed, but might interact with a fishery. This consideration should be a particular priority for fisheries with very low observer coverage levels (<10%), where occurrence of some rare event bycatch could take longer than a decade to document.

The current implementation of *ObsCovgTools* makes several simplifying assumptions. It assumes observer coverage is representative of the fishery, although biases such as nonrandom observer deployment (e.g., Babcock et al, 2003; Benoît and Allard 2009), and observer effects on fishing behavior (e.g., Wahlen and Smith 1985; Babcock et al., 2003; Benoît and Allard 2009) have been documented in fisheries worldwide. It also does not account for unobservable bycatch, e.g., due to dropping out of nets before or during retrieval (e.g., Bisack 1997; Brothers et al., 2010), low detection probability for some species (e.g., due to small size), swimming away with gear (e.g., Knowlton 2005), or fishery interactions that do not involve direct catch in gear (e.g., Ryan 1991). Finally, it does not account for hierarchical sources of variance (e.g., vessel- or trip-level variation). Violating these assumptions is likely to produce negatively biased projections of observer coverage needed to meet specific objectives. Unless hierarchical sources of variance can be ruled out as potentially important, using higher-level units of effort is advised (e.g., mean bycatch per trip and number of trips, instead of mean bycatch per set and number of sets). On the other hand, the mean-per-unit estimator that underlies the estimation CV simulations can serve as conservative benchmark that tends to project a positively biased estimation CV at a given observer coverage level when additional variability can be explained through a model-based approach.

Additional modules can easily be added to the package in the future. One important objective for observer coverage that is not yet implemented is the power to detect a change in bycatch rates in a fishery when a change in management, fishing effort, or environmental conditions occurs. More complex scenarios such as stratified sampling may also be useful additions to a future version of the package. The open-source platform used for *ObsCovgTools* lends itself to maintenance and further development by any user as the need arises.

Current observer program coverage in many fisheries is insufficient to support accurate and precise assessments of bycatch of sensitive species (e.g., Moore et al., 2009; Gilman et al., 2014). In sharing the *ObsCovgTools* package, it is our aim to facilitate and encourage the consideration of sensitive species and rare event bycatch in the design and evaluation of fisheries observer program sampling, as well as of the potential for rare fishery interactions with species not previously observed in the fishery, whose populations are vulnerable to even a small number of removals.

## Funding source

This work was supported by the Protected Species Science Branch of the National Marine Fisheries Service (NMFS) Office of Science and Technology's Assessment & Monitoring Division. The funding source had no role in the design or execution of the study, writing of the report, or submission for publication.

## Acknowledgments

Special thanks to Howard Coleman, who developed the framework and initial version of the Shiny application. Jeff Moore was a PI on the Protected Species Toolbox proposal that funded the work, provided detailed input on the user interface along the way, pushed tool development further, and provided feedback on the manuscript. Eli Holmes was a PI on the proposal, provided guidance on github and package development, and provided an initial home for the Shiny application on a server at NWFSC. Tool development also benefitted from input and constructive feedback from a large number of NOAA Fisheries colleagues, particularly Jason Jannot, Penny Ruvelas, Melissa Soldevilla, and Marjorie Lyssikatos. We also thank an internal reviewer at SWFSC, Jonathan Sweeney, and the handling editor and two external reviewers for strengthening the manuscript through their comments and suggestions.

## Declarations of interest

None.

## References

Amandè, M. J., Chassot, E., Chavance, P., Murua, H., Delgado de Molina, A., and Bez, N. 2012. Precision in bycatch estimates: the case of tuna purse-seine fisheries in the Indian Ocean. ICES Journal of Marine Science, 69: 1501–1510.

Ancelet, S., Etienne, M–P., Benoît, H., and Parent, E. 2010. Modelling spatial zero-inflated continuous data with an exponentially compound Poisson process. Environ Ecol Stat, 17: 347–376.

Babcock E, Pikitch E, and Hudson, Charlotte. 2003. How much observer coverage is enough to adequately estimate bycatch. Oceana, Washington, DC. 36 pp.

Benoît HP and Allard J. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Can. J. Fish. Aquat. Sci. 66: 2025–2039

Bisack, K. D. 1997. Harbor porpoise bycatch estimates in the US New England multispecies sink gillnet fishery: 1994–1995. Reports of the International Whaling Commission 47: 705–714.

Bravington, M.A., C. Burridge, and P. Toscas.2003. Design of observer program to monitor bycatch species in the eastern tuna and billfish fishery, Working Paper 7, Sixteenth Meeting of the Standing Committee on Tuna and Billfish, Mooloolaba, 9-16 July 2003.

Brothers, N., Duckworth, A.R., Safina, C. & Gilman, E.L. (2010) Seabird bycatch in pelagic longline fisheries is grossly underestimated when using only haul data. PLoS ONE 5: e12491.

Brown, S.L., D. Reid, and E. Rogan. 2015. Spatial and temporal assessment of potential risk to cetaceans from static fishing gears. Marine Policy 51: 267-280.

Carretta, J.V. and J. Barlow. 2011. Long-term effectiveness, failure rates, and "dinner bell" effects of acoustic pingers in a gillnet fishery. Marine Technology Society Journal 45(5):7-19.

Carretta, J.V. and J.E. Moore. 2014. Recommendations for pooling annual bycatch estimates when events are rare. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-528. 11 p.

Carretta, J.V., E. Oleson, D.W. Weller, A.R. Lang, K.A. Forney, J. Baker, B. Hanson, K. Martien, M.M. Muto, A.J. Orr, H. Huber, M.S. Lowry, J. Barlow, D. Lynch, L. Carswell, R.L. Brownell Jr., and D.K. Mattila. 2014. U.S. Pacific Marine Mammal Stock Assessments: 2013. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-532. 406 p.

Carretta, J.V., J.E. Moore, and K.A. Forney. 2017. Regression tree and ratio estimates of marine mammal, sea turtle, and seabird bycatch in the California drift gillnet fishery: 1990-2015. NOAA Technical Memorandum, NOAA-TM-NMFS-SWFSC-568. 83 p. doi:10.7289/V5/TM-SWFSC-568.

Carretta, J.V., K.A. Forney, E.M. Oleson, D.W. Weller, A.R. Lang, J. Baker, M.M. Muto, B. Hanson, A.J. Orr, H. Huber, M.S. Lowry, J. Barlow, J.E. Moore, D. Lynch, L. Carswell, and R.L. Brownell Jr. 2019. U.S. Pacific Marine Mammal Stock Assessments: 2018. U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-617.

Chang W, J Cheng, JJ Allaire, Y Xie and J McPherson. 2019. *shiny*: Web Application Framework for *R*. R package version 1.3.0. https://CRAN.R-project.org/package=shiny

Cochran, W.G. 1977. *Sampling Techniques*. Third Edition, New York: John Wiley & Sons.

Crowder, L.B. and S.A. Murawski. 1998. Fisheries bycatch: implications for management. Fisheries 23: 8–15.

Curtis KA. 2019. Zenodo release of github-served *ObsCovgTools* package. *To be completed prior to final submission. Development version of package available at https://github.com/kacurtis/ObsCovgTools*

Curtis KA and H Coleman. 2019. Observer coverage simulator (Version 2.1.1). https://kacurtis.shinyapps.io/obscov/

Curtis KA, Moore JE, Benson SR. 2015. Estimating limit reference points for Western Pacific Leatherback Turtles (*Dermochelys coriacea*) in the U.S. West Coast EEZ. PLoS ONE 10(9): e0136452. doi:10.1371/journal.pone.0136452

Davies, S.L. and J.E. Reynolds (ed.) 2002. Guidelines for developing an at-sea fishery observer programme. FAO Fisheries Technical Paper, No. 414, Rome, FAO. 116p.

Dixon, P. M., A. M. Ellison, and N. J. Gotelli. 2005. Improving the precision of estimates of the frequency of rare events. Ecology 86:1114–1123.

Federal Register. 2013. Fisheries off West Coast states; highly migratory fisheries; California drift gillnet fishery; sperm whale interaction restriction. FR 78:54547-54552.

Gilman, E., K. Passfield, and K. Nakamura. 2014. Performance of regional fisheries management organizations: ecosystem-based governance of bycatch and discards. Fish and Fisheries 15: 327–351.

Hall, M. 1999. Estimating the ecological impacts of fisheries: What data are needed to estimate bycatches? In: Nolan, C. (ed.), Proceedings of the International Conference on Integrated Fisheries Monitoring. Sydney, Australia, 1-5 February 1999. FAO, Rome, pp. 175–184.

Hall S.J. and B.M. Mainprize. 2005. Managing by-catch and discards: how much progress are we making and how can we do better? Fish and Fisheries 6:134–155.

Hilborn, R. and Mangel, M. 1997. *The Ecological Detective: Confronting Models with Data.* Princeton University Press, Princeton.

Hobday, A. J., Smith, A. D. M., Stobutzki, I. C., Bulman, C., Daley, R., Dambacher, J. M., Deng, R. A., Dowdney, J., Fuller, M., Furlani, D., Griffiths, S. P., Johnson, D., Kenyon, R., Knuckey, I.A., Ling, S. D., Pitcher, R., Sainsbury, K. J., Sporcic, M., Smith, T., Turnbull, C., Walker, T. I., Wayte, S. E., Webb, H., Williams, A., Wise, B. S. & Zhou, S. 2011. Ecological risk assessment for the effects of fishing. Fisheries Research 108: 372-384.

Karp, W.A. and H. McElderry. 1999. Catch monitoring by fisheries observers in the United States and Canada. In: Nolan, C.P. (ed.), Proceedings of the International Conference on Integrated Fisheries Monitoring. Sydney, Australia, 1-5 February 1999. FAO, Rome, pp. 261-284.

Kirby, D.S. and P. Ward. 2014. Standards for the effective management of fisheries bycatch. Marine Policy 44: 419-426.

Knowlton AR. 2005. A review of gear modifications and fishing practices aimed at reducing the level and frequency of entanglements with right and humpback whales in the western North Atlantic. Report to National Marine Fisheries Service, Northeast Region, Gloucester Point, MA. NOAA Contract #40EMNF300159

Lennert-Cody, C.E. 2001. Effects of sample size on bycatch estimation using systematic sampling and post-stratification: Summary of preliminary results. IOTC Proceedings WPDCS-01-09, 4: 48–53.

Leydold J and W Hörmann. 2019. *Runuran*: R Interface to the 'UNU.RAN' Random Variate Generators. R package version 0.27. https://CRAN.R-project.org/package=Runuran

Lyssikatos MC, Garrison LP. 2018. Common Bottlenose Dolphin (*Tursiops truncatus*) gillnet bycatch estimates along the US Mid-Atlantic Coast, 2007-2015. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 18-07; 37 p. Available from: http://www.nefsc.noaa.gov/publications/

Manly, B.F.J., A. Seyb, and D.J. Fletcher. 2002. Bycatch of sea lions (*Phocarctos hookeri*) in NZ fisheries, 1987/88 to 1995/96, and observer coverage. DOC Science Internal Series 42, 21 p.

Martin, S.L., Stohs, S.M. and Moore, J.E., 2015. Bayesian inference and assessment for rare-event bycatch in marine fisheries: a drift gillnet fishery case study. Ecological Applications, 25(2), pp.416-429.

Minami, M., C.E. Lennert-Cody, W. Gao, and M. Román-Verdesoto. 2007. Modeling shark bycatch: The zero-inflated negative binomial regression model with smoothing. Fisheries Research 84:210-221.

Moore, J.E., Wallace B.P., Lewison R.L., Żydelis, R., Cox T.M., Crowder L.B. 2009. A review of marine mammal, sea turtle and seabird bycatch in USA fisheries and the role of policy in shaping management. Marine Policy 33: 435– 451.

Müller K and H Wickham. 2019. *tibble*: Simple Data Frames. R package version 2.1.1. https://CRAN.R-project.org/package=tibble

NMFS. 2004. Evaluating bycatch: a national approach to standardized bycatch monitoring programs. U.S. Dep. Commerc. NOAA Tech. Memo. NMFS-F/SPO-66, Silver Spring, MD. On-line version, http://spo.nmfs.noaa.gov/tm

NOAA. 2013. Fisheries Off West Coast States; Highly Migratory Fisheries; California Drift Gillnet Fishery; Sperm Whale Interaction Restriction. Federal Register. 78:54548-54552.

NOAA. 2014. Fisheries Off West Coast States; Highly Migratory Fisheries; California Drift Gillnet Fishery; Sperm Whale Interaction Restrictions. Federal Register. 79: 29377-29379.

Parkes, G., and M. S. Kaiser. 2004. NMFS Fisheries Observer Coverage Level Workshop: Defining a Basis. NMFS, Seattle, WA

Pérez Roda, M.A. (ed.), Gilman, E., T. Huntington, S.J. Kennelly, P. Suuronen, M. Chaloupka and P. Medley. 2019. A third assessment of global marine fisheries discards. FAO Fisheries and Aquaculture Technical Paper No. 633. Rome, FAO. 78 pp.

R Core Team. 2018. *R*: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ryan PG. 1991. The impact of the commercial lobster fishery on seabirds at the Tristan da Cunha islands, South Atlantic. Biological Conservation, 57, 339–350.

Smith, M.H. and S.J. Baird. 2005. Representativeness of past observer coverage, and future coverage required for estimation of New Zealand sea lion (*Phocarctos hookeri*) captures in the SQU 6T fishery. New Zealand Fisheries Assessment Report 2005/5 39 p.

Wade, P.R., 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. Marine Mammal Science, 14(1), pp.1-37.

Wickham H, J Hester, and W Chang. 2018. *devtools*: Tools to Make Developing R Packages Easier. R package version 2.0.1. https://CRAN.R-project.org/package=devtools

Wickham H, R François, L Henry and K Müller. 2019. dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. https://CRAN.R-project.org/package=dplyr

Wigley SE, Rago PJ, Sosebee KA, Palka DL. 2007. The analytic component to the Standardized Bycatch Reporting Methodology Omnibus Amendment: sampling design end estimation of precision and accuracy (2nd edition). U.S. Dep. Commer., Northeast Fish. Sci. Cent. Ref. Doc. 07-09; 156 p.

Wigley, S.E. and C. Tholke. 2019. 2019 Discard Estimation, Precision, and Sample Size Analyses for 14 Federally Managed Species Groups in the Waters off the Northeastern United States. NOAA Technical Memorandum NMFS-NE-254. 179 p.

Zhou, S. and Griffiths, S.P. 2008. Sustainability assessment for fishing effects (SAFE): a new quantitative ecological risk assessment method and its application to elasmobranch bycatch in an Australian trawl fishery. Fisheries Research 91: 56-68.

## Tables and Figures

| **Species** | **Total bycatch** | **BPUE** | **d** | **LRP** | **1 year % (n)** | **5 years % (n)** | **10 years % (n)** |
|---|---|---|---|---|---|---|---|
| Loggerhead Sea Turtle (*Caretta caretta*) | 1 | 0.002 | 1 | – | 95 (76) | 92 (369) | 88 (700) |
| Minke Whale (*Balaenoptera acutorostrata*) | 1 | 0.002 | 1 | 3.5 | 95 (76) | 92 (369) | 88 (700) |
| Bottlenose Dolphin (*Tursiops truncates*) | 1 | 0.002 | 1 | 11 | 95 (76) | 92 (369) | 88 (700) |
| Humpback Whale (*Megaptera novaeangliae*) | 1 | 0.002 | 1 | 16.7 | 95 (76) | 92 (369) | 88 (700) |
| Dall's Porpoise (*Phocoenoides dalli*) | 1 | 0.002 | 1 | 172 | 95 (76) | 92 (369) | 88 (700) |
| Leatherback Sea Turtle (*Dermochelys coriacea*) | 2 | 0.005 | 1 | 0.16 | 95 (76) | 88 (350) | 72 (572) |
| Gray Whale (*Eschrichtius robustus*) | 2 | 0.005 | 1 | 801 | 95 (76) | 88 (350) | 72 (572) |
| Sperm Whale (*Physeter macrocephalus*) | 2 | 0.005 | 2 | 2.5 | 95 (76) | 91 (363) | 82 (658) |
| Short-Finned Pilot Whale (*Globicephala macrorhynchus*) | 3 | 0.007 | 1 | 4.5 | 94 (75) | 81 (322) | 54 (430) |
| Northern Elephant Seal (*Mirounga angustirostris*) | 5 | 0.011 | 1 | 4882 | 92 (74) | 62 (250) | 33 (265) |
| Risso's Dolphin (*Grampus griseus*) | 6 | 0.014 | 3 | 46 | 94 (75) | 78 (312) | 49 (395) |
| Pacific White-Sided Dolphin (*Lagenorhynchus obliquidens*) | 9 | 0.02 | 2.1 | 191 | 92 (74) | 53 (213) | 27 (219) |
| Long-Beaked Common Dolphin (*Delphinus capensis*) | 12 | 0.027 | 1.5 | 657 | 89 (71) | 34 (135) | 17 (135) |
| Northern Right Whale Dolphin (*Lissodelphis borealis*) | 17 | 0.038 | 1.6 | 179 | 84 (67) | 25 (98.1) | 12 (98.1) |
| Northern Fulmar (*Fulmarus glacialis*) | 20 | 0.045 | 3.5 | – | 88 (70) | 33 (132) | 16 (132) |
| California Sea Lion (*Zalophus californianus*) | 92 | 0.208 | 4 | 14011 | 40 (32) | 8 (32) | 4 (32) |
| Short-Beaked Common Dolphin (*Delphinus delphis*) | 94 | 0.213 | 1.4 | 8393 | 21 (17) | 4 (17) | 2 (17) |

Table 1. Minimum observer coverage levels needed in the California drift gillnet fishery to achieve the objective of observing any bycatch that occurs with 95% probability. Mean bycatch per unit effort (BPUE) and dispersion index (d, variance divided by the mean) are based on data from 2002 to 2016, a period with relatively consistent fisheries regulations. Limit reference points (LRP) are drawn from the most recent stock assessments for the relevant populations and from the literature (see Section 4). We included all marine mammal, seabird, and turtle species caught in the fishery during this period. We evaluated the objective for three durations over which data might reasonably be aggregated for evaluation: one year (total effort of 80 trips), five years (400 trips), and ten years (800 trips). For each time period and species, we report minimum observer coverage needed, in terms of percentage and effort (n, trips). Species are sorted by total bycatch, dispersion index, and limit reference point. The current nominal observer coverage is approximately 20%.

| **Species** | **Total bycatch** | **BPUE** | **d** | **LRP** | **1 year % (n)** | **5 years % (n)** | **10 years % (n)** |
|---|---|---|---|---|---|---|---|
| Loggerhead Sea Turtle (*Caretta caretta*) | 1 | 0.002 | 1 | – | – | – | – |
| Minke Whale (*Balaenoptera acutorostrata*) | 1 | 0.002 | 1 | 3.5 | 58 (46) | 16 (13) | 9 (7) |
| Bottlenose Dolphin (*Tursiops truncates*) | 1 | 0.002 | 1 | 11 | 25 (20) | 6 (5) | 4 (3) |
| Humpback Whale (*Megaptera novaeangliae*) | 1 | 0.002 | 1 | 16.7 | 18 (14) | 4 (3) | 2 (2) |
| Dall's Porpoise (*Phocoenoides dalli*) | 1 | 0.002 | 1 | 172 | 2 (2) | 1 (1) | 1 (1) |
| Leatherback Sea Turtle (*Dermochelys coriacea*) | 2 | 0.005 | 1 | 0.16 | 100 (80) | 95 (76) | 82 (66) |
| Gray Whale (*Eschrichtius robustus*) | 2 | 0.005 | 1 | 801 | 1 (1) | 1 (1) | 1 (1) |
| Sperm Whale (*Physeter macrocephalus*) | 2 | 0.005 | 2 | 2.5 | 80 (64) | 30 (24) | 16 (13) |
| Short-Finned Pilot Whale (*Globicephala macrorhynchus*) | 3 | 0.007 | 1 | 4.5 | 49 (39) | 14 (11) | 8 (6) |
| Northern Elephant Seal (*Mirounga angustirostris*) | 5 | 0.011 | 1 | 4882 | 1 (1) | 1 (1) | 1 (1) |
| Risso's Dolphin (*Grampus griseus*) | 6 | 0.014 | 3 | 46 | 11 (9) | 2 (2) | 1 (1) |
| Pacific White-Sided Dolphin (*Lagenorhynchus obliquidens*) | 9 | 0.02 | 2.1 | 191 | 2 (2) | 1 (1) | 1 (1) |
| Long-Beaked Common Dolphin (*Delphinus capensis*) | 12 | 0.027 | 1.5 | 657 | 1 (1) | 1 (1) | 1 (1) |
| Northern Right Whale Dolphin (*Lissodelphis borealis*) | 17 | 0.038 | 1.6 | 179 | 2 (2) | 1 (1) | 1 (1) |
| Northern Fulmar (*Fulmarus glacialis*) | 20 | 0.045 | 3.5 | – | – | – | – |
| California Sea Lion (*Zalophus californianus*) | 92 | 0.208 | 4 | 14011 | 1 (1) | 1 (1) | 1 (1) |
| Short-Beaked Common Dolphin (*Delphinus delphis*) | 94 | 0.213 | 1.4 | 8393 | 1 (1) | 1 (1) | 1 (1) |

Table 2. Minimum observer coverage levels needed in the California drift gillnet fishery to achieve the objective of a 95% upper confidence limit no higher than the LRP when no bycatch has been observed. Mean bycatch per unit effort (BPUE) and dispersion index (d, variance divided by the mean) are based on data from 2002 to 2016, a period with relatively consistent fisheries regulations. Limit reference points (LRP) are drawn from the most recent stock assessments for the relevant populations and from the literature (see Section 4). We included all marine mammal, seabird, and turtle species caught in the fishery during this period. We evaluated the objective for three durations over which data might reasonably be aggregated for evaluation: one year (total effort of 80 trips), five years (400 trips), and ten years (800 trips). For each time period and species, we report minimum observer coverage needed, in terms of percentage and effort (n, trips). Species are sorted by total bycatch, dispersion index, and limit reference point. The current nominal observer coverage is approximately 20%.

| Species | Total bycatch | BPUE | d | LRP | 1 year % (n) | 5 years % (n) | 10 years % (n) |
|---|---|---|---|---|---|---|---|
| Loggerhead Sea Turtle (*Caretta caretta*) | 1 | 0.002 | 1 | – | 98 (78) | 92 (370) | 86 (690) |
| Minke Whale (*Balaenoptera acutorostrata*) | 1 | 0.002 | 1 | 3.5 | 98 (78) | 92 (370) | 85 (680) |
| Bottlenose Dolphin (*Tursiops truncates*) | 1 | 0.002 | 1 | 11 | 97 (78) | 92 (370) | 87 (690) |
| Humpback Whale (*Megaptera novaeangliae*) | 1 | 0.002 | 1 | 16.7 | 98 (78) | 92 (370) | 87 (700) |
| Dall's Porpoise (*Phocoenoides dalli*) | 1 | 0.002 | 1 | 172 | 97 (78) | 93 (380) | 87 (700) |
| Leatherback Sea Turtle (*Dermochelys coriacea*) | 2 | 0.005 | 1 | 0.16 | 96 (77) | 87 (350) | 77 (620) |
| Gray Whale (*Eschrichtius robustus*) | 2 | 0.005 | 1 | 801 | 96 (77) | 87 (350) | 76 (610) |
| Sperm Whale (*Physeter macrocephalus*) | 2 | 0.005 | 2 | 2.5 | 97 (78) | 92 (370) | 84 (680) |
| Short-Finned Pilot Whale (*Globicephala macrorhynchus*) | 3 | 0.007 | 1 | 4.5 | 96 (77) | 81 (330) | 68 (550) |
| Northern Elephant Seal (*Mirounga angustirostris*) | 5 | 0.011 | 1 | 4882 | 93 (75) | 71 (290) | 56 (450) |
| Risso's Dolphin (*Grampus griseus*) | 6 | 0.014 | 3 | 46 | 95 (76) | 87 (350) | 74 (590) |
| Pacific White-Sided Dolphin (*Lagenorhynchus obliquidens*) | 9 | 0.02 | 2.1 | 191 | 92 (73) | 74 (300) | 60 (480) |
| Long-Beaked Common Dolphin (*Delphinus capensis*) | 12 | 0.027 | 1.5 | 657 | 89 (71) | 62 (250) | 44 (350) |
| Northern Right Whale Dolphin (*Lissodelphis borealis*) | 17 | 0.038 | 1.6 | 179 | 86 (69) | 54 (220) | 35 (280) |
| Northern Fulmar (*Fulmarus glacialis*) | 20 | 0.045 | 3.5 | – | 92 (74) | 68 (280) | 52 (420) |
| California Sea Lion (*Zalophus californianus*) | 92 | 0.208 | 4 | 14011 | 75 (60) | 35 (140) | 22 (180) |
| Short-Beaked Common Dolphin (*Delphinus delphis*) | 94 | 0.213 | 1.4 | 8393 | 47 (38) | 16 (63) | 9 (72) |

Table 3. Minimum observer coverage levels needed in the California drift gillnet fishery to achieve the objective of an estimation CV of 0.3. Mean bycatch per unit effort (BPUE) and dispersion index (d, variance divided by the mean) are based on data from 2002 to 2016, a period with relatively consistent fisheries regulations. Limit reference points (LRP) are drawn from the most recent stock assessments for the relevant populations and from the literature (see Section 4). We included all marine mammal, seabird, and turtle species caught in the fishery during this period. We evaluated the objective for three durations over which data might reasonably be aggregated for evaluation: one year (total effort of 80 trips), five years (400 trips), and ten years (800 trips). For each time period and species, we report minimum observer coverage needed, in terms of percentage and effort (n, trips). Species are sorted by total bycatch, dispersion index, and limit reference point. The current nominal observer coverage is approximately 20%.

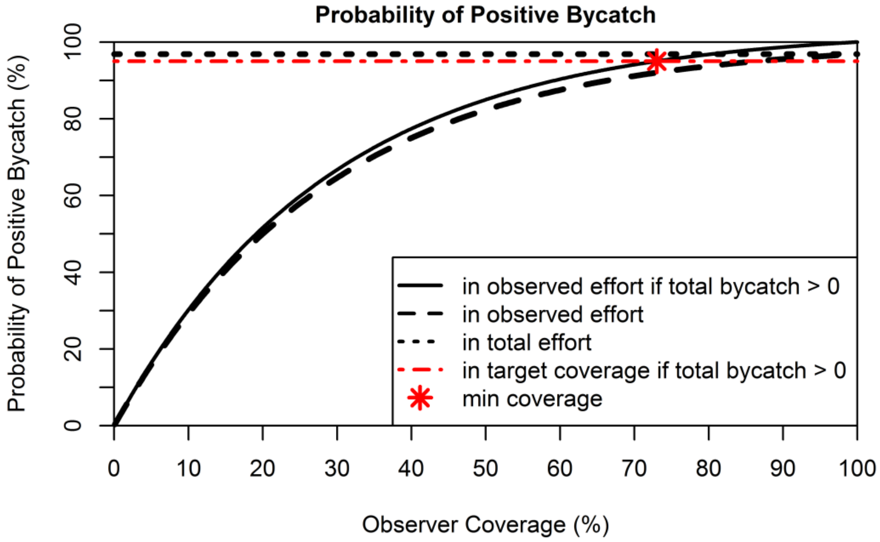

Figure 1. Example plot output from *ObsCovgTools* package for observer coverage metric of probabilities of observing any bycatch and of any bycatch in the total effort. User inputs were total fishery effort of 10,000 (arbitrary units, e.g., trips, days, sets), bycatch per unit effort of 0.0005, and dispersion index of 2. When target probability is specified, corresponding minimum observer coverage is based on the conditional probability of observing any bycatch if it occurs (solid black line), obtained by dividing the absolute probability of observing any bycatch (black dashed line) by the probability that any bycatch occurs in the given total effort.

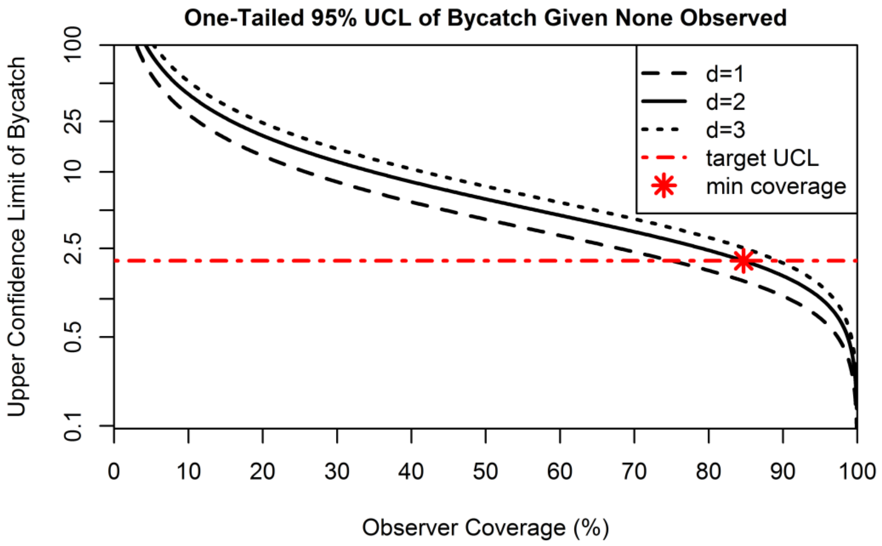

Figure 2. Example plot output from *ObsCovgTools* package for observer coverage metric of one-tailed upper confidence limit of bycatch given none observed. User inputs were total fishery effort of 10,000 (arbitrary units, e.g., trips, days, sets), dispersion index of 2, and target UCL equal to 2.

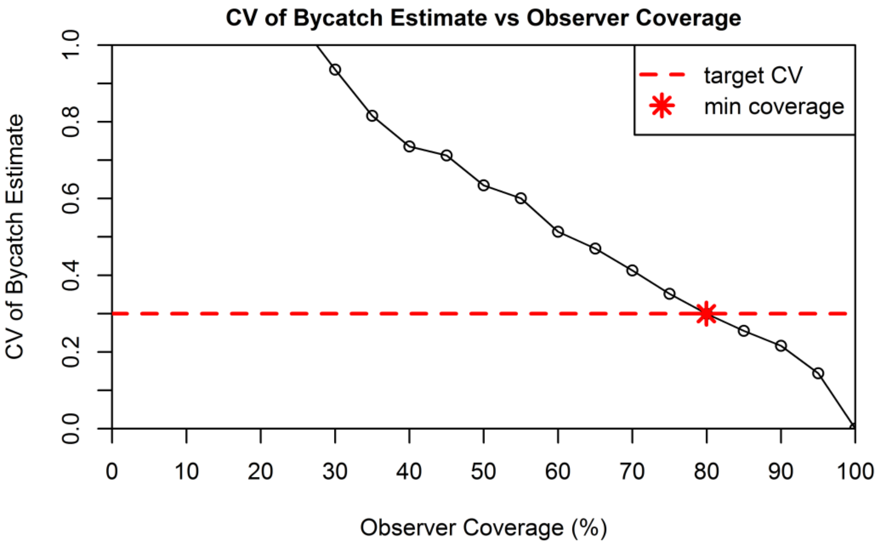

Figure 3. Example plot output from *ObsCovgTools* package for observer coverage metric of bycatch estimation CV. User inputs were total fishery effort of 10,000 (arbitrary units, e.g., trips, days, sets), bycatch per unit effort of 0.0005, and dispersion index of 2.

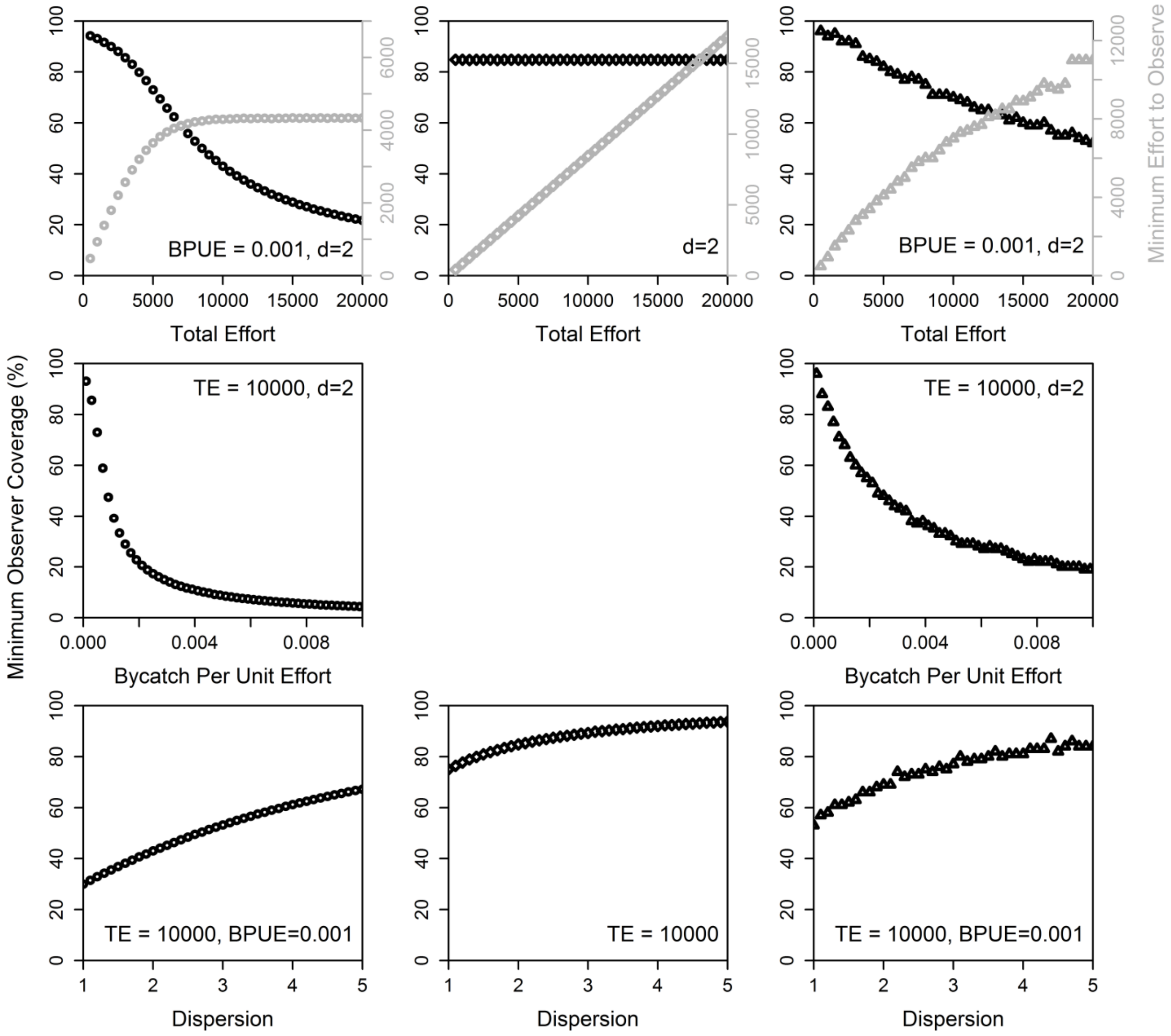

Figure 4. Variation of observer coverage needs with total effort (TE), BPUE, and dispersion index (d). Minimum observer coverage needed, in terms of percentage (black symbols) and effort to observe (gray symbols), in order to meet (1) a target probability of 95% of observing any bycatch if bycatch occurs (left panels, open circles), (2) a target upper confidence limit (UCL) given no positive bycatch in observed effort (middle panels, open diamonds) or (3) a target estimation CV of 0.3 (right panels, open triangles). Variation of observer coverage targets for each objective is explored with respect to changing TE (top panels), BPUE (middle panels), and d (bottom panels). Minimum effort to observe (gray) is omitted from the middle and bottom rows, since it varies linearly with percentage coverage when total effort is fixed. No plot exists for the UCL objective with varying BPUE, because this function solves for BPUE.