Charter-Boat Logbook Reporting Pilot Study

Estimators for Use with Logbook Data

Mark S. Kaiser

Dept. of Statistics, Iowa State University

Prepared for

Marine Resources Assessment Group Americas

MRAG Americas

November 2012

1 Introduction

A major conclusion in the final report "For-Hire Electronic Logbook Pilot Study in the Gulf of Mexico" of the Marine Recreational Information Program (MRIP) was that a logbook program such as that investigated in the pilot study could not be considered to provide a census of trips in the for-hire fishery of the Gulf of Mexico, but might be a useful data source for estimation of catch and catch disposition. This report develops a basic estimator that has potential for use with logbook data. This estimator was initially described in a report from the Marine Resources Assessment Group - Americas (MRAG) titled "Charter-Boat Logbook Reporting Pilot Study: Initial Examination of Data", prepared for MRIP and dated August 2011. Here, details needed for implementation of the estimator are examined, and examples of its use are given for red snapper and vermilion snapper. Attention will be restricted to the portion of Florida included in the MRIP pilot study. One of the primary objectives of this work was to compare the use of logbooks to provide information on catch with the use of dockside sampling to provide information on catch.

This report assumes that the reader is familiar with the basic theory of Bayesian estimation, the operation of Markov Chain Monte Carlo (MCMC) algorithms, the production of Bayesian credible intervals from simulated posterior distributions, and the use of posterior predictive distributions for model assessment.

2 The Estimator

The basic structure of the estimator is given in this section. Many estimators in fisheries science have the form of effort multiplied by catch-per-unit-effort (cpue). This forms the basis of the estimator to be considered here. Assume that the quantity to be estimated is a total denoted by τ. This could be the total harvest in the fishery (or a well-defined portion of the fishery) for a given species or taxonomic group over a specified period of time, or the total number of individuals released over a period of time, or the total number of mortalities. The time period for which an estimator is applicable could be a year, a season, what is called a "wave" in some programs, or a week. The only requirement is that the relevant time period be long enough to contain sufficient fishing activity to provide data for estimation.

Let N denote the total number of trips taken in the population for which the total τ is defined. Let κ denote the average effort per trip (in specified units) and let ψ denote the average cpue relative to those same units. In the examples presented in this report, effort will be in units of hours fished (one angler for one hour) and cpue will be in number of individuals per fishing hour. Given values for N, κ and ψ, the desired total is

$$\tau = N\kappa\psi. \tag{1}$$

An point estimator of τ, $\hat{\tau}$, would result from replacing N, κ and ψ in (1) with estimators $\hat{N}$, $\hat{\kappa}$ and $\hat{\psi}$. In estimation of number of releases or mortalities, the definition of "catch" in cpue is modified appropriately – for harvest, "catch" is in fact catch, while for releases it is the number released. In addition to point estimates, however, we desire a procedure for quantifying uncertainty in those estimates.

Obtaining a measure of uncertainty about $\hat{\tau}$, such as a standard error, based on survey sampling theory can be difficult, and is sometimes approached through the use of Taylor expansions and/or asymptotic results such as what is often called the delta method (or the continuous mapping theorem). A straightforward alternative taken in this report is the use of a Bayesian strategy for analysis. If a posterior distribution for τ can be obtained, either analytically or through simulation, uncertainty is automatically captured in the quantiles of that distribution. For this application, the posterior will be obtained through simulation using a combination of direct sampling from known distributions and what are called Markov Chain Monte Carlo (MCMC) methods. Notice that what will be needed are simulated values from the distribution of the *expected values* of effort and cpue on individual trips. The next sections describe models used to obtain posterior distributions necessary for such simulation.

It should be noted that the use of effort and cpue in (1) differs from what might be considered a typical approach in fishery science. In many fisheries applications, the sources of information for estimation of effort and cpue differ. Effort is presumed available from a more expansive or extensive source of information (or perhaps even known in total without estimation), while cpue is estimated from a smaller source of information due to greater difficulty in obtaining accurate data on this quantity. In addition, if estimation of cpue proceeds from the theory of survey sampling, that estimator will typically take total catch from a data source (total over all sampling units) divided by total effort from that same source as the point estimator. The estimator of (1) differs from that approach. The estimator proposed in this report will be equal to the product of expected effort (on a per trip basis) and the expected cpue (again on a per trip basis) then also multiplied by the estimated (rather than known) number of total trips. From a survey sampling perspective this would seem a less than optimal approach, so it is important to realize that the estimator proposed here does not derive its motivation or justification from the theory of survey sampling. In other work on estimation of discard in the groundfish fishery off the northeastern portion of the United States, the approach embodied by the estimator of this report has been demonstrated to be substantially superior to the traditional survey sampling (ratio) estimator used by NMFS (MRAG report "Development of an Estimator of Discard for the Northeast Observer Program", January 2006). While that of course does not imply that the estimator proposed here should necessarily be superior to sampling-based alternatives in this problem, it does motivate its consideration. The fundamental character of the approach proposed in this report is that the distribution of what is known about a population total (τ in (1)) can be obtained by examining the distribution of random draws from distributions that reflect what is known about its three component quantities (N, κ, and ψ in (1)). And, once a distribution has been determined that describes what is known about τ, intervals and other inferential quantities are immediately available.

As indicated in the previous paragraph, effort and cpue were modeled using the same data even though cpue is defined (on a trip-by-trip basis) as the ratio of catch (or release) to effort. An immediate concern is then the potential for induced correlation between these quantities and the effect of that potential correlation on (estimation of) uncertainty in a composite estimator that follows the form of expression (1). This concern will be addressed in Section 9 after the overall estimation strategy has been developed.

3 Estimating Number of Trips

Estimation of a population total using either logbook or dockside sampling as data sources for information on catch requires a value for the total number of trips taken. A part of the MRIP logbook pilot study involved what was called "prevalidation" sampling by the project team and was referred to as "activity monitoring" in previous MRAG reports on data analysis from the pilot study. These data consist of counts of the number of trips that could be verified as fishing trips by port agents, and the number of those that resulted in corresponding logbook reports. Given the total number of logbook reports filed in a given period and these data it is possible to estimate the total number of trips taken. Let m denote the number of verified fishing trips taken in a period of concern, and let X denote a random variable connected with the number of those trips that had corresponding logbook reports filed. A binomial data model is assigned to X having probability mass function, for $0 < \theta < 1$,

$$f(x|\theta) = \frac{m!}{x!\,(m-x)!}\theta^x\,(1-\theta)^{m-x}; \quad x = 0, 1, \ldots, m \tag{2}$$

In (2) θ represents the proportion of all fishing trips in the population for which logbook reports were filed, and is to be estimated from the sample of verified fishing trips. To accomplish this, assign a prior distribution to what we believe about the value of θ as a beta distribution,

$$\pi_t(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, \theta^{\alpha-1}\,(1-\theta)^{\beta-1}; \quad 0 < \theta < 1, \tag{3}$$

where $\alpha > 0$ and $\beta > 0$ in (3) are specified numbers, not unknown parameters to be estimated. For example, choosing $\alpha = \beta = 1$ produces a uniform prior on the interval (0, 1), assigning equal probability to any equal interval of possible values of θ.

Given the data model (2) and the prior (3), the posterior distribution for what we believe about θ becomes

$$p_t(\theta|x) = \frac{\Gamma(\alpha+\beta+m)}{\Gamma(\alpha+x)\Gamma(\beta+m-x)}\, \theta^{\alpha+x-1}\,(1-\theta)^{\beta+m-x-1}; \quad 0 < \theta < 1. \tag{4}$$

The posterior (4) is again a beta distribution. A point estimate of the proportion of trips in the population for which a logbook report was filed may then be taken as the expected value (mean) of the posterior distribution,

$$\hat{\theta} = \frac{\alpha+x}{\alpha+\beta+m}, \tag{5}$$

and a point estimate of the total number of trips taken as the number of logbook reports M divided by this value,

$$\hat{N} = M/\hat{\theta}. \tag{6}$$

As an illustration, in week 23 of 2011 (the first full week of the red snapper season in that year) there were 69 verified fishing trips of which 54 had corresponding logbook reports filed. There were also a total of 711 logbook reports filed for that week. If one begins with a uniform prior, $\alpha = \beta = 1$ in (3), the posterior expected values of θ and N are,

$$\begin{aligned} \hat{\theta} &= \frac{1+54}{1+1+69} = 0.775 \\ \hat{N} &= M/\hat{\theta} = 711/0.775 = 917. \end{aligned}$$

Graphs of the prior and posterior densities for this example are presented in Figure 1. The 0.025 and 0.975 quantiles of a beta distribution with parameters 55 and 16 (the posterior parameter values) are 0.671 and 0.863, respectively, so that a 95% credible interval for the number of fishing trips taken in week 23 is,

$$(711/0.863,\ 711/0.671) = (824,\ 1060).$$

To simulate values from the posterior distribution of N requires only simulation of values from the posterior distribution of θ and then division of the known number of logbook reports filed by the simulated values of θ.

4 Estimating Effort

Both logbook and dockside sampling data sources recorded number of anglers and hours fished. The product of these quantities was taken to be an observed value of effort for the associated trip. In (1) κ is the expected value for effort across all trips in the population. To estimate this quantity, we assign a distribution to values of effort, specify a prior distribution for the parameters of that model, and derive the posterior distribution for the parameters (and hence also in principle the expected value, which will be a function of those parameters).

Modeling effort proved to be challenging because of the presence of extreme values. Some of these extremes could reasonably be considered errors, but others could not. For example, consider week 22. One trip in the logbook data (departed 6 pm on 2 June 2011, and returned at 6 pm on 5 June 2011) recorded 6 anglers and 72 hours fished for an effort of 422 fishing hours. It seems doubtful that these 6 anglers all fished for 72 hours in a period of 72 hours. The next largest value for effort in these data was 182 which resulted from 7 anglers fishing for 26 hours, also a suspect value. The next largest value for effort in logbooks for week 22 was 170.5, corresponding to a trip with 31 anglers fishing for 5.5 hours, which cannot be discounted as an actual possibility. Because of the difficulty of examining each record individually to determine plausibility of effort values, any logbook record with effort greater than 200 fishing hours was excluded from the estimation procedure used here. This was an arbitrary decision for the current analysis. The issue of extreme values (here for effort, but the same will be true for catch categories) is one that MRIP would be well advised to examine in greater detail if logbook data are accepted as a source of information for estimation in the for-hire fishery (see also item 3 in the Conclusions of Section 10).

Even with exclusion of extremely large values, it proved challenging to find a theoretical distribution that could adequately describe distributions of effort, which proved highly skew to the right but with small to moderate modes. A generalization of a gamma distribution was selected to model effort in this project. A random variable Y has a generalized gamma distribution if $Z = Y^\delta$ follows a standard gamma distribution. Using this model, the densities of Z and Y are, for parameters $\alpha > 0$ and $\beta > 0$,

$$\begin{aligned} g_z(z|\alpha, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} \exp(-\beta\, z); \quad z > 0, \\ g_y(y|\alpha, \beta) &= \frac{\delta\beta^\alpha}{\Gamma(\alpha)} y^{\delta\alpha-1} \exp(-\beta y^\delta); \quad y > 0. \end{aligned} \tag{7}$$

In (7) the parameters α and β are the same in $g_z(\cdot)$ and $g_y(\cdot)$.

As an illustration of the description this model provides for data on effort, let $\{Y_i : \ i = 1, \ldots, n\}$ denote random variables connected with trips taken in week 22, after deletion of the one extreme value of 422 (i.e., $n = 654$). Assume that these variables are independent and identically distributed following the densities given in (7) with $\delta = 0.5$. This value of δ was selected through some exploratory analysis as a value generally reasonable for effort in the available data, but certainly must be considered arbitrary in this example. Although not the estimates that will be used in the overall Bayesian estimator developed in this report, moment estimates of α and β based on transformed data $Z_i = y_i^{0.5}; \ i = 1, \ldots, n$ gave $\hat{\alpha} = 7.709$ and $\hat{\beta} = 1.445$. Using these estimated parameters, an estimated density $g_y(y|\hat{\alpha}, \hat{\beta})$ as in

(7) is plotted over a histogram of the observed effort values in Figure 2.

For use with the overall estimator, the model for effort needs to be updated in a Bayesian framework. To accomplish this, prior distributions are assigned to the parameters α and β of (7). Two different prior distributions were investigated, a product of independent normals each truncated at 0, and a bivariate normal truncated at $\mathbf{0}$. The posteriors obtained through the use of these priors were essentially the same. The bivariate prior was introduced in an attempt to improve the acceptance rate in a Metropolis-Hastings algorithm, but showed little improvement over the simpler product of independent normals, which are what is presented here. The joint prior for α and β in (7) is taken as a product form,

$$\pi_e(\alpha, \beta) = \pi_{e,a}(\alpha|m_a, v_a)\, \pi_{e,b}(\beta|m_b, v_b), \tag{8}$$

where

$$\begin{aligned}
\pi_{e,a}(\alpha) &= \frac{1}{I_a}(2\pi v_a)^{-1/2} \exp\left[-\frac{1}{2v_a}(\alpha - m_a)^2\right]; \quad 0 < \alpha < \infty, \\
\pi_{e,b}(\beta) &= \frac{1}{I_b}(2\pi v_b)^{-1/2} \exp\left[-\frac{1}{2v_b}(\beta - m_b)^2\right]; \quad 0 < \beta < \infty, \\
I_a &= (2\pi v_a)^{-1/2} \int_0^\infty \exp\left[-\frac{1}{2v_a}(\alpha - m_a)^2\right] d\alpha \\
I_b &= (2\pi v_b)^{-1/2} \int_0^\infty \exp\left[-\frac{1}{2v_b}(\beta - m_b)^2\right] d\beta
\end{aligned} \tag{9}$$

In (9) m_a, v_a, m_b and v_b are chosen values, not parameters to be estimated. These priors will typically be chosen to be rather diffuse, that is, with large values for the variances v_a and v_b.

The posterior distribution of α and β is,

$$p_e(\alpha, \beta|y_1, \ldots, y_n) \propto \pi_{e,a}(\alpha)\, \pi_{e,b}(\beta) \prod_{i=1}^{n} g_y(y_i|\alpha, \beta). \tag{10}$$

The posterior (10) is not available in closed form. A Metropolis-Hastings algorithm was used to simulate values from this posterior. A bivariate normal random walk was used to generate jump proposals in this algorithm, allowing elements of the covariance matrix to be adjusted in order to tune the algorithm.

To simulate from the posterior distribution of τ in (1) we need to simulate from the posterior distribution of the expected value of effort, for which a "double" simulation strategy was employed. In principle, the expected value of effort is some function of the parameters, $h(\alpha, \beta)$ say. If this function is applied to each of a set of draws from the posterior (10) one obtains draws from the posterior of the expected valued. But the function that gives expected values $h(\cdot)$ is difficult to obtain analytically due to the power δ in (7). Thus, for each draw of α and β from their posterior (10), a Monte Carlo approximation is obtained for expected effort as the average of an additional set of simulated values from the model (7). This is detailed in step 3 of the algorithm of Section 6.2.

5 Estimating Catch Per Unit Effort

Determination of a generally applicable theoretical model for cpue proved elusive. Fairly obviously, the values assumed by this variable depended on the definition of "catch" (e.g., harvest versus release at $<$ 120 feet, etc.), but ranges and empirical distributions of values also changed depending on the length of the time window used for estimation, the season of the year involved, and the species under consideration. An mild example is presented in Figure 3, which shows histograms of harvest cpue for red snapper in week 22 (upper panel) and week 27 (lower panel). At first glance these histograms may not appear greatly different, but the amount of probability placed at values greater than 3 differs substantially. Although the probabilities of larger cpue values is small in both weeks, the relative difference is great, being 1.8^{-3} for week 22 and 1.0×10^{-2} for week 27, which is nearly 6 times greater. Both weeks have a sufficient number of trips, with 565 logbook records in week 22 and 689 in week 27.

Various data models were investigated for use with cpue in different situations, including gamma, lognormal, generalized gamma, and extreme value distributions.

While continued examination of alternative distributions would be appropriate and is recommended (including the possibility of what are called "nonparametric" Bayesian density estimates) the approach taken here was to put all cpue values into categories and make use of multinomial distributions across those categories. A portion of the motivation for this approach is mathematical convenience, but the lack of precision of values for cpue in logbook records was also a contributing factor. This does necessitate defining a "largest category", which will vary depending on the quantity to which cpue applies (e.g., harvest versus released) and the species under consideration. An effort was made in this work, however, to avoid changing the cpue categories used when applying the estimator to time windows of different lengths. For example, the greatest value of harvest cpue in any logbook record for red snapper was 6.762, so the maximum cpue category was defined with an upper endpoint of 7.0. The same categories were then used in estimating total harvest of red snapper in any time period, regardless of what the maximum value was for that time period. Thus, the largest harvest cpue category retained positive probability even for estimation of week 22 in which the largest observed value was 3.667, although the probability of this category in the posterior would be reduced from what it was in the prior.

The model for cpue was then formulated as follows. Let $\mathcal{C} = \{(u_j, v_j] : j = 1, \ldots, k\}$ denote a set of k half-open intervals such that $u_1 = 0$, $u_j = v_{j-1}$ for $j = 2, \ldots, k$, and v_k is greater than the maximum observed cpue in the estimation problem under consideration. For example, in estimation of harvest cpue for red snapper, the categories were defined as

$$[0, 0.5],\ (0.5, 1.0],\ (1.0, 2.0],\ (2.0, 3.0],\ (3.0, 4.0],\ (4.0, 5.0],\ (5.0, 6.0],\ (6.0, 7.0]$$

Let $\{S_i : i = 1, \ldots, n\}$ denote random variables connected with the appropriate cpue (e.g., harvest or release) for trip i and define the categorical random variables, for $j = 1, \ldots, k$,

$$W_j = \sum_{i=1}^{n} I(u_j < S_i \leq v_j), \tag{11}$$

where $I(A)$ is the indicator function that assumes a value of 1 if A is true and a value of 0 otherwise.

The random variable $\boldsymbol{W} = (W_1, W_2, \ldots, W_k)^T$ (here, the superscript T is for transpose) is assigned a multinomial distribution. With parameters $0 < \eta_j < 1$, the probability mass function of $\boldsymbol{W}$ may be written as, for $\boldsymbol{w} = (w_1, \ldots, w_k)^T$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)^T$,

$$f_w(\boldsymbol{w}|\boldsymbol{\eta}) = \frac{n!}{w_1!\, w_2! \ldots w_k!} \prod_{j=1}^{k} \eta_j^{w_j}; \quad w_j = 0, 1, \ldots, n, \tag{12}$$

where $\sum_{j=1}^{k} w_j = n$, the number of trips in the relevant portion of logbook (or dockside sampling) records and $\sum_{j=1}^{k} \eta_j = 1$. Expression (12) is written as a function of k arguments $w_1, \ldots, w_k$ but defines only a $k-1$ dimensional probability distribution due to the constraint that the value of the sum of these arguments equal n, the known number of trips included in the estimation. For the Bayesian estimation strategy developed here this representation of the probability mass function of $\boldsymbol{W}$ as a form with less than full rank will not pose any difficulties (although it could cause difficulties if we were basing inference on asymptotic properties of, for example, maximum likelihood estimators).

The natural prior distribution to specify for the parameters $\boldsymbol{\eta}$ is the conjugate Dirichlet distribution having probability density function, for $\alpha_j > 0;\ j = 1, \ldots, k$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)^T$,

$$\pi_c(\boldsymbol{\eta}) = \frac{\Gamma(\alpha_1 + \ldots + \alpha_k)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\ldots\Gamma(\alpha_k)} \prod_{j=1}^{k} \eta_j^{\alpha_j - 1}; \quad 0 < \eta_j < 1, \tag{13}$$

such that $\sum_{j=1}^{k} \eta_j = 1$. In the same way that (12) defines only a $k-1$ dimensional distribution, the density in (13) is defined on the $k-1$ unit simplex in $k-$dimensional space. In this case, we know that the arguments sum to 1.0. The Dirichlet distribution may be considered an extension of the beta distribution of expression (3), to which it reduces if $k = 2$. As for other prior distributions, $\alpha_1, \ldots, \alpha_k$ in (13) will be specified numerical values, not parameters to be estimated.

Given the multinomial data model (12) and the Dirichlet prior (13), observation of values $\boldsymbol{w} = (w_1, \ldots, w_k)^T$ produces the posterior

$$p_c(\boldsymbol{\eta}|\boldsymbol{w}) = \frac{\Gamma(\alpha_1 + \ldots + \alpha_k + n)}{\Gamma(\alpha_1 + w_1)\Gamma(\alpha_2 + w_2)\ldots\Gamma(\alpha_k + w_k)} \prod_{j=1}^{k} \eta_j^{\alpha_j + w_k - 1}; \quad 0 < \eta_j < 1, \quad (14)$$

such that $\sum_{j=1}^{k} \eta_j = 1$. The posterior (14) may be recognized as a Dirichlet density with parameters $\{\alpha_j + w_j : \ j = 1, \ldots, k\}$. The posterior expected value of η_j is then, for $j = 1, \ldots, k$,

$$\hat{\eta}_j = \frac{\alpha_j + w_j}{\alpha_1 + \alpha_2 + \ldots + \alpha_k + n}. \quad (15)$$

As in the case of effort, what is needed to simulate from the posterior distribution of the population total τ in (1) are draws from the posterior distribution of expected cpue, ψ. Given simulated values from the posterior (14), these are obtained by direct calculation from (12) and the definition of W_j in (11) as detailed in step 4 of the algorithm outlined in Section 6.2.

6 The Estimation Algorithm

The entire procedure by which simulated values from the posterior distribution of τ in (1) are obtained is presented in this section in algorithmic form. This algorithm is presented under the assumption that an appropriate window in time and space has been identified for which estimation is desired and that window defines the extent of the data.

6.1 Input Data

Data needed for the estimation procedure are as follows:

1. The number of verified trips from monitoring or "prevalidation" sampling, denoted as m.

2. The number of verified trips for which a logbook report was filed, denoted as x.

3. The total number of logbook reports filed denoted as M.

4. Either logbook reports or dockside samples with the following information for each recorded trip:

   (a) number of anglers

   (b) hours fished

   (c) number of individual fish in relevant catch disposition category (harvest, released < 12), released > 120 or mortality)

From the data available in item 4, variables of effort $\{y_i : i = 1, \ldots, n\}$ and catch-per-unit-effort categories $\{w_j : j = 1, \ldots, k\}$ are created as discussed in Sections 4 and 5, respectively.

6.2 Algorithm

Given data available as outlined in the previous subsection, the estimation algorithm is as follows.

1. Choose prior parameters

   (a) Select α and β for the prior distribution of θ, the proportion of verified trips filing logbook reports. This prior is given in expression (3).

   (b) Select m_a, v_a, m_b and v_b for the prior distribution of the parameters of the generalized gamma data model (7) for effort. This prior is given in expressions (8) and (9).

   (c) Select $\alpha_1, \ldots, \alpha_k$ for the prior distribution of the parameters of the multinomial data model (12) for cpue. This prior is given in expression (13).

2. Simulate from posterior of number of trips.

   Draw a large number of values θ_m^*; $m = 1, \ldots, \mathcal{M}_1$ from the posterior (4). This is accomplished by simulating values from a beta distribution with parameters $\alpha + x$ and $\beta + m - x$. For each of these values, simulate an additional value from the posterior distribution of the total number of trips as $N_m^* = M/\theta_m^*$, where M is the known number of logbook reports.

3. Simulate from posterior distribution of expected effort.

   (a) Using the observed values of effort $y_1, \ldots, y_n$ and a Metropolis-Hastings algorithm, simulate $\mathcal{M}_1$ values (α_m^*, β_m^*); $m = 1, \ldots, \mathcal{M}_1$ from the posterior distribution (10). The construction of an appropriate Metropolis-Hastings algorithm is discussed in greater detail in the section on Technical Details in this report.

   (b) For each value (α_m^*, β_m^*) simulate $\mathcal{M}_2$ values $y_{m,q}^*$; $q = 1, \ldots, \mathcal{M}_2$ from a generalized gamma distribution (7) having parameters (α_m^*, β_m^*).

   (c) One value from the posterior distribution of expected effort is then given as,

   $$\kappa_m^* = \frac{1}{\mathcal{M}_2} \sum_{q=1}^{\mathcal{M}_2} y_{m,q}^*.$$

   The $\mathcal{M}_1$ values κ_m^*; $m = 1, \ldots, \mathcal{M}_1$ then constitute an approximate sample from the posterior distribution of expected effort.

4. Simulate from the posterior distribution of expected cpue.

   (a) Using the observed values for cpue categories $w_1, \ldots, w_k$ simulate $\mathcal{M}_1$ values $\boldsymbol{\eta}_m^* = (\eta_{1,m}^*, \ldots, \eta_{k,m}^*)^T$; $m = 1, \ldots, \mathcal{M}_1$ from the posterior (14). This is most easily accomplished through composition of simulated gamma variates. Additional discussion of this point is contained in the section on Technical details in this report.

(b) For each value $\boldsymbol{\eta}_m^*$, a value from the posterior distribution of cpue is obtained as

$$\psi_m^* = \sum_{j=1}^{k} \eta_{j,m}^* \frac{u_j + v_j}{2},$$

where $(u_j,\, v_j]$ are values used to define the cpue categories and random variables in expression (11). The $\mathcal{M}_1$ values $\psi_m^*;\;\; m = 1, \ldots, \mathcal{M}_1$ then constitute an approximate sample from the posterior distribution of expected cpue.

5. At the completion of steps 1-4, one has $\mathcal{M}_1$ values simulated from the posterior distributions of number of trips N_m^*, expected effort κ_m^* and expected cpue ψ_m^*. From (1) then, simulated values from the posterior distribution of the desired population total as, for $m = 1, \ldots, \mathcal{M}_1$

$$\tau_m^* = N_m^* \kappa_m^* \psi_m^*. \tag{16}$$

7 Application to Red Snapper

To illustrate the estimation procedure developed in this report, the algorithm summarized in Section 6 was applied to the harvest of red snapper during the season in 2011 which ran from 1 June 2011 through 18 July 2011. This includes part or all of weeks 22 through 29. Estimates will be produced for each week separately, and also as a cumulative total in each week of the season. Week 22 and week 29 contained days that did not fall into the red snapper season in 2011, and so these "weeks" were shortened to include only days that were included in the season. Nevertheless, it should be noted that both logbook reports and dockside samples included some positive values for red snapper harvest for days that were not officially part of the season. For example, logbook reports for trips listed as taken in May 2011 indicate that about 150 red snapper were harvested on 13 trips in this month. There is a legitimate question as to how such records should be handled. For the illustration here they are simply ignored.

Along with illustrating the use of the proposed estimator, an objective of this application is to compare the use of logbook data and dockside sampling data as sources of information on effort and cpue. In both cases, estimation of the total number of trips taken will rely on use of the prevalidation or activity monitoring data gathered as part of the pilot study. We will step through the algorithm of Section 6.2, once using logbook data to provide information on catch, and once using dockside sampling to provide catch information. Estimation of the number of trips is common to estimation with both of these data sources.

7.1 Data

The number of verified trips, the number of those that had corresponding logbook reports, the total number of logbook reports filed, and the number of dockside samples taken are given for the weeks of the 2011 red snapper season in Table 1. The columns of this table gives the variables m, x and M in the first three items listed as necessary input data in Section 6.1. The individual trip records from the logbook reports and the dockside samples provide values listed in item 4 of Section 6.1. For the logbook data source $M = n$, that is, the total number of logbook reports is the same as the sample size for estimation of effort and cpue. For the dockside sampling data source the sample sizes for catch data are considerably smaller. Over the entire season there were 4760 trips with logbook reports and 289 dockside samples. Note that estimation in week 29 was not possible for the dockside data source as there were no dockside samples taken before 19 July in that week.

7.2 Stepping Through the Algorithm

The steps of the estimation algorithm of Section 6.2 are considered in order.

1. Selection of Prior Parameters

(a) The parameters for the prior distribution of θ were chosen as $\alpha = 1$ and $\beta = 1$. While more sophisticated choices would be possible, particularly for weeks greater than 22, the number of verified trips is sufficiently great for all weeks (with the possible exception of week 28) that the prior will have little effect on estimation of the total number of trips taken.

(b) The prior distribution for parameters of the generalized gamma data model of effort was selected to have parameter values of,

$$m_a = 6.0 \quad v_a = 50.0 \quad m_b = 1.5 \quad v_b = 25.0$$

These values were selected based on tuning of the MH algorithm used for simulation of the posterior distribution for parameters of the effort data model, as described in Section 8.1 on Technical Details.

(c) The multinomial data model for harvest cpue was formulated with 8 categories as given immediately prior to expression (11). The prior distribution for the parameters of the multinomial data model was chosen as a Dirichlet with parameter values

$$\begin{array}{llll} \alpha_1 = 0.5000 & \alpha_2 = 0.4000 & \alpha_3 = 0.0800 & \alpha_4 = 0.0130 \\ \alpha_5 = 0.0045 & \alpha_6 = 0.0015 & \alpha_7 = 0.0004 & \alpha_8 = 0.0006 \end{array}$$

These values were selected on the basis of some exploratory analysis with harvest of red snapper over the entire season. While selection of prior parameter values based on examination of the data is generally frowned upon because it can lead to under-assessment of uncertainty, in this proof-of-concept exercise it is not entirely without merit. Sensitivity analysis indicates that the selection of these prior parameter values can have some effect (more so than other prior choices) on the final estimator. This effect is more pronounced for estimation of releases and mortalities than it is for estimation of harvest, but assigning too large a parameter for the larger cpue categories, combined with relatively small data availability (such as might occur in some weeks or even months) can lead to estimates that are highly effected by the prior probabilities of large cpue categories. It is only in this initial exercise, however, that no information is available to guide the selection of this prior distribution. That is, using values selected on the basis of data gathered in this pilot study for analysis of any subsequent data would not suffer the potential effects of "using the data twice", as the pilot study will be previous data for any actual application of these estimators. It should also be noted that final estimation results become dramatically less sensitive to the selection of these prior parameter values as the volume of available data increases (i.e., as the time period of estimation increases).

2. Simulate from Posterior of Number of Trips

   Parameters of the posterior distributions of the proportion of trips filing logbook reports in each week of the red snapper season are presented in Table 2. As an illustration of what a sample from one of these posterior distributions looks like, 50000 values were simulated from the posterior distribution of the proportion of trips filing logbook reports in week 23, and these were used to produce values from the posterior distribution of the total number of trips, as described in Section 3 and step 2 of the algorithm of Section 6.2. Histograms of these values (proportion reporting in upper panel and total number of trips in lower panel) are presented in Figure 4. For the histogram of 50000 values simulated from the posterior distribution of the proportion of trips reporting in the upper panel of Figure 4, the theoretical posterior is overlaid as a solid curve – this is the same distribution as depicted in Figure 1, although the scales of the horizontal axis differ

3. Simulate from Posterior of Expected Effort.

(a) Construction of an appropriate Metropolis-Hastings algorithm to sample from the posterior of the effort data model parameters is perhaps the most technically involved step in the estimation procedure. More extensive discussion of this issue is contained in Section 11. Continuing to use week 23 of 2011 as an example, 50000 values of α and β were simulated from the posterior (10) using the MH algorithm described in Section 11.1 with a bivariate random walk jump proposal distribution. The simulated values are presented in Figure 5, values of α in the upper panel and values of β in the lower panel.

(b) Although not obvious from the histograms of Figure 5, values of α and β are simulated from the joint posterior as pairs of values and, as eluded to in Section 11, tend to be highly correlated in the posterior distribution. For each pair of values simulated from the posterior, $\mathcal{M}_2 = 10000$ values were then simulated from the data model (7) using those values as parameters.

(c) For each of the 50000 pairs of values α and β simulated from their joint posterior in step 3(a), the average of the 10000 values simulated from the data model with those parameter values was computed, yielding 50000 values of expected effort. This computationally intensive procedure resulted in the posterior distribution of expected effort presented in Figure 6, which has a mean of 31.1 and a small standard deviation of 0.86. It is the values of this distribution that are the κ_m^* in step 3c of the algorithm of Section 6.2.

4. Simulate from Posterior of Expected Catch per Unit Effort.

(a) Simulation from the posterior distribution of expected cpue will also be illustrated with week 23 of 2011, now making use of red snapper harvest as the pertinent response. The observed frequencies with which trips fell into the various cpue categories used for this response are presented in Table 5 for both logbook and dockside sampling data sources. Using these values as the w_j, 50000 sets of values $\{(\eta_{1,m}, \ldots, \eta_{8,m} : \ m = 1, \ldots, 50000\}$ were simulated from the Dirichlet posterior (14). These simulated values were produced using the composition of gamma variates outlined in Section 11.2.

(b) For each of the 50000 values simulated in step 4(a), the expected cpue was computed as in step 4(b) of the algorithm of Section 6.2. The values of u_j and v_j are the lower and upper endpoints of the intervals in Table 5, which are listed as $(u_j, v_j]$. Histograms of simulated values from the posterior distributions of cpue for logbook (upper panel) and dockside sampling (lower panel) data sources are presented in Figure 7. These histograms are plotted on the same horizontal scale for purposes of comparison. Visually, it appears as if the posterior distribution of harvest cpue for dockside sampling data is shifted to the right (larger values) than is that based on logbook data, and this is true. But the more substantial difference between these posterior distributions is in the amount of spread. The mean for the distribution based on logbook data is 0.53 while that for dockside sampling data is 0.63. The standard deviations, however, are 0.019 for logbook and 0.052 for dockside data, the value for dockside sampling data being 2.8 times larger.

7.3 Initial Results for Red Snapper

Running the entire algorithm for each week of the red snapper season separately produces the results presented in Table 6 and the upper panel of Figure 8. In separate estimation for each week of the red snapper season, estimates based on logbook data and dockside sampling data compare favorably. In only one week (week 27) did 95% credible intervals fail to overlap. While not the most sensitive assessment for differences, this does indicate that estimates based on the two data sources are generally not highly discrepant. The mean absolute difference in point estimates (as posterior means) was 2760. The credible intervals are noticably more narrow for estimates based on logbook data than for those based on dockside samples. This is because many more logbook reports than dockside samples are available. Although sample size does not enter the procedure for determining these intervals in the same manner as for confidence intervals for means (as taught in most introductory statistics classes) the effect is similar – the more data are avaliable, the less uncertainty exists in estimates.

Cumulative estimates across the 2011 red snapper season are presented in Table 7 and the lower panel of Figure 8. Estimates across weeks are not necessarily additive in the sense that the sum of the estimates of Table 6 will not be equal to those of Table 7, although they should not be highly discrepant unless the fishery is extremely volatile from week to week. That this does not appear to be the case is evidenced by the fact that cumulative sums of mean values from Table 6 are actually quite similar to the values of Table 7. For example, in week 22 estimation based on logbook data returns point estimates of harvest of 13238 and 13242 in Tables 6 and 7, while the corresponding values for estimation based on dockside samples are 13675 and 13688. These are estimates of the same quantity, given that the cumulative harvest at week 22 was the harvest for week 22. The values from the two runs of the simulation-based estimation procedure differ by only 4 and 13 individuals for logbook and dockside data sources, respectively. It appears from both Table 7 and Figure 8 that cumulative estimates based on logbook reports increase over time more rapidly than do the values based on data from dockside samples, for this particular application at least. Credible intervals do overlap for estimates based on the two data sources for all time points for which both are available (up to week 28). It would be premature to conclude from this one example that there will be a consistent difference between the data sources. What is important is to judge whether the two data sources are giving sufficiently similar values to consider logbook data as a reliable source of information for the purposes of estimation.

7.4 Alternative Results for Red Snapper

Throughout this exercise in developing estimators for population totals of important quantities in the for-hire or charter recreational fishery in the Gulf of Mexico, the most challenging statistical task has been developing models that can accommodate extreme values in distributions of effort and/or catch (defined relative to harvest or releases or mortalities). Extreme values of these variables appear in logbook reports and, to a lesser extent, also in dockside sampling records, and can have a strong influence on estimated values. Several decisions about "acceptable" data records have been previously indicated in this report (e.g., deletion of records with effort greater than 200 fishing hours, see the discussion of Section 4). As work on examination of the proposed estimator progressed, concern about these extreme values increased. If real, such values are critical to the determination of total harvest or releases. If not real, however, such values can lead to substantially biased and misleading estimates.

Some records, especially logbook records, are clearly suspect in accuracy. For example, a logbook record for a trip taken on 10 July 2011 listed 35 anglers, 3 fishing hours, and red snapper harvest of 710. This amounts to roughly 20 red snapper per angler during a season with an individual bag limit of 2 red snapper (one might suspect this was supposed to be 71 fish harvested, but that is not what was in the data and so the value cannot just be changed by someone conducting data analysis). The cpue for this trip was 6.76, the highest value of any trip. This one record had a substantial influence on simulated values of expected cpue and, hence, on the overall estimate of total harvest for this week. The legitimacy of other records, however, was less clear. For example, a trip taken on 6 June 2011, listed 11 anglers fishing for 2 hours and having a harvest of red snapper of 24, which is only 2 fish over the sum of individual bag limits. If crew members also fished and were not counted in the number of anglers reported, this is entirely possible. Suspect records are more frequent in logbook data than in dockside data, but some dockside sampling records also raise concern. For example, a dockside sampling record for a trip taken on 16 July 201 reported 3 anglers fishing for 4 hours and 12 harvested red snapper, so that each angler harvested 4 fish, twice the daily bag limit.

The point of the previous paragraphs is that a decision process for determining which logbook (and dockside sampling) records to accept as legitimate data needs to be determined. This will be true regardless of whether the estimator proposed in this report is adopted, or some other estimator is used. Included in this need is a decision process for how to handle trip records (either logbook or dockside samples) that include harvest of species at times outside the open season such as previously noted for red snapper at the beginning of Section 7.

To illustrate the difference that using a rule to eliminate suspect data records can have, both logbook and dockside sampling data sources had trips eliminated for which red snapper harvest was greater than twice the number of anglers. While this may have eliminated some "legitimate" trips, it also seemed to eliminate all of the records with potentially inflated cpue that one might identify as suspect by inspection. Of the 4760 logbook records filed during the red snapper season, 658 were eliminated, leaving 4102 data records for estimation. Of the 289 dockside samples taken during the red snapper season, 36 had greater harvest than allowed by this rule and were eliminated, leaving 254 records for estimation. The results of estimation with these reduced data sets are presented in Tables 8 and 9 and Figure 9. Estimates based on logbook data and dockside sampling data are now even more similar than previously (i.e., Tables 6 and 7 and Figure 8). This would suggest that for the unthinned data, dockside and logbook data sources differ primarily in terms of the number of extreme values for cpue that occur. The estimated harvests are also decreased for estimation based on the thinned data. For example, cumulative harvest at the end of the red snapper season was estimated as 118796 using logbook data (see Table 7) and now with the thinned logbook data is estimated as 91890, a difference of 26906 which is a substantial amount of difference. The estimate of total harvest based on dockside samples also decreased, but not as dramatically, changing from 90067 at the end of week 28 (as close as one can get to the end of red snapper season with dockside sampling data) to 78479, a difference of 11588.

The importance of the question posed at the end of the opening paragraph in this subsection should now be clear. What appear to be exceptionally large (or extreme) values for effort and/or cpue can exert a good deal of influence on estimates. If these values are accurate it is important to use and correctly model them. If these values are not accurate (for whatever reasons), their inclusion in the data used for estimation will give rise to misleading estimates.

8 Application to Vermilion Snapper

Subsequent to the original application to red snapper, the project team for the MRIP logbook reporting pilot study requested an additional application to a species without the stringent individual bag limit and short open season that exist for red snapper. The project team selected vermilion snapper as such a species. In the Gulf, there is no closed season for vermilion snapper, and the daily bag limit is 20 (along with a total of 20 for all reef species).

8.1 Data

Given that vermilion snapper may be harvested all year in the Gulf, estimation for individual weeks of a season, such as conducted with red snapper in Section 7, is not feasible. Instead, estimates were produced for individual months, aside from January and February which were combined, and November and December, which were also combined. The number of verified trips by month (from the activity monitoring or "prevalidation" portion of the pilot study) and the number of those filing logbook reports are presented in Table 10.

For red snapper, the number of trips during the season that did not have any harvest was extremely small, and was absorbed into the smallest category of cpue (which was [0, 0.5]). In contrast, a much larger number of trips throughout the year reported no harvest of vermilion snapper. The number of logbook reports, the number of dockside samples, and the number of each of those with nonzero harvest of vermilion snapper are reported in Table 11. The proportion of logbook records having harvest of vermilion snapper ranges from 0.25 (October) to 0.56 (May and August), while the proportion of dockside samples with harvest range from 0.27 (April) to 0.88 (Jan/Feb) (but note that this is the result of only 8 trips, a better upper value is 0.66 in May). The point is that a substantial proportion of trips record no harvest of vermilion snapper throughout the year, which has a decided effect on expected cpue.

Drawing on the experience of analyzing red snapper harvest, records (either logbook or dockside samples) were deleted if (1) hours fished was less than 1, (2) effort was greater than 200 angler hours, or (3) harvest divided by number of anglers was greater than the individual bag limit of 20.

8.2 Modifications to the Estimator

Because of the large number of trips with no harvest of vermilion snapper, the estimator of the form (1) was modified to follow

$$\tau = N\,\kappa\,\lambda\,\psi, \tag{17}$$

where N is the total number of trips and κ the expected effort as before, λ is the proportion of trips having non-zero harvest of vermilion snapper, and ψ is the expected cpue *given non-zero harvest*. The proportion of trips having harvest of vermilion snapper (λ in (17)) is modeled in exactly the same manner as the proportion of verified trips having corresponding logbook reports (see Section 3). To make this explicit, let s denote the total number of trips for which records are available. For months, this will be either the column labeled "Logbook Total" or "Dockside Total" from Table 11, depending on which data source is being used. Also, let Z be a random variable connected with the number of trips having harvest of vermilion snapper (observed values are given in the "With Harvest" columns of Table 11). Assign Z the probability mass function

$$f_h(z|\lambda) = \frac{s!}{z!\,(s-z)!}\,\lambda^z\,(1-\lambda)^{s-z}; \quad z = 0, 1, \ldots, s. \tag{18}$$

In the same way that θ of (2) was assigned a beta prior (3), so too is λ. That is, a prior distribution for what we believe about the value of λ is specified as

$$\pi_h(\lambda) = \frac{\Gamma(\alpha_h + \beta_h)}{\Gamma(\alpha_h)\,\Gamma(\beta_h)}\,\lambda^{\alpha_h - 1}\,(1-\lambda)^{\beta_h - 1}; \quad 0 < \lambda < 1, \tag{19}$$

for some chosen values of α_h and β_h. In this application a prior uniform on the interval $(0, 1)$ was specified by setting $\alpha_h = \beta_h = 1$.

In the same manner as for the probability that a verified trip has a corresponding logbook report, the posterior distribution of λ in this model is again a beta distribution with parameters $\alpha_h + z$ and $\beta_h + s - z$.

The algorithm of Section 6.2 is then modified by adding a Step 2A as

2A. Simulate from posterior of probability of harvest.

    Draw a large number of values $\lambda_m^*; m = 1, \ldots, \mathcal{M}_\infty$ from the posterior distribution of λ.

Values are simulated from the posterior distribution of expected cpue exactly as in Step 4 of the algorithm outlined in Section 6.2, except that the observed values for cpue categories are taken only from data records that have non-zero harvest of vermilion snapper.

8.3 Results for Vermilion Snapper

Results of estimation for harvest of vermilion snapper by month are presented in Table 12 and the upper panel of Figure 10. Cumulative results over months are presented in Table 13 and the lower panel of Figure 10. Recall, for interpretation of this figure, that the pilot study began collecting data in September of 2010 and concluded in August of 2011. Thus, the months $1 - 12$ in Figure 10 and Table 13 and not one calendar year, but span two years. In addition, it should be kept in mind that the red snapper season was re-opened in October of 2010 (at least on weekends) and this may have impacted the level of fishing activity for that month (which is reflected, for example, in the number of logbook reports for October in Table 11). Also, it was decided that the number of logbook reports in November (319) and December (54) were sufficient to warrant separate updates of these months in the cumulative total (although not enough for individual estimation using only those months) but the same was not true for dockside samples which numbered 32 in November and only 1 in December.

Overall, the results of Tables 12 and 13 and Figure 10 indicate that the harvest of vermilion snapper (number of individuals) in the portion of the fishery covered by the data sources is about twice that of red snapper on an annual basis (roughly 12000 for red snapper and 250000 for vermilion snapper). Secondly, there is considerably more variability in estimates for vermilion snapper than there was for red snapper. And third, while agreement between logbook and dockside sampling data sources is still reasonable, there are greater differences than were seen with red snapper. In particular, a discrepancy arises between estimates based on the two data sources in July, and this discrepancy is sufficient to cause a difference in cumulative estimates for the remaining months of the year (lower panel of Figure 10), although the individual months after July do not continue to give different results (upper panel of Figure 10).

The cause of the discrepancy between the estimate of July harvest based on logbook data and that based on dockside sampling data appears to have been in expected cpue. Of the 1269 logbook records for July with non-zero harvest, 14 cpue values (0.11%) were greater than 5. Of the 129 dockside sampling records with non-zero harvest, 1 (0.08%) was greater than 5. So, the dockside and logbook data sources are reasonably in concert regarding the frequency of large cpue values. But, the one value for dockside samples was a cpue of 5.2, while logbook records contained cpue values between 6 and 7 (5 such values), between 7 and 8 (3 such values) and between 8 and 9 (3 such values). These observations increased the probabilities for larger cpue categories enough to make expected cpue substantially greater for the analysis based on logbook data than for that based on dockside sampling data. Months other than July did not contain as many large cpue values in the logbook data, and estimates are more similar between the two data sources. For example, logbook records for June contained only 6 records with cpue over 5 out of 1364 with non-zero harvest (or 0.4%). The estimate based on dockside samples is, again, lower than that based on logbook records, but not to the extent that ocurred in July. As indicated in the first paragraph of Section 7.4, a few quite large values for cpue can have a good deal of influence on the estimator. If these records are accurate it is important to detect such occurrences.

The modification of the basic estimator for use with vermilion snapper was to include a term for the probability that this species is harvested on a trip. For red snapper (during the red snapper open season) this probability was not only high, but consistently high over weeks, and any effect of trips that failed to harvest red snapper was incorporated in the distribution of cpue. For vermilion snapper, however, this probability was not necessarily high, and it is valuable to examine how much it may have varied over months. Figure 11 presents posterior means and 95% credible intervals for the proportion of trips with harvest, by month (i.e., posterior means and credible intervals for the value of λ in (17). There is a considerable amont of variability over months exhibited in these values, and there appear to be groups.

In particular, the months of May, June, July, and August appear similar to each other and to form one group, while the other months could be placed in a second group (recall from Table 11 that the dockside value for this proportion in Jan/Feb is based on only 8 samples and the logbook record is also obtained from the smallest number of logbook reports filed in any two-month period). Alternatively, one could view Figure 11 as having 3 groups, January through April, May through August, and September through December. The estimator could potentially be improved by taking these groups as defining "strata" in the population under consideration. This would then result in an estimator for an overall population total as of a sum of pieces, each of which would have the form of (17) similar to an estimator proposed for discard in the northeast groundfish fishery (see an MRAG report titled "Development of an Estimator of Discard for the Northeast Observer Program", dated January 2006). A key point is that in a simulation-based procedure for estimation and inference based on posterior distributions, splitting an estimator for a total into a sum over sub-populations introduces little in the way of additional complexity, although it does increase computational burden to some extent and certainly renders tasks of "book-keeping" more involved.

9 Further Development

The methodology proposed in this report departs from approaches that NMFS has relied on in the past. Although I have no first-hand knowledge of efforts by MRIP to this effect, information available on the web seems to indicate that methodologies used in estimation of recreational catch and mortality have been undergoing considerable scrutiny in recent years. It is difficult, however, to discern whether these efforts are actually targeted at producing a coherent comprehensive estimation program, or simply focus on individual components of the problem such as estimation of effort only. As indicated in Section 2 of this report, derivation of a justifiable measure of uncertainty for estimators of population totals based on the theory of survey sampling can be difficult when those totals are estimated from data that arises from a number of disparate sampling plans. A major strength of the estimator proposed in this report is that defensible interval estimates (credible intervals) are produced automatically as part of the estimation procedure. The statistical procedures relied on in this report are not “experimental” or “unproven”, although they do require a certain level of statistical and mathematical sophistication to apply with confidence.

Although the proposed estimator offers a defensible estimation strategy for population totals and the associated uncertainty that could be used as soon as suitable data become available, there are certainly aspects of the procedure that need additional investigation and development. The more obvious of these are listed here.

1. If logbook data are to be used to provide catch information for the estimator, a review of which vessels qualify as part of the charter sector of the recreational fishery is critical. It is a reasonable presumption, although not a certainty, that inclusion of large capacity boats produce values for effort, and possibly cpue, that are not in concert with the majority of trips taken in this sector of the fishery. Knowingly leaving these vessels in this same portion of the fishery amounts to introducing a source of heterogeneity *on purpose*, which cannot be defended in a serious estimation program.

2. The capacity of MRIP and its affiliated agencies for data cleaning and preliminary examination needs additional attention and development. A few anomalous occurrences have been mentioned in this report relative to recorded effort (e.g., Section 4), but there are suspect data records in other areas as well, particularly for the logbook data source. For example, there were logbook records in 2011 that reported positive values of harvest for red snapper in February (1 trip), March (5 trips), April (7 trips), May (15 trips) and August (4 trips). There were trips both leaving and returning on the same day that recorded greater than 24 hours fished. There were a few trips that reported hundreds of

red snapper released (e.g., 400, 150, 100) some with much smaller harvest (either in or out of season) and some with no harvest. These extreme values have a decided influence on estimation which, if they are correct, is important to include but, if they are not correct, can produce misleading results. Section 7.4 illustrates how estimates could change depending on how records are judged as legitimate or not to use in an estimation procedure. It should be noted that this is not unique to the estimator proposed in this report. Nearly any estimator using logbook or dockside sampling data will be subject to similar effects.

3. The component of the proposed estimator that provides the greatest scope for improvement is modeling of cpue. This is particularly true for "catch" that corresponds to releases or mortalities. To a greater extent than harvest, releases tend to have distributions that are governed by a small number of extreme values, which can increase the estimated value of expected cpue by an amount sufficient to cause dramatic shifts in the estimation of population totals when multiplied by the (estimated) number of total trips taken. The previous point is again highly relevant to this issue.

4. If the estimator is adapted for a range of species, attention will need to be given to defining cpue categories and specifying prior parameters for the probabilities of those categories.

5. It was mentioned in the last paragraph of Section 2 that the potential exists for induced correlation between the variables of effort and cpue. This does, in fact, appear to be the case for these variables for individual trips. Figure 9 presents a scatterplot of cpue against effort for trips taken during the red snapper season of 2011. Linear correlation is not strong in these data, -0.31 for logbook values and -0.32 for values from dockside samples, but there is a typical hyperbolic relation that often results from defining two variables in

this manner.

A question is to what degree this relation between effort and cpue should be of concern for use of the proposed estimator. The estimators (1) and (17) multiply values of number of trips, expected effort, and expected cpue (and, in the case of (17) the probability of harvest) that are all simulated from independent sources. It is not entirely clear what relation between expected effort and expected cpue might be induced as a result of the induced relation between effort and cpue on individual trips. The primary effect of ignoring any relation in the simulation of independent values of expected effort and expected cpue for construction of the estimator as in (1) should be to increase variability. One could argue, then, that ignoring this relation should tend to result only in an overestimate of uncertainty and thus produce a more conservative procedure. Nevertheless, the fact that there is demonstrable induced correlation between effort and cpue at the level of individual trips, and the fact that this is not accounted for in the proposed estimator constitutes an aspect of the estimation strategy that it would be desirable to improve on.

10 Conclusions

The work that forms the basis for this report was motivated by the conclusion of the MRIP project team for the pilot study of a logbook system for the charter fishery in the Gulf of Mexico that a logbook system could not reasonably be made to function as a census. The project team also concluded that logbook records could not be considered equivalent to dockside samples at the level of individual trips, but that aggregated logbook and dockside values might be similar enough to motivate the use of logbook data in an estimation procedure. This report is the result of developing one such estimator, demonstrating its potential, and comparing what would result from its use with logbook data versus dockside sampling data to provide information on catch and catch disposition. The following conclusions are warranted.

1. It is possible to develop estimators that provide a coherent strategy for assessing catch in the Charter fishing sector of the Gulf of Mexico fishery. Although such material may exist, the author of this report has seen no documentation on other estimators that produce both population totals and justifiable quantification of uncertainty (here in the form of intervals) despite having made several requests to see such documentation. The estimator proposed here can produce estimates for any window in space and time large enough to provide sufficient data (which is related to the previous conclusion that logbook and dockside reports should not be taken as equivalent for individual trips can do appear similar in aggregate). Depending on the time allowed for filing logbook reports, estimates for some species of particular concern could be produced on a weekly basis during peak fishing periods, as illustrated in Section 7. Other species with different regulatory status may require temporal windows of one or even multiple months, such as the case of vermilion snapper considered in Section 8.

2. The proposed estimator returns results from the use of logbook data that compare favorably with those obtained from the use of dockside sampling data. Interval estimates resulting from the use of logbook data are more narrow than those resulting from the use of dockside samples, due to the greater number of logbook reports available.

3. There is a pressing need for NMFS and/or MRIP to set standards for detecting and dealing with extreme values in data. This is true for both logbook and dockside sources, but is especially needed if logbooks are to be used as sources of information for estimation. While there may be some dockside sampling records that appear suspect, the frequency is much higher and the extremes more extreme in logbook reports.

Two final points that might be made are motivated as responses to the misconceptions that a faulty census is somehow to be preferred to a scientifically sound estimation procedure, and that estimation procedures that might be improved upon are somehow not sound. Both of these opinions are simple fallacies. Along with illustrating the value of random sampling, the failure of the Literary Digest 1936 presidential poll (of 10 million people, with 2.4 million returns) relative to the Gallup poll of about $50,000$ people demonstrates the misguided belief that quantity at the cost of quality is a meaningful goal. Similarly, the belief that there is "A Correct" statistical estimator for a problem is a misconception of what statistical analysis is all about. Statistical methods and estimators for any but the most simple of problems undergo continual examination and attempts at improvement. Methods used in analysis of what is now called the USDA National Resources Inventory (NRI), for example, have been undergoing continual development and improvement since 1956, yet no one takes this to mean that results produced in 1985 should have been dismissed in development of the Farm Bill in that year (one of many uses made of estimates produced in the NRI). The points of all of this for estimation of catch in the for-hire Gulf fishery are that logbooks should be viewed as a sampling device rather than an imperfect census, and that the fact an estimator leaves room for additional development should not be taken as a reason for rejecting its use out of hand. The use of an estimator such as that proposed in this report in conjunction with a logbook reporting system, an ongoing activity monitoring program, and periodic validation of logbook reports represents one viable option for approaching the assessment of the for-hire fishery in the Gulf of Mexico. Initial work on design and sample sizes needed for a monitoring program and a periodic validation effort for logbook data has been documented in previous MRAG reports as referenced in the final report of the pilot study project team.

11 Technical Details

This section contains some technical details that should be of use to those attempting to adapt the algorithm presented in this report to different species or situations.

11.1 MCMC for Expected Effort

To simulate from the posterior distribution of expected effort requires, first, simulation from the posterior distribution of parameters α and β in the generalized gamma data model for effort. The data model is given in (7), a prior consisting of two truncated normal distributions in (9) and the resulting posterior in (10). The posterior of α and β may be simulated from using a Metropolis-Hastings (MH) algorithm, but this intermediate step in obtaining values from the posterior distribution of the population total τ in (1) is perhaps the most involved portion of the overall algorithm and thus merits some additional discussion.

At iteration t of a MH algorithm let $(\alpha_t,\, \beta_t)$ denote the current state of the chain (current values of α and β). A proposed jump is produced as (α^*, β^*) from a proposal distribution $q(\cdot)$ (an effective proposal distribution for making these proposals is described later in this subsection). A new value $(\alpha_{t+1},\, \beta_{t+1})$ is determined as follows:

$$(\alpha_{t+1},\, \beta_{t+1}) = \begin{cases} (\alpha^*,\, \beta^*) & \text{with probability } \lambda \\ (\alpha_t,\, \beta_t) & \text{otherwise .} \end{cases} \tag{20}$$

In (20),

$$\lambda = \min\left\{1,\, \frac{p(\alpha^*, \beta^*|\boldsymbol{y})\, q(\alpha_t, \beta_t|\alpha^*, \beta^*)}{p(\alpha_t, \beta_t|\boldsymbol{y})\, q(\alpha^*, \beta^*|\alpha_t, \beta_t)}\right\}. \tag{21}$$

where $\boldsymbol{y} = (y_1, \ldots, y_n)^T$ and $p(\alpha, \beta|\boldsymbol{y})$ is given in (10). As usual, the normalizing constant needed to make $p(\cdot)$ a density cancels in this ratio so that the right hand side of (10) may be used in place of $p(\cdot)$ in (21).

A random walk jump proposal proved effective for generating values $(\alpha^*,\, \beta^*)$ in (20) and (21). Initially, independent normals were used for this proposal. The acceptance rate of such proposals was quite low (generally $< 5\%$) and values of α and β simulated from the posterior were highly correlated (generally > 0.90). The acceptance rate and overall performance of the algorithm was improved by using a bivariate normal proposal distribution with a fairly high degree of correlation in the jump steps. In particular, let (z_a^*, z_b^*) denote a value simulated from a bivariate normal distribution with expected value $\mathbf{0}$ and covariance matrix Σ. The proposed jump was then

$$\alpha^* = \alpha_t + z_a^* \qquad\qquad \beta^* = \beta_t + z_b^*$$

This jump proposal is symmetric so that the acceptance probability λ in (21) simplifies to contain only the ratio of target densities $p(\cdot)$.

An appropriate burn-in period for the MH chain to simulate values from the posterior of α and β was determined through inspection of trace plots for dispersed starting values and computation of the scale reduction factor Gelman and Rubin (see nearly any modern text on Bayesian analysis for an explanation of this diagnostic). As an illustration, consider estimation of effort in week 23 (the first full week of the 2011 red snapper season) based on logbook data. There was one effort value in the logbook data that exceeded 200 (effort of 320) and that record was deleted from the analysis as explained in Section 4 of this report. Tuning of the MH algorithm to produce an acceptance rate of 21% resulted in the selection of the jump proposal covariance matrix as

$$\Sigma = \begin{pmatrix} 0.25 & 0.10 \\ 0.10 & 0.05 \end{pmatrix}.$$

Trace plots of three chains started at widely varying values of α and β are presented in Figure 4 for the first 1000 iterations of the chains. Mixing appears adequate by iteration 400 to 500 for both parameters (scale reduction factors had stabilized near 1.0 by iteration 300 for both α and β). Conservatively doubling this amount, the burn-in period was set to 1000 iterations.

Running an additional chain with a burn-in of 1000 iterations and collecting the next 50000 values produced the posterior distributions for α and β presented in Figure 5. Posterior expected values and 95% credible intervals are shown in Table 3.

The posterior predictive distribution of effort can be used to assess the appropriateness of the model and prior distribution. This is not a part of the overall algorithm presented in Section 6 because it is not necessary for each application. But, model assessment is important in development (and should be conducted periodically in application as well), particularly those for effort and cpue. A posterior predictive distribution of effort can be simulated using the following steps. For each value (α_t, β_t) simulated from the posterior distributions of these parameters,

1. Simulate one value z_t^* from a gamma distribution with parameters (α_t, β_t)

2. Let $y_t^* = (z_t^*)^{1/\delta}$, where δ is the power used to define the generalized gamma distribution in (7). For harvest of red snapper, δ was chosen as $\delta = 0.50$.

Simulating 50000 values from the posterior predictive distribution of effort in week 23 of 2011 resulted in a mean of 31.6, compared to the observed average of 32.0 from logbook data. Five number summaries of observed and posterior predictive values are presented in Table 4. The posterior predictive appears to give a reasonable fit to these data. The 90^{th} and 95^{th} percentiles of the observed data were 60.0 and 78.0, respectively, while those same percentiles from the posterior predictive distribution were 59.4 and 74.0.

11.2 Simulating Values of Expected cpue

In step 4a of the general algorithm of Section 6.2, it is indicated that draws from the posterior Dirichlet distribution (14) may be obtained by composition of simulated gamma variates. Most statistical software packages contain functions or routines to simulate values from gamma distributions, but not Dirichlet distributions. This subsection gives details for simulating from a Dirichlet distribution.

Values may be simulated from a Dirichlet distribution with k variables and parameters $\alpha_1, \ldots, \alpha_k$ as follows.

1. Simulate k values from independent gamma distributions with parameters α_j and 1.0, for $j = 1, \ldots, k$. That is, the gamma parameter usually denoted by α has the value α_j for the j^{th} component of the desired Dirichlet distribution and the gamma parameter usually denoted by β has the value 1.0. Let these values be denoted as $x_1, \ldots, x_k$.

2. For $j = 1, \ldots, k$, let

$$\eta_j^* = \frac{x_j}{\sum_{i=1}^{k} x_i}.$$

   Then the value $\boldsymbol{\eta}^* = (\eta_1^*, \ldots, \eta_k^*)^T$ is one value simulated from the desired Dirichlet distribution.

3. Repeating steps 1 and 2 $\mathcal{M}_1$ times results in the values $\boldsymbol{\eta}_m^*; \quad m = 1, \ldots, \mathcal{M}_1$ of step 4a in the algorithm of Section 6.2 in this report.

   Given values $\boldsymbol{\eta}_m^*$, producing values of expected cpue is straightforward as described in step 4b of the algorithm.

12 Tables and Figures

Week	Verified Trips	With Logbook	Total Logbook	No. Dockside
22	96	77	565	21
23	69	54	711	57
24	92	66	697	33
25	50	38	663	28
26	110	76	721	42
27	85	65	689	61
28	28	19	618	47
29	82	63	106	0

Table 1: Number of verified trips, number of those having logbook reports, total number of logbook reports, and number of dockside samples for the weeks of the red snapper season 2011.

Week	$\alpha + x$	$\beta + m - x$	Expected Proportion
22	78	20	0.796
23	55	16	0.775
24	67	27	0.713
25	39	13	0.750
26	77	35	0.688
27	66	21	0.759
28	20	10	0.667
29	64	20	0.762

Table 2: Posterior parameters and expected values for the proportion of trips filing logbook reports during the red snapper season of 2011.

Parameter	Expectation	Interval
α	8.35	(7.54, 9.22)
β	1.58	(1.42, 1.76)

Table 3: Posterior expected values and 95% credible intervals for parameters α and β in the model for effort.

Source	Min	Q1	Q2	Q3	Max
Observed	3.0	16.0	24.0	36.0	192.5
Posterior Predictive	3.7	15.6	25.6	40.5	217.5

Table 4: Five number summaries for observed effort (logbook data) and the posterior predictive distribution of effort in week 23 of 2011.

	Category							
Data Source	$(0, 0.5]$	$(0.5, 1.0]$	$(1.0, 2.0]$	$(2.0, 3.0]$	$(3.0, 4.0]$	$(4.0, 5.0]$	$(5.0, 6.0]$	$(6.0, 7.0]$
Logbook	441	213	42	7	4	2	0	0
Dockside	24	26	7	0	0	0	0	0

Table 5: Observed frequencies of trips falling into cpue categories for harvest of red snapper in week 23 of 2011.

	Logbook Data			Dockside Data		
Week	Mean	Lower	Upper	Mean	Lower	Upper
22	13238	11668	15333	13675	8764	22466
23	14993	13033	17743	15412	11495	21025
24	16120	13985	19055	14878	10080	22010
25	14770	12616	18047	10048	7127	14399
26	16213	14140	18971	14382	10902	19263
27	16971	14815	19912	10881	8275	14666
28	16606	13296	22826	12424	8640	18897
29	19188	14776	25853			

Table 6: Weekly estimated total red snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data sources to provide information on catch.

	Logbook Data			Dockside Data		
Week	Mean	Lower	Upper	Mean	Lower	Upper
22	13242	11655	15302	13688	8743	22336
23	28064	25503	31257	29491	23190	37564
24	44103	40725	48213	44506	36392	55359
25	58814	54706	63736	54454	45626	65855
26	75580	70789	81128	68490	59082	80090
27	92381	87038	98453	78864	68964	91174
28	106772	100832	113541	90067	79401	102629
29	118796	112562	125879			

Table 7: Cumulative estimated total red snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data sources to provide information on catch.

	Logbook Data			Dockside Data		
Week	Mean	Lower	Upper	Mean	Lower	Upper
22	10361	9167	11864	11453	7081	18491
23	11919	10376	13937	12098	9058	16114
24	12962	11256	15102	14051	8813	22351
25	12098	10287	14541	9879	6652	14364
26	12989	11297	15085	14042	10289	19008
27	12155	10629	14058	9992	7531	13297
28	12527	9836	16909	11158	7538	16422
29	9209	7267	11678			

Table 8: Weekly estimated total red snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data after deletion of records with greater than 2 fish harvested per angler.

	Logbook Data			Dockside Data		
Week	Mean	Lower	Upper	Mean	Lower	Upper
22	10294	9151	11797	11485	7014	18804
23	21966	20111	24318	23298	18210	29446
24	34796	32183	37855	36904	29458	46579
25	46547	43375	50187	46530	38490	56686
26	59707	56048	63839	60120	51288	70673
27	71657	67485	76222	69230	60081	80103
28	83143	78524	88253	78479	69376	89745
29	91890	87096	97214			

Table 9: Cumulative estimated total red snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data after deletion of records with greater than 2 fish harvested per angler.

Month	Verified Trips	With Logbook	Without Logbook
1	7	2	5
2	3	1	2
3	64	45	19
4	131	87	44
5	114	85	29
6	326	247	79
7	303	222	81
8	129	86	43
9	24	13	11
10	55	36	19
11	16	11	5
12	2	1	1

Table 10: Number of verified trips having and not having corresponding logbook records, by month (September - December 2010 and January-August 2011)

.

	Logbook		Dockside	
Month	Total	With Harvest	Total	With Harvest
1-2	82	45	8	7
3	572	211	17	8
4	952	382	69	19
5	1183	658	109	72
6	2997	1364	167	89
7	2534	1269	248	129
8	1128	631	108	61
9	436	122	53	19
10	1063	271	106	43
11-12	373	100	33	14

Table 11: Number of logbook reports and dockside samples by month, and the number of each with harvest of vermilion snapper.

	Logbook Data			Dockside Data		
Month	Mean	Lower	Upper	Mean	Lower	Upper
1-2	3839	1464	10138	8719	2092	27759
3	9220	7134	11926	3440	1129	8306
4	21873	18182	26400	9801	5023	17374
5	33899	29428	39285	33553	24408	44969
6	58159	52897	63862	49719	37263	64862
7	59816	53915	66285	39965	30972	50647
8	30765	26333	35990	31370	21790	43778
9	5873	3709	9496	7464	3371	14830
10	10841	8365	14191	10865	6690	16893
11-12	3856	2432	6102	4378	1824	9114

Table 12: Monthly estimated vermilion snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data.

	Logbook Data			Dockside Data		
Month	Mean	Lower	Upper	Mean	Lower	Upper
2	3891	1501	10256	8798	2108	28083
3	11556	9024	14818	10763	4742	21047
4	33132	28477	38458	18417	11240	28628
5	68265	61524	75691	59002	45057	76029
6	127520	118874	136881	108018	88979	129243
7	187284	176589	198692	146755	125610	170630
8	218136	206544	230740	178272	154881	204290
9	224289	212358	236912	183618	159484	211085
10	233935	221562	247140	191900	167935	218118
11	237347	225065	250633			
12	237627	225117	250633	187234	164803	212324

Table 13: Cumulative estimated vermilion snapper harvest and 95% credible interval lower and upper endpoints based on logbook and dockside data.

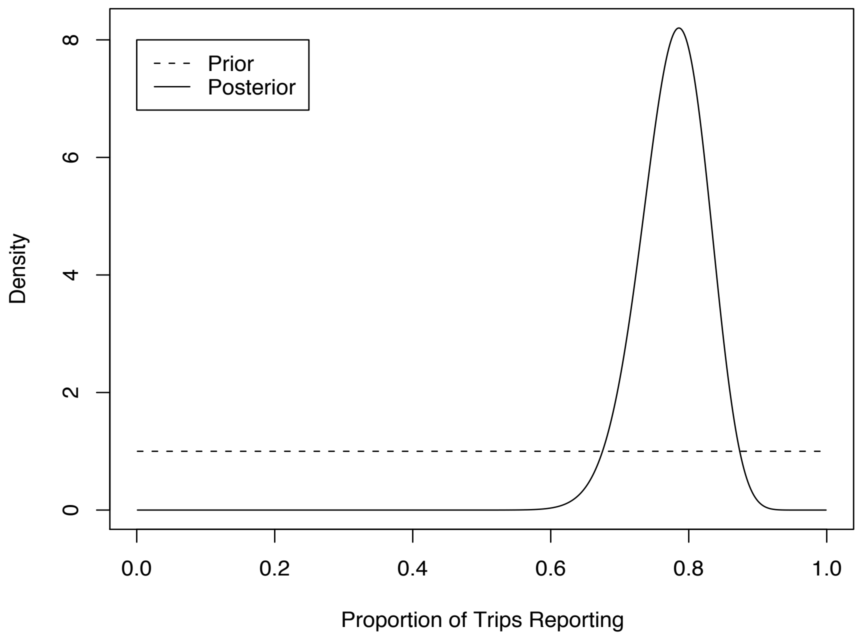


Figure 1: Prior and posterior distributions for the proportion of fishing trips with corresponding logbook reports for week 23 of 2001.

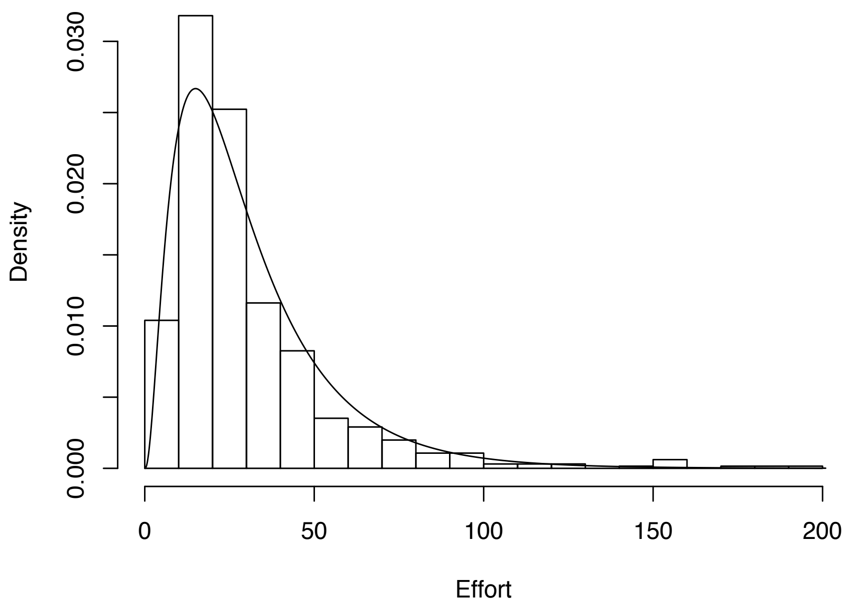


Figure 2: Histogram and fitted generalized gamma density for logbook effort data from week 22 of 2001.

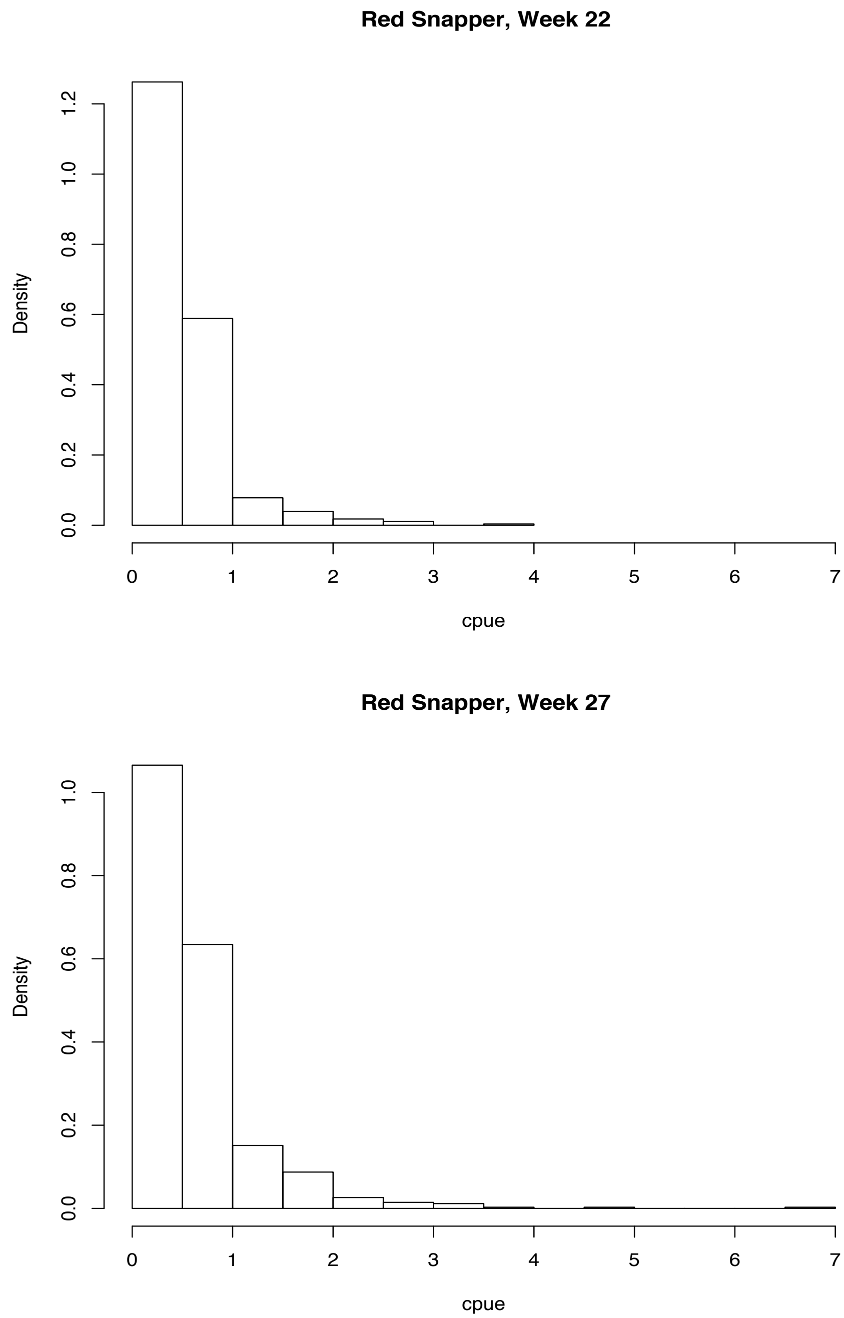
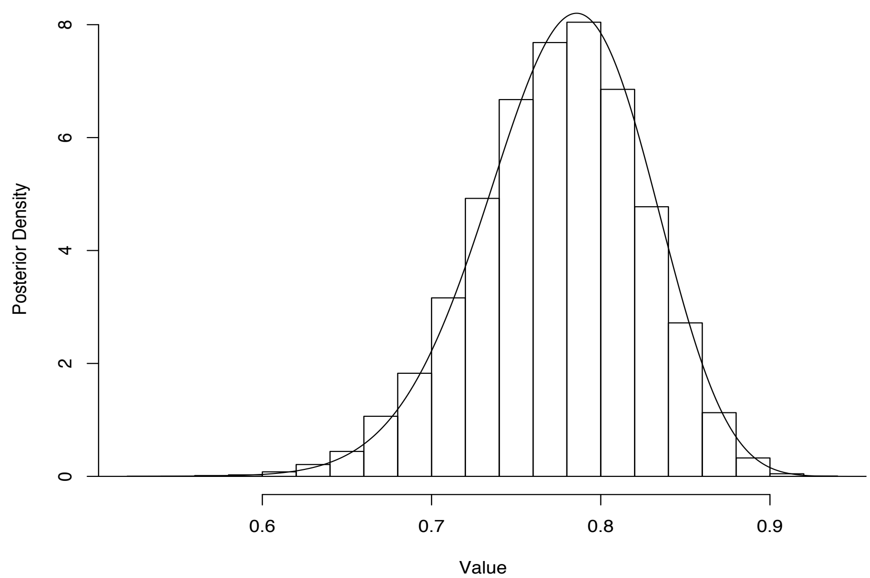


Figure 3: Histogram of red snapper harvest cpue for week 22 and week 23 in 2011 from logbook data.



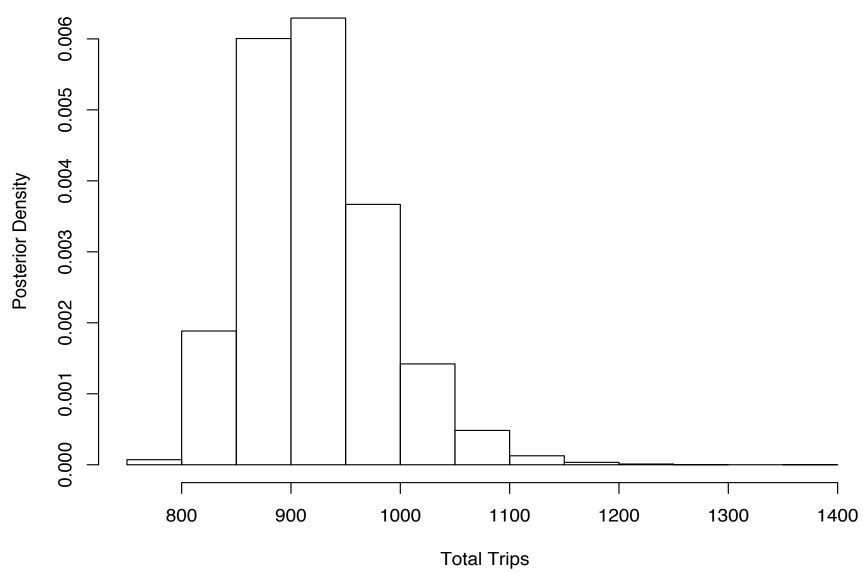


Figure 4: Histograms of simulated values of proportion of trips filing logbook reports and total number of trips from their respective posterior distributions week 23 of 2011.

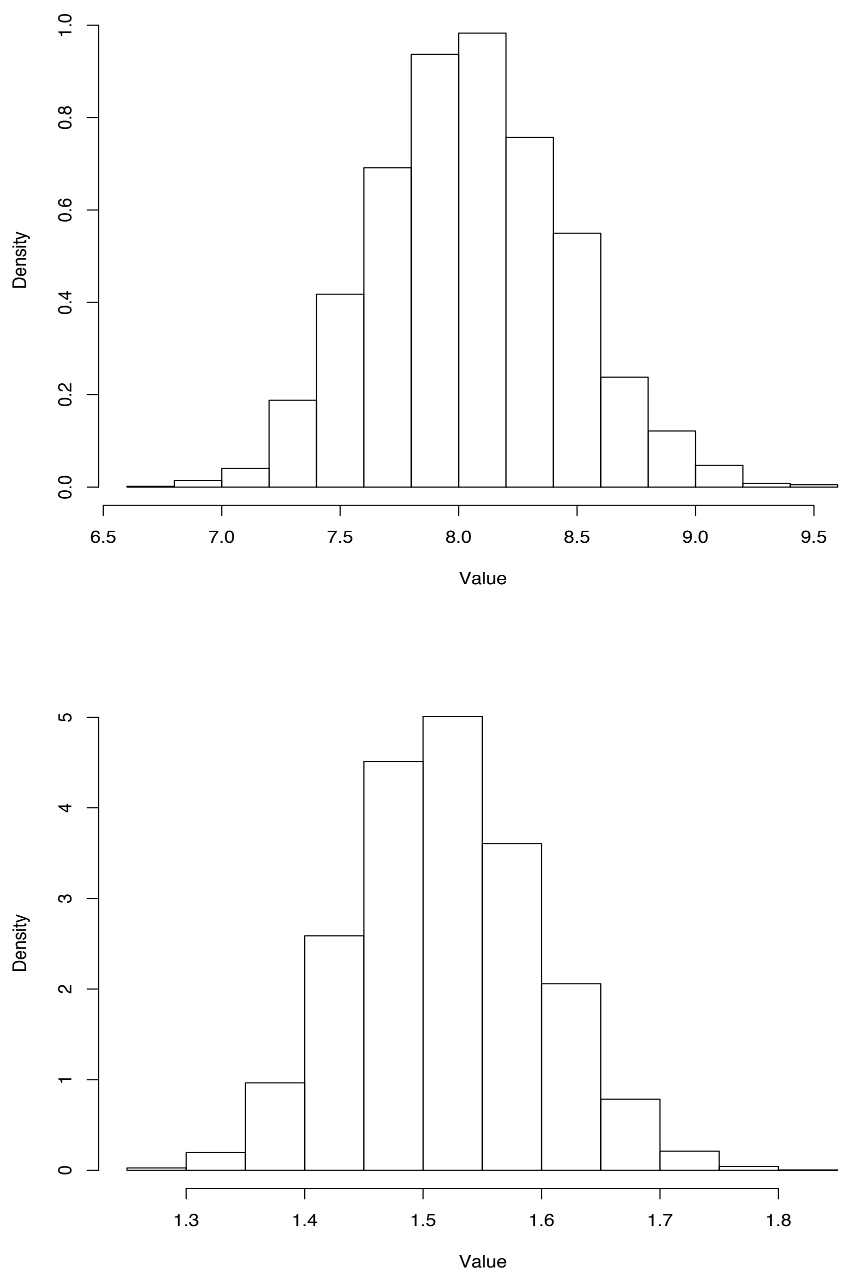


Figure 5: Posterior distributions of effort model parameters α (upper panel) and β (lower panel) for red snapper harvest in week 23 of 2011.

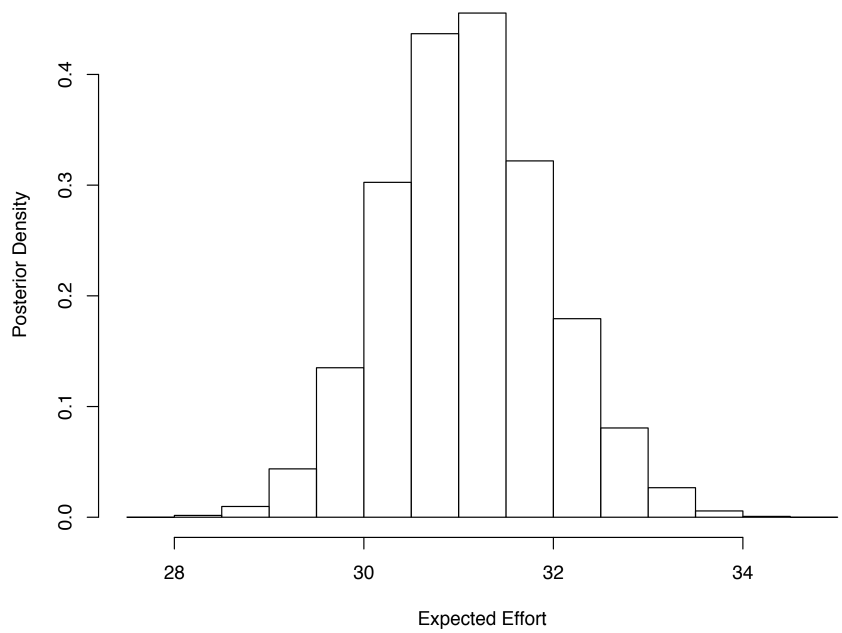


Figure 6: Posterior distribution of expected effort for week 23 of 2011 based on logbook data.

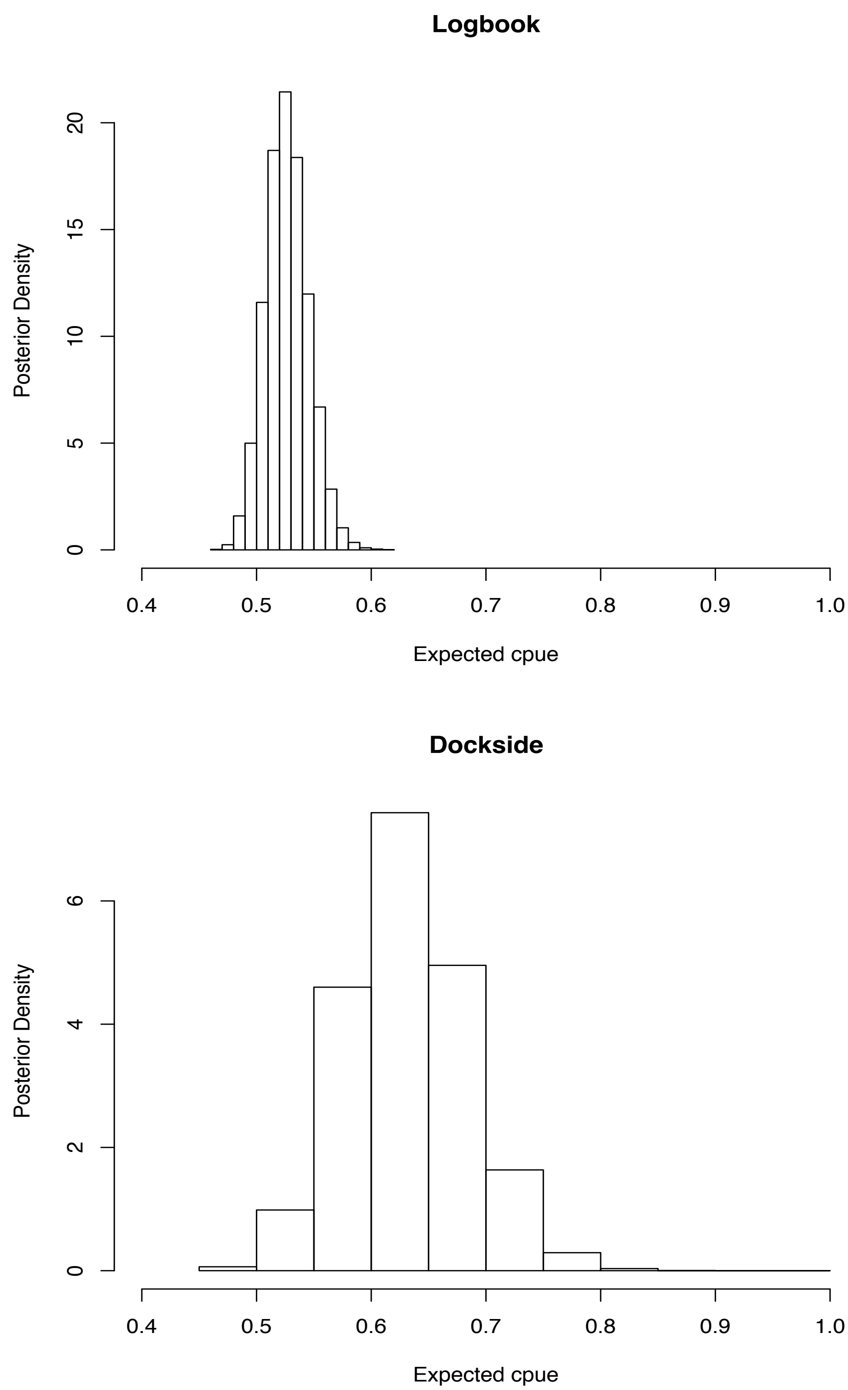


Figure 7: Posterior distributions of expected harvest cpue for red snapper in week 23 of 2011 based on logbook (upper panel) and dockside sampling (lower panel) data.

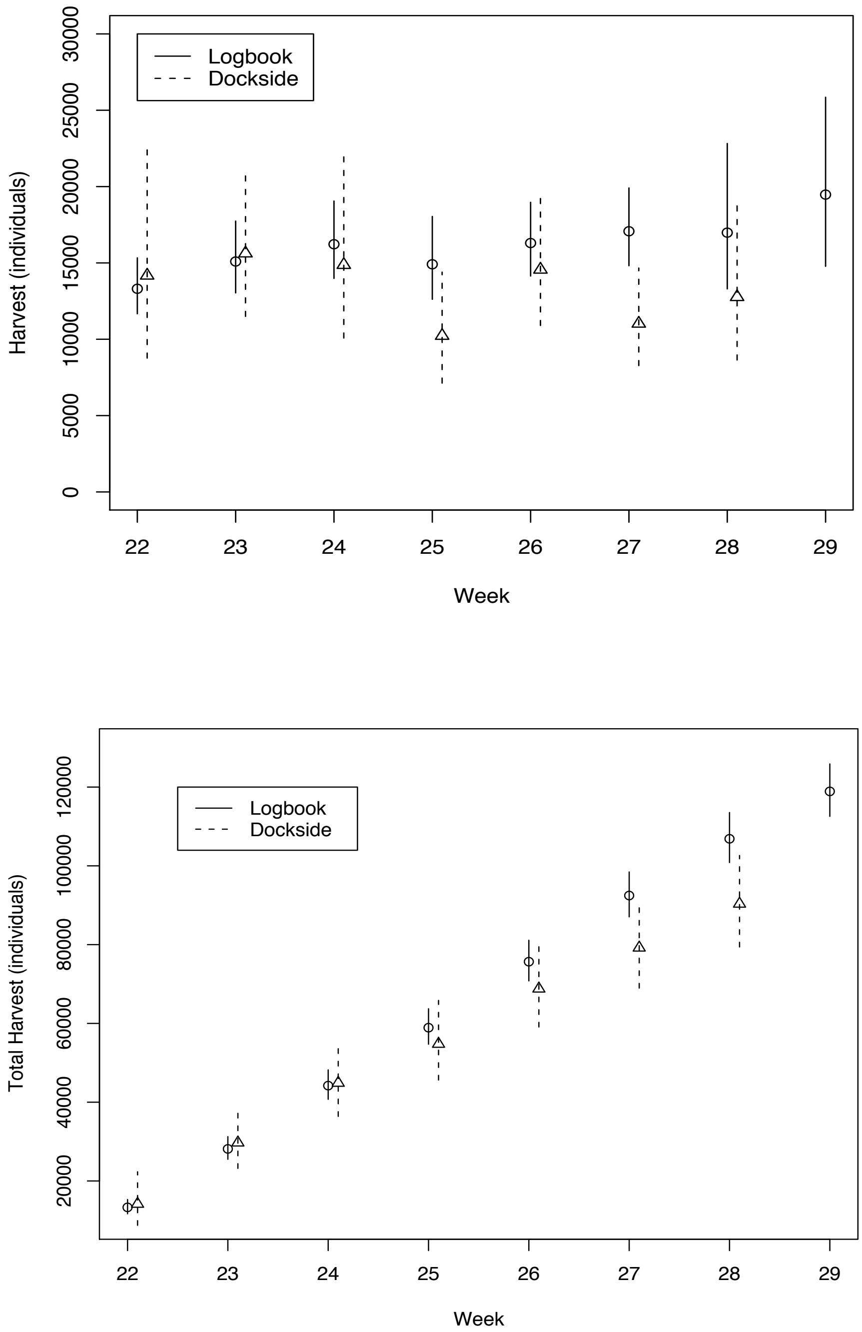


Figure 8: Estimates of total harvest of red snapper by week (upper panel) and cumulative (lower panel) during the season in 2011.

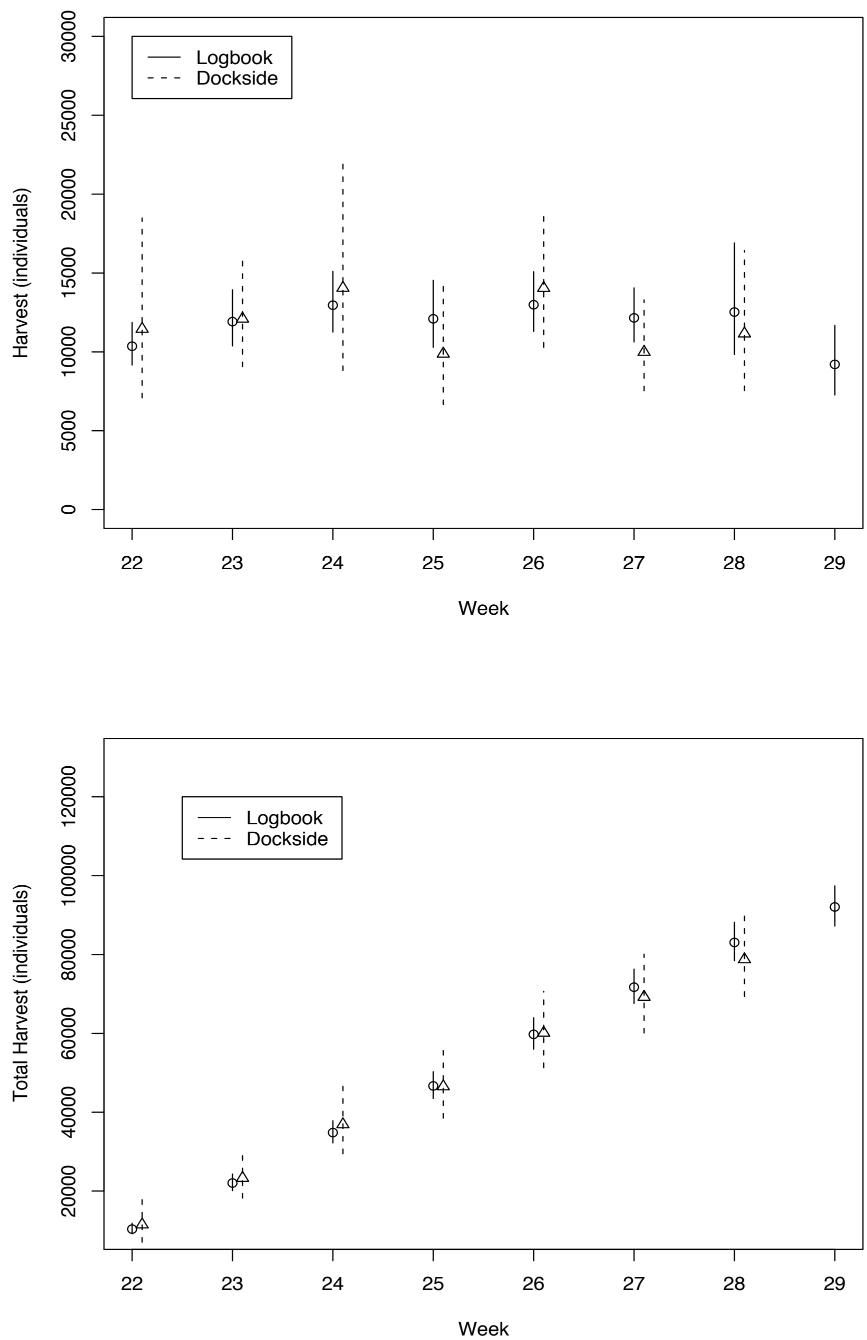


Figure 9: Estimates of total harvest of red snapper by week (upper panel) and cumulative (lower panel) during the season in 2011 after deletion of records with greater than 2 fish harvested per angler.

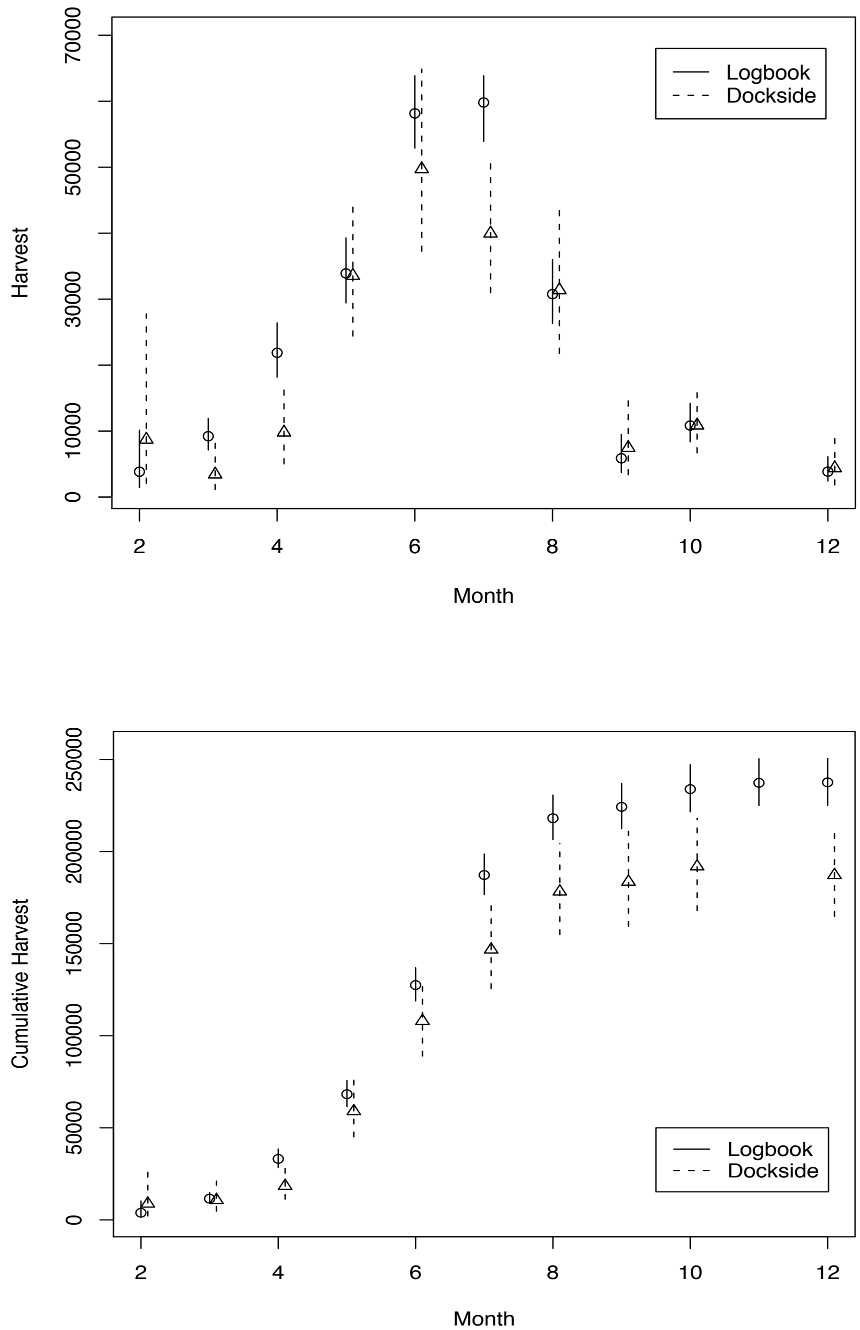


Figure 10: Estimates of harvest of vermilion snapper by month (upper panel) and cumulative harvest (lower panel).

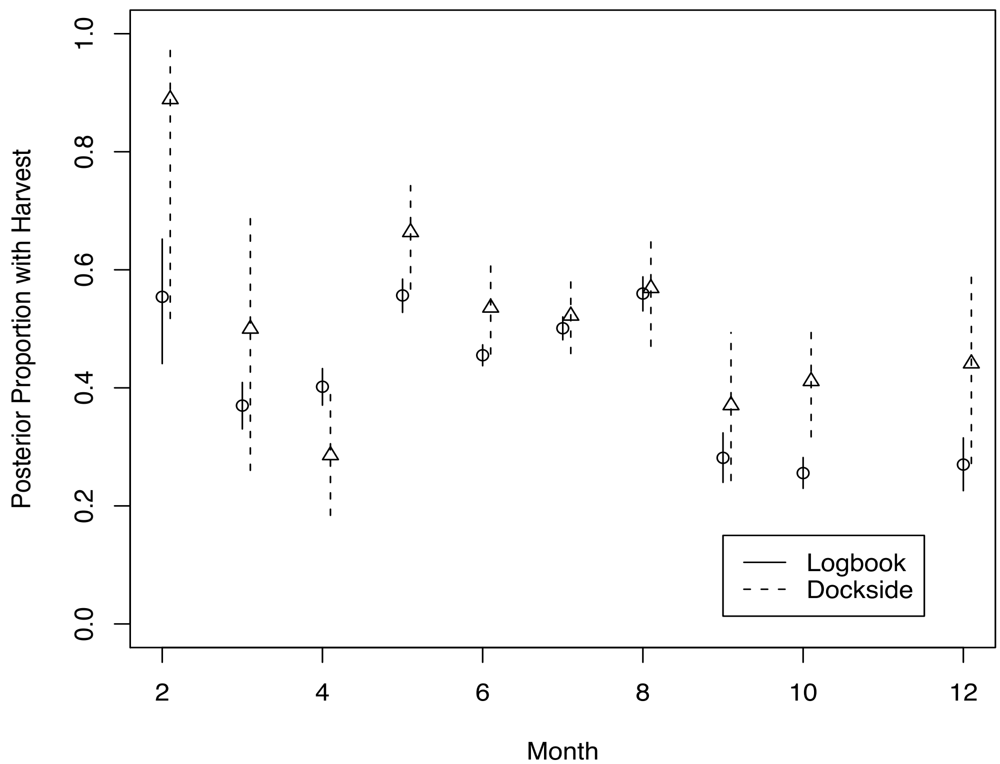


Figure 11: Posterior means and 95% credible intervals for the probability that vermilion snapper was harvested on trips, by month.

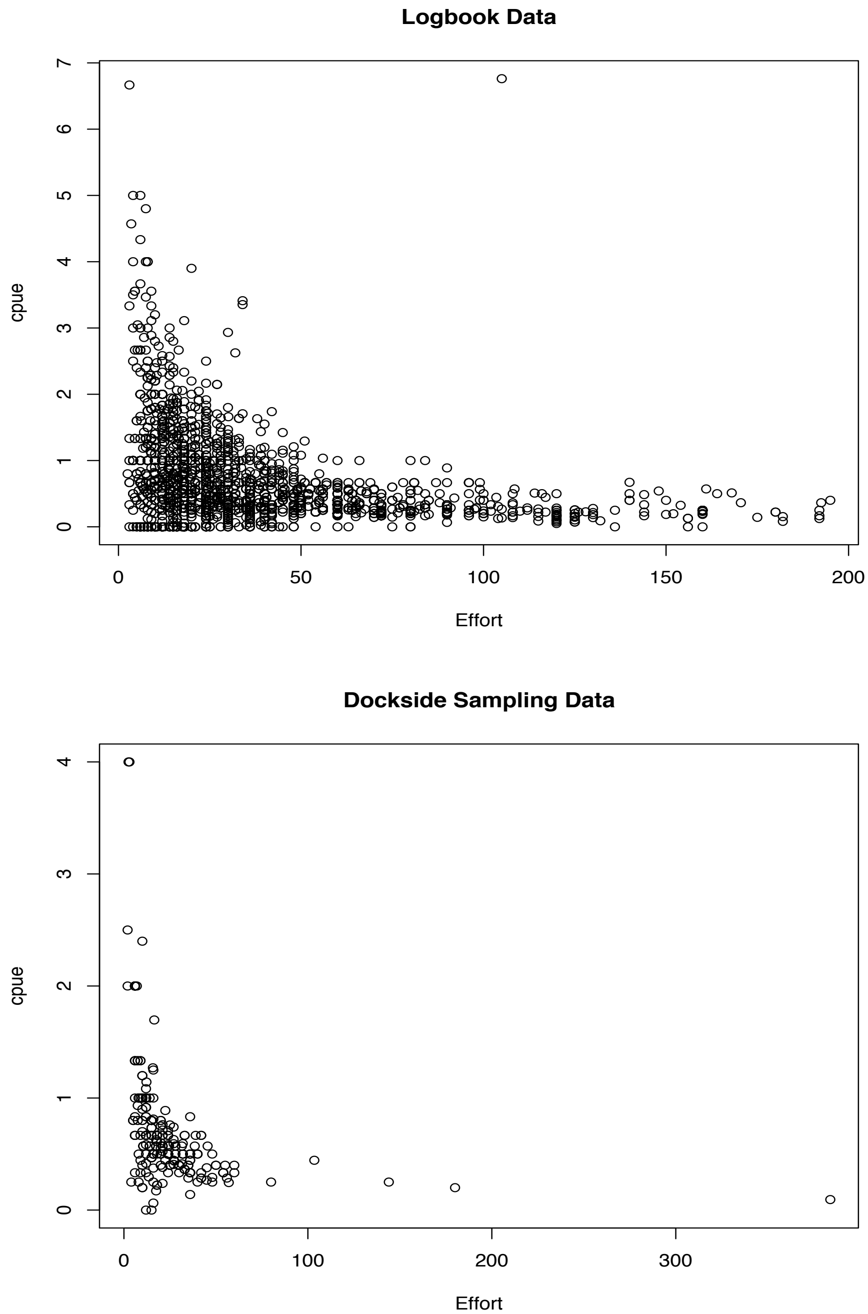


Figure 12: Relations between effort and harvest cpue for red snapper in logbook (upper panel) and dockside sampling (lower panel) data.

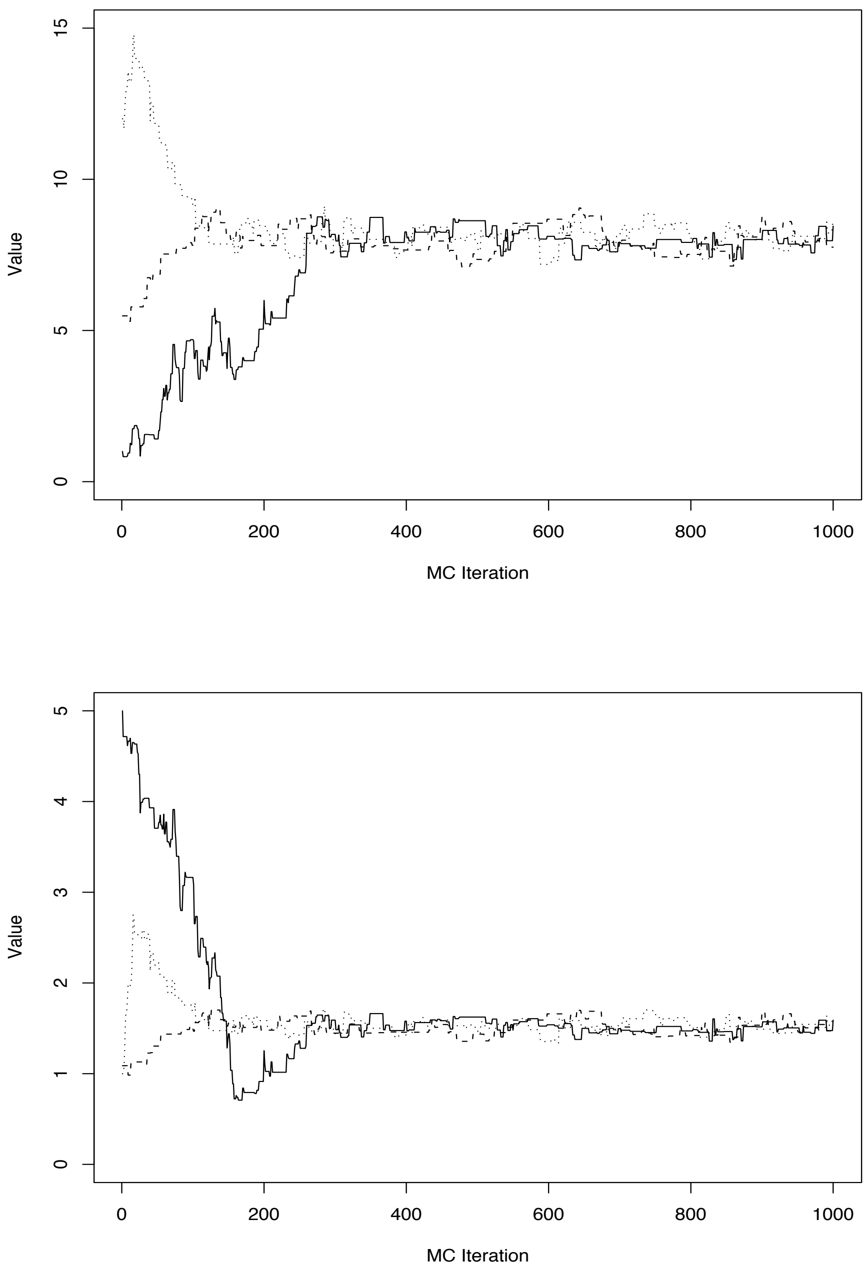


Figure 13: Trace plots of simulated values of effort model parameters α (upper panel) and β (lower panel) for red snapper harvest in week 23 of 2011.

Peer Review Report for

"Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data"

Reviewed by
Dr. Alicia Carriquiry, Iowa State University
Dr. Jean D. Opsomer, Colorado State University
Dr. Stephanie Eckman, Institute for Employment Research, Nuremberg, Germany

Introduction

This document combines the comments provided by three different peer reviewers of the MRIP Project Report entitled "Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data." The document provides verbatim reviewer comments without identifying the source of each comment.

Reviewer 1

In this report, an approach for estimating charter-boat catch and related characteristics based on logbook data is proposed. The approach uses Bayesian modeling, making it convenient to specify separate models for number of trips, average effort per trip and average catch per unit effort, and combining their estimates and associated measures of precision into a single estimation procedure. The approach is demonstrated on two species, red snapper and vermillion snapper. Finally, the report gives a set of recommendations and conclusions. The overall estimation approach is carefully described and statistically valid, and there is a large literature on implementing Bayesian models such as those proposed here.

While the approach is statistically valid and appropriate, I have three somewhat interrelated concerns regarding its implementation in the context of NOAA's overall goal of producing official estimates of catch and related characteristics from these data. I will describe those first, followed by further discussion, and I will number my comments for ease of reference.

1. The first and most important concern has to do with the non-random selection of the logbook data. This was recognized early on, when it was clear that the logbook records could not be treated as a census of trips. However, they are treated here as a random sample of trips, which is better than a census but which still might not be appropriate, because of non-random selection of reported trips. Looking specifically at the three identified components of catch, I see the estimation of the first one (number of trips) as being fine, but I am not sure about the remaining two, because they assume that the logbook trips are a representative sample of all trips in the estimation procedure. I realize the estimates are being compared with those obtained from dockside sampling in this report. But the proposed method does not actually provide any way to incorporate the dockside sampling estimates, and they are only used as an external check to detect potential large discrepancies. I will return to this issue in comments (4) and (5) below.

2. The second concern has to do with the fact that, being model-based, appropriate models need to be picked for each of the components of the estimation procedure. This needs to be potentially repeated for every variable (catch vs. released), species, each time period, etc. This was illustrated in this report with the change in the models between red and vermillion snapper, for instance. Doing this properly for the range of estimates that are to be produced is a very significant amount of work, which will need to be done and documented to justify the estimates. In general, such modeling effort is usually only undertaken if design- based methods cannot be used, for instance because the sample sizes are too small or because issues such as informativeness, nonresponse or measurement errors are present to such a degree that they need to be adjusted explicitly via modeling. It is not clear to me that the logbook data fall in either of those categories. Let me reiterate that I do not think the proposed estimation method is not statistically valid or otherwise incorrect. The issue is that implementing this for the official for-hire catch estimation will entail a very substantial amount of work.

3. The third concern relates to the measures of uncertainty that are produced as part of the Bayesian estimation. As explained on p.30 of the report, the Bayesian estimation approach provides these measures of uncertainty as part of the estimation procedure, which is a indeed major strength of the approach. However, these measures of uncertainty are valid under the critical assumptions that (1) the parametric models for the various components of the final estimator are correct, and that (2) the observed and unobserved parts of the target population follow the same distribution (i.e. that the nonrandom selection of the observed units did in fact produce a representative sample). Whether these assumptions are met or not has a major effect on the precision and reliability of the estimates, so the concern is therefore that the precision of the estimates as produced by the posterior distributions could be overstated, potentially severely so.

4. Returning to the first concern, I want to discuss the use of the dockside sampling data in validation and possibly estimation of the for-hire catch. As done in this report, these data can be used to compare with the estimates coming from the logbook data to assess informativeness of the latter, as was done in this report. However, what I would have preferred is not to compare the model fits using both data, but to compare the model fits using the logbook data to \un-modeled" (design-based) estimates from the dockside data. While having both model fits agree is certainly a good sign, there is the potential for both being wrong in the same direction if the assumed model is not appropriate. Note that I think this is unlikely for these species, which were carefully modeled, but as a recommendation for future comparisons, I think that comparing a model prediction with a model-free estimate is a better approach.

5. But in addition to using the dockside data for comparison, I am wondering whether a procedure that incorporates them in estimation might not be preferable. Ideally, I would like to be able use the logbook data to estimate the average effort per trip and average catch per unit effort for the logbook trips, and the dockside data to estimate the same quantities for the

non-logbook trips. This would correspond to splitting the population in two components, with a census of one (logbook trips) and a random sample of the latter (non-logbook trips). Another possibility is to treat this as a dual-frame problem, with the dockside sampling reaching both the logbook and the non-logbook trips. This would completely remove the selection bias concerns, but I realize both of these approaches might be very difficult to do in practice, because they require a determination of the \logbook status" of individual dockside observations. If that is not feasible, I still think it would be worthwhile to investigate ways to incorporate the dockside data directly into the overall estimation procedure, possibly by calibrating relevant quantities in the logbook estimation procedure to the dockside results. The goal here is to remove potential sources of informativeness to the extent possible. As a simple example of this, one could compute the estimated average number of anglers and/or the average trip length for the dockside data, and use those to calibrate the corresponding quantities for the logbook data. I am not sure whether this can be incorporated into a Bayesian estimation framework, but conceptually, this would remove potential bias in the logbook data that would be caused by the logbook trips being either of different size or of different length from the non-logbook data. Even if the dockside data are not used in estimation, a set of relevant measures of comparison between both the dockside and the logbook data should be developed and tracked over time, to ensure (and confirm to the data users community) that the logbook data are representative of the overall for-hire catch.

6. An important issue identified in several places in the report is that of data review and editing. The presence of outliers, such as those discussed in the report, clearly has a large effect on the resulting estimates of catch. This will need to be addressed regardless of the estimation method eventually adopted.

7. Finally, I found a number of minor typos that should be fixed: “allude" instead of “elude" (p.19), “available" not “avaliable" (p.21), “amount" not “amont" (p.28).

Reviewer 2

The author has written a report proposing a new technique to estimate catch, release and mortality of fish species in the Gulf coast of Florida. The report is very well and clearly written. Bayesian estimation techniques are not easy to describe in non-technical language, and the author has done a very good job with the task. I am also convinced, based on the material available in this report, that the estimation technique is a sound one, though the report would be more persuasive if it also included a comparison to more standard estimation methods. I have a few comments on how the text could be further improved.

General Comments:

I strongly recommend that the report include a paragraph or two in the introduction on the use and importance of estimates on catches and releases from for-hire boats. While it is obvious to the author why this data is needed, it is not necessarily obvious to readers and the report will be stronger if it (briefly) makes a case that this data is crucial to government and industry.

The report repeatedly gives 95% credible intervals but does not explain how such intervals are to be interpreted. One or two examples in the report of how to interpret such intervals should be added.

The report should include an outline to let the reader know how it is structured.

Please use commas in numbers with more than 3 digits before the decimal point – this will make the text and tables easier to read.

I was expecting to read in the conclusion some recommendation about whether collecting dockside data was worth the cost. Does the author feel that the proposed technique using only the logbook data is sufficiently robust that the dockside data is not needed? Would the logbook data be as good as it is if the dockside data were not collected alongside? (That is, might boat operators give lower quality data if they know that no dockside validation will be done?)

Specific Comments:

Page 3 – I would like to see a bit more discussion of the results of the MRAG Jan 2006 report that found the estimation approach in this report performs better than a survey based estimator.

Page 11 – spell out in words what the values in the w and eta vectors are. I believe w_j is the number of trips where the cpue fell in the the jth bin and eta_j is the proportion of trips falling into this bin.

Page 13 – similarly, more words are needed here about what the values of y_i and w_j are and what n and k are.

Page 14 – please don’t use m here to index the number of simulations, as m is used on page 12 and elsewhere as the number of verified trips.

Page 15 – here, or perhaps above, a discussion is needed about why we want M_1 different values of the total catch. The reader may be worried that the algorithm produces not a single number but a large vector of numbers. Explain that one can use this vector of values to get a sense of the distribution and make statements about the estimate’s precision.

Page 16 – first full paragraph, last sentence. The use of “information on catch” and “catch information” confused me. Please use the same phrase in both parts of the sentence to emphasize that the important difference is between logbook and dockside sources. I also suggest italicizing logbook and dockside to draw the reader’s attention to the fact that this is the important distinction you’re making.

Page 17 – point b. Above you recommended using a diffuse distribution here as prior. You might make that point again here.

Page 18 – point 2. When you give the value of 50,000 please indicate that this is M_1 in the algorithm.

Page 19 – point a. M_1 is also 50,000 here, but it needn’t be the same as above, right?

point b. again, include references to the notation used in other parts of the text. “For each pair of values (\alpha^*_m, \beta^*_m) simulated from the posterior, M_2 = 10,000 y^*_{m,q} values were then simulated from the data model (7) using those values as parameters.”

As above, though, avoid the use of m as an index.

point c. here you do a nice job referring back to the earlier notation.

Page 20 – please add a point 5 which states that you multiplied the three estimates together to get tau.

Page 25 – first full paragraph. This paragraph needs a topic sentence to make it clear to the reader what the point is. The last sentence starts with “the point is” but by then the reader is far too lost. The first sentence of the paragraph is about red snapper, but the rest of the paragraph is not. Please rewrite this paragraph.

Also, the text talks about % of trips with any vermillion snapper caught, but the table presents only numbers. Percents seem much more relevant to the discussion, so redo the table in terms of percents.

Conclude this paragraph with a statement that the next section will address how the high proportion of trips without vermillion snapper catches necessitates a change in the estimator.

Page 26 – first line. Refer the reader back not only to section 3 but also to a specific step in the algorithm.

Page 27 – why do you give such rough numbers here? You have better numbers in the table and numbers in the text that match the tables would help the reader see where you're pulling the numbers from. Also, 250,000/12,000 is more than twice as many; it's 20 times as many. Or does the "about twice" in the text refer to a different ratio?

Page 28 – missing ")" near the end of the page. Amount misspelled as amont. "there appear to be groups" is very vague. Trends would be a better word here.

Page 30 – first paragraph. This paragraph is quite defensive. The author seems to be arguing against critics that are not known to the reader. Either cite some articles or correspondence from these critics or tone down the defensiveness in this paragraph.

point 1 – this point also seems to be arguing against a suggestion that the reader has no knowledge of. What kind of size eligibility criteria are currently used for the logbook study? What does the author propose for those criteria? In contrast to the rest of the report, this section is not clear.

Page 31 – point 5. The reference here should be to Figure 12, not Figure 9.

Page 33 – "can do" – something is wrong in this sentence.

Page 34 – top of page. Again the author is responding to criticisms that the reader is not aware of. First tell us what these criticisms are, with references or quotes, and then rebut them.

Towards the end of this page there are a few typos.

Tables and Figures

Table 1: the text mentions that the first three columns are m, X and M, but these labels should be included in the table.

Table 6 (and all others): include commas in numbers.

Table 11: make "With Harvest" columns into %s, not numbers

Figure 2 (and others): include kappa in the horizontal axis label to make it clear to your reader what is represented.

Figure 4: again, label top horizontal axis as theta hat posterior and bottom horizontal axis as Nhat

Figure 6. The horizontal axis here is kappa*

Reviewer 3

I enjoyed reading Mark Kaiser's report on estimation using logbook data. Kaiser does a very good job of motivating the methods and explaining assumptions and limitations of the estimators he proposes for the various quantities of interest.

At the outset, I will say that I fully agree with a conclusion in this report. Using logbooks, at least according to the information obtained from the pilot study, appears to be a reasonable sampling strategy, which, in conjunction with the appropriate statistical methodology will almost certainly result in more accurate estimates of catch, release and other totals than a poorly conducted census.

There is little discussion in the report about whether everyone is expected to fill the logbook in every trip or whether at some point a more formal sampling approach might be conducted to select a subset of the trips for logbook completion. This point, which can be discussed at a later time, in no way would change the methods or the conclusions that are drawn in this report.

If there is a limitation to this methodology, it is that it requires some expertise by the person in charge of producing the estimates. This type of analysis is difficult (but not impossible) to carry out using "canned software", so at least initially, MRAG might need to depend on a consultant to carry out analyses. One other limitation that received little attention in Kaiser's report is that for some of the parameters (e.g., cpue), it might be necessary to develop a different model for each different species. As a consequence, it is probably not possible to write a "black box" type of program that MRAG can use on all types of fish. I revisit this issue briefly in my specific comments.

Overall, I find that the methods proposed in the Kaiser report are correct, appropriate and well justified and that Kaiser's approach offers an excellent alternative to the methods that appear to be in use today. My specific comments are mostly picky and offered as suggestions for the future version of the methodology

Specific comments

- On page 5, something might be said about other choices for the prior parameters α,β. How sensitive are the results (given the data) to different choices of these two parameters? Probably not terribly sensitive because of the large sample sizes, but even a small change in the posterior mean of θ could have a noticeable effect on τ because of the multiplicative structure of (1).

- Page 7, equation (7): why not estimate δ together with the other parameters in the model? Perhaps I am missing something here, but it does not seem like including δ in the MCMC process would be too difficult. If nothing else, at least it would be good to understand why the choice of $\delta = 0.5$ and also get a sense of the sensitivity of the estimates of this (different) set of parameters denoted α,β to different values of δ.

- Page 10, toward the top: I very much appreciate the difficulties with modeling a variable such as cpue. One alternative that is not mentioned is to use a semi-parametric mixture approach. The binning approach proposed by Kaiser is fine, but the bins are arbitrary and would need to be revisited every time a new dataset is analyzed. Mixtures, in contrast, can be formulated so that the number of components in the mixture is estimated from the data. Without going to such a "high tech" extreme, a simpler model in which the number of components is limited to, e.g., three, ought to be flexible enough to account for almost any shape in the distribution of cpue. One advantage of using a mixture model is that developing "black box" software is easier.

- I missed mention of δ in the applications to red and vermilion snapper data. Was it fixed at 0.5 always?

- Page 30, point 1: I do not know what proportion of fishing is done in small recreational boats versus large vessels. Kaiser is correct in that mixing apples and oranges is typically a bad statistical idea. However, rather than eliminating the large fishing vessels from the dataset, I would probably carry out the analysis for the two types of boats separately and then combine the results using some kind of weighting that depends on the relative contribution to catch by the different types of boats.

- Page 31, point 9: The scatter plots clearly show a non-linear association between effort and cpue. What if cpue were transformed (using, e.g., a square root transformation). Would the association become more linear?

- Page 32, first full paragraph: I agree with the points raised by Kaiser, but a question is why not simulate the two variables using a bivariate proposal instead of assuming independence. If a transformation could be found to linearize the association between the two variables, carrying out a bivariate simulation would not be difficult.

Response to Review Comments Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data

31 July 2014

Mark S. Kaiser

Dept. of Statistics

Iowa State Universtiy

Prepared for

Marine Resources Assessment Group Americas

1 Introduction

This document contains responses to review comments provided on the MRIP Project Report titled "Charter-Boat Logbook Reporting Pilot Study Estimators for Use with Logbook Data." In that report, an estimation strategy was developed for use with logbook data from the for-hire fishery of the Gulf of Mexico. The proposed estimator was applied using logbook data to provide information on catch and effort, and the results compared to results from using dockside (intercept) sampling to provide information on effort and catch. The document provided to the author by MRIP was organized in sections titled Reviewer 1, Reviewer 2 and Reviewer 3, and that organization is followed in these responses.

These responses deal only with comments of substantive nature from the reviews. Comments regarding details of writing and suggestions to improve communication are appreciated, but are not responded to in this document. The report that was reviewed was a follow-up to issues that arose in analysis in the MRIP Charter-Boat Logbook Reporting Pilot Study. It was intended as an initial "proof-of-concept" regarding the use of logbook data in estimation of catch. It was not intended to be a completely polished and self-standing development of estimation strategies for quantities desired in the recreational fisheries in the Gulf of Mexico. Nor was it intended to address all of the possible issues that exist in these estimation problems. It was meant to illustrate one potential estimator that could profitably be applied to estimation in the Gulf using logbook data, and to suggest that this estimation strategy be pursued by MRIP in the future.

While the report under consideration was not meant to address the overall estimation problem of catch (and/or discard) it does suggest one potential estimation approach that lends itself to combining data from different data collection programs and from various sectors of the fishery, including both commercial and recreational. It does appear that the reviewers recognized this potential, and that may be one of the more important outcomes of this work. The reviewers do not appear to have been given any detailed context within which the reviewed report was to be conducted, and that was perhaps unfortunate. I assume that reviewers were not provided extensive context in a effort to maintain "objectivity". This may be a reasonable objective, but it can also lead to reviewers expending a fair amount of energy commenting on issues a report it was never intended to address.

2 Reviewer 1

Numbered responses in what follows refer to the same numbered comments offered by this reviewer.

1. This comment makes an important point that could be applied to every data collection effort connected with monitoring of the recreational fishery in the Gulf of Mexico. Any source of data that is accepted for use in an estimation strategy is assumed to be representative of a larger population, whether that be trips, time intervals, vessels, fishers, or some combination of these. MRIP has employed some data collection and estimation strategies based on

classical principles and procedures of survey sampling, which are designed to ensure samples are representative. However, these strategies assume that the procedures can be implemented as designed, without excessive nonresponse, unavailability, inaccurate recall, lack of resources to implement a plan at a given time in the manner it was designed, and so forth. The point is that any statistical estimation procedure assumes the data it relies on are representative of the situations about which inferences are to be made. This is important, and is true for the proposed estimator just as it is for any existing or other potential procedures.

2. It is true that in a model-based approach to statistical analysis one must specify a probability model, including distributional forms. It is also true that these aspects of model formulation are important and can influence the results of an analysis. It is not true, however, that "such modeling effort is usually only undertaken if design-based methods cannot be used". There are any number of situations in which a design-based approach could be used but would prove inferior to a (properly specified) model-based approach. This is one of the reasons the use of what are called "model-assisted survey estimators" have become popular in the area of survey sampling. The comment that implementation of the proposed estimation strategy will require work is most certainly true, and there are any number of issues that would need to be addressed. This is, however, not unique to the proposed strategy.

3. Statistical analysis based on a probabilistic model proceeds under the assumption that the model structure is correctly specified. Statistical analysis based on a design-based approach proceeds under the assumption that the population has been correctly identified, sampling units appropriately specified and identified, a sampling frame constructed correctly, any strata or sub-population groupings meaningfully defined and identified, and implementation issues are

not present. Estimates of uncertainty under any approach can be adversely affected by severe violation of the assumptions made. This comment, as the previous commments, is valid and a proper concern for any analysis using any statistical approach, but is no more severe for a model-based approach than for a design-based approach.

4. The comparison described in this comment would be a valuable undertaking, but was not the purpose of the project under review. The purpose of the project was to develop an estimation strategy that could be used with logbook data (in conjunction with sampling for activity or number of trips) and to compare the results with results from using dockside sampling data. Overall estimation strategies for various sectors of the fishery and, especially, how to combine those strategies into overall estimators of catch and discard should constitute an ongoing effort. It might be worthy of mention that in a different project concerned with the estimation of discard in the commercial groundfish fishery in the Northeastern United States, an estimator similar to that proposed here was shown to outperform the standard design-based estimator for each of seven species used in an evaluation. That does not, of course, indicate that the model-based approach proposed with outperform a design-based one in this situation, but it does indicate that the opposite is not a valid foregone presumption.

5. Developing one or more strategies for combining both logbook and dockside sampling data into one estimation procedure is an appealing idea that should be pursued. I would not approach this from within the constraints of a design-based framework, but the objective is desirable. If it would be possible, as hinted at in the comment, to make use of dockside sampling as both a means of verifying logbook data and in constructing a combined estimator from these different data sources, that would be ideal.

6. I agree completely with this comment.

3 Reviewer 2

This reveiwer offered unnumbered General Comments and page-referenced Specific Comments. Although the General Comments were not numbered, they were separated with white space into 5 discrete thoughts. My numbered responses in what follows correspond to these paragraphs. Most of the specific comments concern suggestions for improving clarity in presentation. While valuable, I will not respond to these in the manner that one might in a manuscript revision. The purpose and scope of this report was discussed in the Introduction.

1. This report was written as an addition to a project conducted on the evaluation of a logbook reporting system in the Gulf of Mexico. I refer to the overall report for that project concerning the value of estimating catch in various sectors of the fishery.

2. The report assumes a reasonable familiarity with the fundamentals of Bayesian analysis. Credible intervals are to be interpreted as intervals that contain the specified proportion of our belief space about the value of the quantity under consideration.

3. Editorial

4. Editorial

5. In the primary report on the pilot project "For-Hire Electronic Logbook Pilot Study in the Gulf of Mexico" it was indicated that logbook data should not be assumed to provide similar indications about catch and releases as dockside sampling data in perpetuity, and that a dockside verification program should be implemented. A separate MRAG report dated January 2012 and

titled “Charter-Boat Logbook Reporting Pilot Study Verification Sampling” addresses this in some detail.

4 Reviewer 3

This reviewer offered General Comments in narrative form separated into paragraphs, and Specific Comments some of which are more than editorial in nature. I will try to identify which comment is being responded to in each of the following ennumerated responses.

1. Fourth Paragraph of General Comments.

   It is true that some understanding of the statistical procedures developed will be needed to implement the estimation strategy proposed, and it is true that the strategy proposed cannot be conducted using spreadsheet software. The problem being considered is difficult and complex. Solving difficult and complex problems often requires more than a trivial approach.

   As also mentioned by Reviewer 1, there well may be a need for species-specific modifications to various components of the overall model (such as distributions of catch). This is one aspect of the situation that causes the overall problem of estimating catch to be difficult and complex.

2. First and Second Specific Comments.

   All aspects of the model should continue to be developed and evaluated. The point that the multiplicative nature of the model can magnify errors is well made.

3. Third Specific Comment.

   The use of a finite mixture with a fixed number of components was attempted, but problems were encountered with extreme values for cpue and catch. Several long-tailed distributions were also considered, but with similar problems. This motivated the use of the discrete binning used. Distributional forms are certainly prime candidates for additional improvement. A mixture with an unknown number of components was not attempted, as estimating the number of components remains an involved problem (and certainly not amenable to a "black box" solution).

4. Fifth Specific Comment (reference to page 30).

   The suggestion in the report was not to ignore larger vessels, but to consider them to be contained in a different sector of the fishery, such as headboats.

5. Seventh Specific Comment (reference to page 32).

   Agreed. The only catch is "If a transformation could be found . . .". If one could, then this comment could be enacted and that would be beneficial.