Marine Recreational Information Program
FY-2014

Implementation of the iSnapper smartphone application to collect data across all recreational sectors in the Gulf of Mexico

Project: Implementation of the iSnapper smartphone application to collect data across all recreational sectors in the Gulf of Mexico

Gregory Stunz - Author, *Texas A&M University - Corpus Christi*
David Yoskowitz - Author, *Texas A&M University - Corpus Christi*
Mark Fisher - Author, *Texas Parks and Wildlife Department*
Megan Robillard - Author, *Texas A&M University - Corpus Christi*
Tara Topping - Author, *Texas A&M University - Corpus Christi*

Table of Contents

1. Acknowledgement

Acknowledgments
We thank the National Marine Fisheries Service Marine Recreational Information Program for funding this study. This study would not have been possible without the support of all the recreational anglers that voluntarily submitted their catch data during and after the federal season. These individuals were willing to take the time to submit trip information to provide a more robust data series to fisheries managers. We also thank Rob Andrews, Ken Brennan, Dave Donaldson, and Andy Strelcheck for their advice about data collection and their support of this project, as well as Gregg Bray and the Gulf States Marine Fisheries Commission. We thank Texas Parks and Wildlife creel agents for helping inform the public about *iSnapper* and a special thanks to Jeremey Leitz who designed a variety of print media which TPWD made available at boat ramps and at all Coastal Fisheries offices. We also thank Lynne Stokes and Bingchen Liu for providing their statistical expertise in both development and in calculating the total harvest of Red Snapper based on angler reporting. Their advice regarding the best methods for trip validation proved to be a critical feature to determine the accuracy of the self-reported data. Finally, we would like to thank Michael Christopher of Elemental Methods, LLC for redesigning and improving a new version of *iSnapper* along with working in several special features and designs creating a more appealing application for the private anglers to use.

Additional funding to run *iSnapper* was provided through Texas Parks and Wildlife for the 2016 season and the National Fish and Wildlife Foundation for 2017 and 2018. Thanks to the continued funding, we are updating the final report to include data and findings from 2016 and 2017 to provide a more comprehensive view of the utility of the application.

2. Executive Summary

Red Snapper is one of the most highly targeted species in the northern Gulf of Mexico. This fishery has been classified as overfished since the late 1980’s, and this designation has led to drastic reductions in both season and bag limits. Uncertain and unpredictable recreational catch rates along with a federal court ruling required managers to build in large harvest buffers to prevent overfishing, and these measures have created even shorter seasons with only 10 days in 2015. Ironically, this situation has also created a much narrower window for fisheries managers to collect catch data, often shorter than historical survey methods were designed to accommodate, further compounding the harvest estimate issues.

A major challenge to improving recreational season length is the lack of timely catch data. Supplementing the Marine Recreational Information Program (MRIP) with an electronic data collection system would provide more timely and robust information, thereby allowing managers to refine catch estimates and reduce buffers that could lead to optimizing the harvest and allow for longer recreational seasons. Thus, we created *iSnapper*, a smart device application designed for private anglers to log their catch and effort data. Mobile applications provide a unique opportunity to provide better data that will work in combination with current MRIP survey protocols to bolster rapid in-season (and out of season) catch information that would otherwise be unavailable. In addition to catch statistics, this type of reporting mechanism generates additional data that are typically difficult to collect such as fish discard rates, fishing depth, effort estimates, and socioeconomic parameters that will help optimize the fishery’s full potential from

both a harvest and an economic perspective. The concept of electronically collected and self-reported data certainly has many challenges and limitations; however, this pilot showed *iSnapper* has the ability to overcome many of these obstacles while generating realtime, validated, and reliable private recreational catch data for fisheries managers – something needed by all groups involved to improve access to the Red Snapper fishery.

A key component to this study was the validation of catch data submitted by anglers, with creel interviews conducted to verify accuracy of user-entered information. These data were compared to trips submitted via *iSnapper* using the vessel registration number as the key linking parameter. A majority of the data analyzed within this report occurred during the federal Red Snapper season (June 1 – 10), because there was an increased number of creel surveys and fishing activity during this time allowing for the best probability of intercepting an *iSnapper* user for validation. From the validated trips we calculated a reporting and harvest error rate and developed an estimator (specifically for this project) to calculate the total number of Red Snapper harvested, and the total number of private recreational angler trips in Texas during 2015.

iSnapper was introduced to anglers via extensive outreach and advertising campaign employing TV, radio, print, and social media. A total of 163 trips were submitted using *iSnapper* by Texas private recreational anglers harvesting a total of 2,012 fish during the federal season. Red Snapper was the most dominate species captured (1,519; 75.5%), with all trips reporting at least one Red Snapper harvested. A variety of other species were also caught, with King Mackerel and Dolphinfish as the other most commonly harvested species. The mean reported discard rate for Red Snapper was 56.1%, with the highest rate of 74.3% occurring at depths between 21 – 30 m.

A total of 969 private recreational Red Snapper angler-trips were encountered during surveys at Texas boat ramps. The creels from these anglers represented 259 fishing trips and harvested a total of 2,268 Red Snapper. To validate the self-reported data, these trips were compared with those submitted using *iSnapper*. The sampling was done using a mark and recapture approach, with *iSnapper* users being recaptured during the creel survey. Of the *iSnapper* trips, a total of 18 were validated during the dockside interviews, generating an 11% validation rate. From these validated trips, we calculated an overall reporting rate of 4.1% as well as an error rate of +5.1%. Using the reporting rate and error estimates from the federal season, along with creel data provided by TPWD for the entire year, the estimated total number of Red Snapper harvested by Texas private recreational anglers was 58,251 fish (25,344 SE) weighing an estimated 277,752 lbs by 23,358 angler-trips (6,660 SE) in 2015. Although not the focus of this study, we had an additional 13 charter for-hire trips submitted using *iSnapper* throughout summer 2015, which was surprising given the popularity of previous versions of *iSnapper* piloted with the for-hire Federal Reef Fish permit holders. Different than private recreational anglers, charter captains had a longer federal Red Snapper season (40 days; June 1 – July 14) and during this time reported harvesting a total of 76 Red Snapper, 25 Dolphinfish, 6 King Mackerel, 1 Warsaw Grouper, and 1 Yellowedge Grouper (seven total trips). Six additional trips were submitted in August from state waters with a harvest of 128 Red Snapper, and no other species were reported.

In addition to catch and effort data, the app also collected socioeconomic information from participants. The survey was a separate feature built into the app that allowed anglers to fill out and submit the survey, but was not linked to particular fishing trips. A total of 100 surveys were completed with 95 unique respondents. Approximately 98% of the respondents were male, with 93% residing in Texas. The average distance traveled per trip was 89.3 miles (7.3 SE) with a mean expenditure in bait and tackle for the trip was $197.40 ($34.22 SE), and boats consumed on average 82.4 gallons of fuel (12.0 SE). These results demonstrate the utility for smart devices to collect these types of socioeconomic data that are essential to valuating the fishery.

Overall, this project demonstrated that smart device applications can be successfully designed to collect catch and effort data from the private recreational fishing sector to enhance and supplement current data collection approaches. An advantage of this approach is the timeliness of the data collection provided, particularly in circumstances of shorted seasons, where traditional survey methods may not be as reliable. We were able to streamline data collection by gathering all the pertinent information in only a few screens making the data entry quick and easy, allowing anglers to report catch information for multiple species, as well as discard mortality, depth fished, and a variety of socioeconomic components of the fishery. The program was also voluntary, and that may contribute to a lower than desired reporting rate. Although mandatory reporting certainly does not guarantee 100% compliance, comparisons should be made across states where reporting is mandatory to determine how this influences the accuracy of the estimates. However, even without mandatory reporting, we found that smart device app technology has great potential to collect valuable catch and effort data quickly and efficiently from the private sector. In addition, harvest and catch estimates can be calculated using data collected from *iSnapper* if there is a robust validation component, wherein anglers are also creeled following their trip entry. Certainly, submitted trips should meet a thorough validation process to determine both reporting and error rates, which can then be used to estimate the total harvest. Through our partnership with TPWD, we had access to creel data for trips harvesting Red Snapper which was essential to the process. Considerations (both for effort and expense) must be given regarding collecting this data if a partnership such as ours with a state agency is not available, as the cost for additional validation surveys could become substantial for these types of initiatives. Finally, while this pilot was specifically targeting Red Snapper anglers, *iSnapper* has the potential to improve management for a variety of fisheries, and anglers are willing to enter this additional information as well. By combining these smart device technologies with traditional fish survey methods, managers have improved tool sets to gather more information and make the most informed decisions.

3. Background

A. Description of the problem

One of the major challenges to fisheries management is the ability to collect timely catch data from the recreational fishing sector. Recently, management measures have led to an increasing need for more timely and accurate estimates of recreational catch and effort data for assessing stocks (Griffiths et al. 2010). The problem is further compounded with shortened seasons and the need for rapid in-season measurements of catch. The lack of timely and robust data from this sector has created problems when fisheries managers calculate the annual harvest quota well after the season has ended and this has led to major conflicts among users for species such as Red Snapper (*Lutjanus campechanus*). This species is highly sought after by both commercial and recreational fishers and its management is one of the most controversial in the United States. For example, the inability to rapidly gauge recreational catch has resulted in the sector exceeding the allocation for the past 21 of 24 years. A recent 2014 federal court ruling resulted in federal managers implementing a 20% buffer to prevent the overages. Compounding the problem, anglers are also catching larger snapper each year; thus, reaching the quota faster than in previous years. All of these factors have led to very short federal fishing seasons (10 days in 2015, 9 days in 2014), despite anglers seeing a resurgence of Red Snapper. These shortened seasons hinder the ability of traditional approaches to collect accurate data from recreational anglers, because they were not designed to collect data in this manner. The Marine Recreational Information Program (MRIP) has modified their sampling protocols to increase the amount of data collected from anglers during these short windows. However, there is still a need for rapid in-season and near real-time data collection. Here we developed a novel data collection tool for the private recreational sector using smartphone/tablet applications ("apps") and web-based data entry portals.

In the past decade there has been a dramatic increase in the number of people using mobile phones. With this emerging technology, a new avenue of data collection was developed using apps as the platform for data collection. The concept was initially tested in the fisheries field by having recreational anglers submit text messages of their catch and effort data (Baker and Oeschger 2009). However, the 160 character maximum severely limited what could be included in the message, making it difficult to report an entire day of fishing. Currently, nearly two-thirds of Americans own a smartphone (Pew Research Center 2015), so the next logical step is to move from a text message data collection system to creating a specialized app that can be created and then downloaded onto the phones. More recently, smart devices have redefined the technological market allowing users to do a multitude of operations including accessing the internet using cellular data; thus, there is a high potential to use smart devices to collect more informative fisheries catch data.

Collecting data via smart device technology incorporates another recent trend using citizen scientists (individuals that are amateurs or nonprofessionals) to collect a substantial amount of data for relatively little cost. The data submitted by these citizen scientists are considered "self-reported," because these individuals are reporting without any direct validation from state or federal managers. The benefit in collecting and using self-reported data is that these citizen scientists provide managers with data that would otherwise be unavailable and they feel a sense of empowerment by being able to contribute to the conservation and management of their natural resources (Cohn 2008). Scientists are also recognizing the potential of self-reported data from smart device apps that have been created merely for entertainment purposes. One such app, iFish, is essentially a catch log for freshwater anglers throughout the United States and Canada. Catch and effort data is submitted by the user and this information can be used by fisheries managers to better understand local hotspots or how fishing pressures change depending on season (Papenfuss et al. 2015). We proposed and tested the potential to use this technology in the private recreational Red Snapper sector.

The Red Snapper fishery in the Gulf of Mexico is an ideal testbed to examine the feasibility of a voluntary smart device data collection app for private recreational anglers. Due to the limited federal Red Snapper season, fishery managers must rely on collecting as much data as possible from this sector while the brief season is open. Creating an app not only provides fisheries managers with near real-time data, but it can also collect a multitude of other important information (e.g. socioeconomic data, release mortality, etc.). A lack of timely data hinders the management of this fishery because data generated from directed creel surveys takes months before it is transcribed, edited, reviewed, and available for management advice. During this time, data submitted using an app could be analyzed and the total harvest could be estimated in a much more rapid fashion, allowing for in-season monitoring of the recreational harvest, which could potentially increase the season length, and reduce the 20% management buffer.

While an app provides the ability to collect more robust and timely harvest data, validating the quality is of paramount importance. The most critical and informative validation measure is to visually inspect the entire catch when it is landed dockside to confirm that submitted catch reports are consistent. This validation allows managers to calculate the reporting rate as well as error estimates based on anglers who may have, for example, misidentified certain species of fish or have inaccurately reported their total harvest. These estimates can then be extrapolated to all of the trips that were interviewed at the boat ramps to calculate more accurate total harvest estimates.

Given the potential of smart device apps to improve the management for many fisheries, our aim in this study was to supplement the current MRIP data collection program by developing and testing new technologies to enhance recreational fisheries data collection. In 2011, Harte Research Institute for Gulf of Mexico Studies (HRI) released *iSnapper* for use in the charter for hire industry and had overwhelming success with the project with major buy-in and support from participants (Stunz et al. 2016, in prep). Due to this success, the original concept was redesigned to create an app that could be used in the private recreational sector as well as for charter captains.

B. Objectives of the project
The specific objectives of this study were to:
1) Develop and implement *iSnapper* as a data collection app (for Apple, Android,and Windows platforms including a web portal) for private recreational anglers in the Gulf of Mexico;
2) Compare *iSnapper* data from panels of private anglers to TPWD creel survey data to validate the applicability of electronic data collection;
3) Collect and assess socioeconomic data from reef fish fishery participants using *iSnapper*;
4) Provide *iSnapper* as a data collection tool for NOAA-approved programs targeting Red Snapper in the Gulf of Mexico.

4. Methods

<u>*iSnapper* development</u>
Despite creating a very successful prior version of *iSnapper* (v1.0), it was necessary to redesign the platform to create a submission process easy and aesthetically pleasing for private recreational anglers that were not as incentivized as for-hire captains to enter catch data. Thus, working with Elemental Methods, LLC we recreated *iSnapper* (v2.0) on Apple's iOS®, as well as two new platforms, Android® and Windows®. Both smartphone and tablet versions were created for each platform to give individuals different and the most comprehensive entry options.

<u>Application Architecture</u>
The *iSnapper* v2.0 app was created and built upon *iSnapper* v1.0 and used to collect catch and effort data from private recreational angler boat owners, as well as charter boat operators, throughout the Gulf of Mexico. This version was specifically adapted to be most suitable to private recreational anglers; however, most of the features and data collection options were

available to continue to gather these data in the for-hire sector. Anglers were asked to enter their catch data by “adding” the species captured from the provided list of all the commonly caught species in the Gulf of Mexico (Figure 1) that included easy to select images. Once a species was selected, they provided the number harvested and released individuals. The average depth fished and general fishing area were also required fields and indicated by clicking on an image of the Gulf of Mexico. Effort data was gathered by providing the number of anglers on the vessel, and fishing times were also collected. Several new features were also implemented to build a multi-functional app in a very user-friendly environment to promote use by private anglers. Some of these features include current weather and tide information based on location, the ability to submit pictures of unidentified fish directly to researchers, and each trip can be shared on the individual’s social media networks (Facebook® and Twitter®).

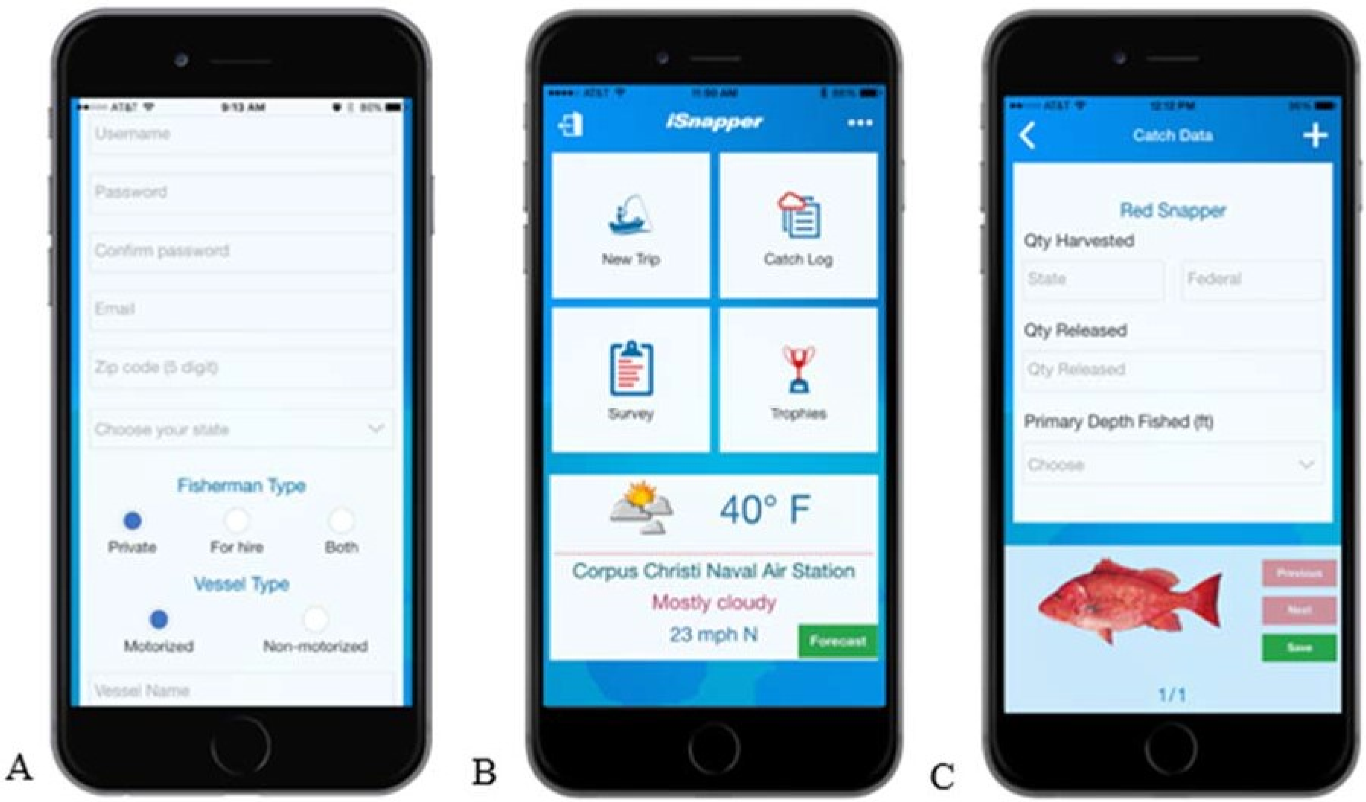

Figure 1. Screenshots of the registration screen (A), home screen (B), and the species catch screen (C).

<u>Web Portal</u>

Anglers were not limited to only using smart devices to submit their trips. We also created an online *iSnapper* webportal (https://isnapperonline.org, Figure 2) that anglers could use if they did not have a smart device, or encountered problems submitting their catch using the app. This option provided anglers with the opportunity to register or login using the same username they created when registering on the *iSnapper* app and enter their catch information online. Additionally, the webportal allowed anglers to login and view their catch data and saved photos from previous trips. The webportal was also designed to store all user and trip information for administrator access and data download as needed throughout the season.

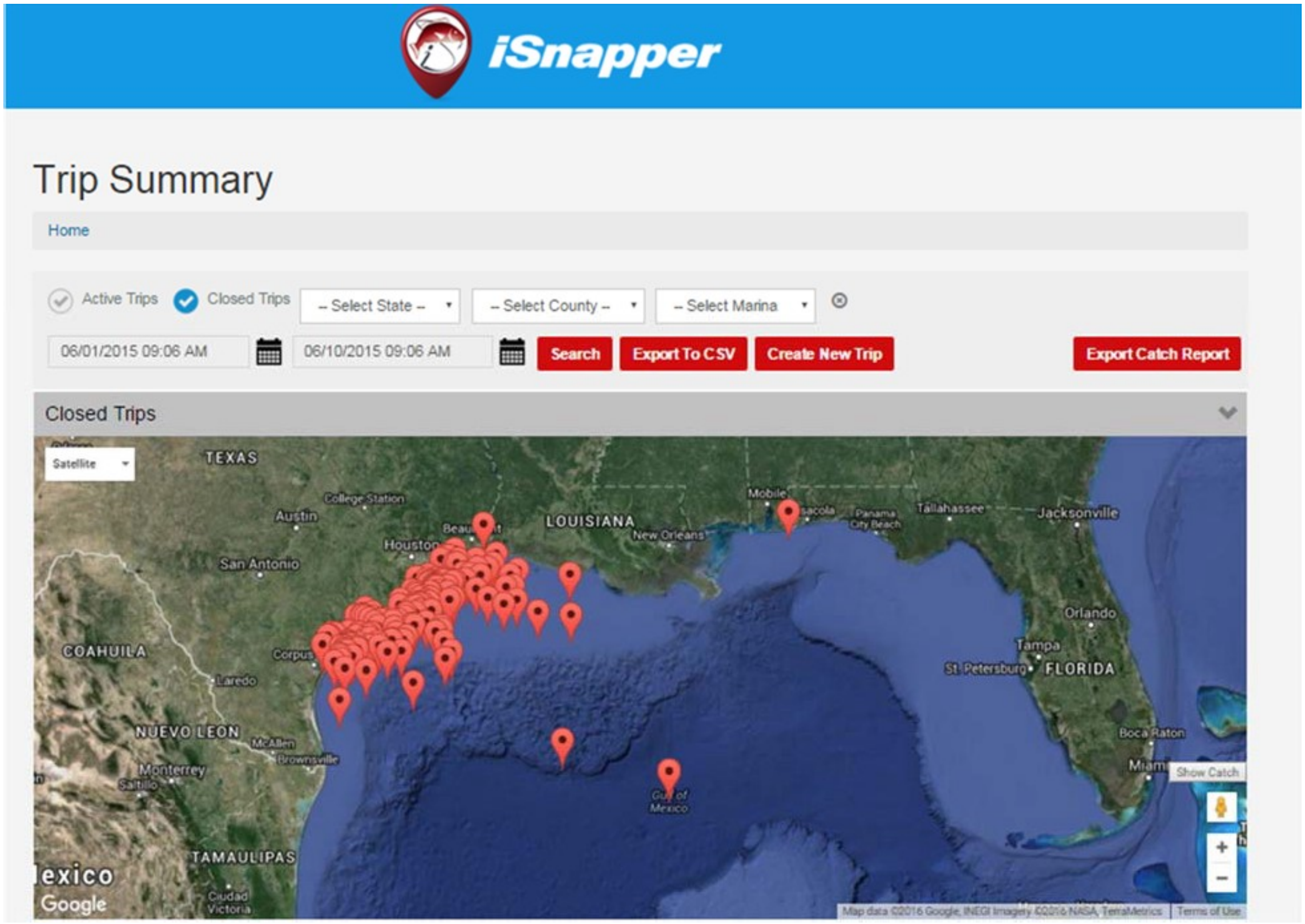

Figure 2. Image of the webportal with all trips from the federal season displayed.

<u>Panelist Selection</u>

A crucial portion of this study involved the validation of self-reported data. Originally, we planned on selecting a defined list of panelists from known Red Snapper anglers that represented a variety of angling frequency types using a stratified systematic random sampling design and then target these panelists at boat ramps during the Red Snapper season. However, this plan represented some insurmountable obstacles given this was the first pilot of this type. For example, we quickly discovered that a robust database of individual Red Snapper anglers that we could draw from did not exist, these anglers could not be assigned to ports of origin, and there was no information regarding their angling frequency. Thus, this initial plan was reconsidered and a new approach, and in hindsight a much better method, was developed after extensive consultations with MRIP statisticians (Lynne Stokes, Ph.D.) wherein we created a more inclusive design that involved collecting data from all *iSnapper* users, which was used as the initial “panel.” Now future studies have the ability to draw from this created panel of users that was built from known Red Snapper anglers. Briefly, the overall design was based on the well-established mark and recapture theory (Laplace 1786), where *iSnapper* registered users were considered “marked,” and then “recaptured” during boat ramp creel surveys after completing a fishing trip. The initial identification of anglers was randomized by sending out postcards, media outreach, and distributing informational wallet cards to as many anglers as possible without prior knowledge of their willingness to participate in the study or with what frequency they fish for Red Snapper.

Outreach

In collaboration with media and outreach partners at TPWD, we sent two postcard flyers that provided information about *iSnapper*, to as wide of an audience as we could identify. This included the 610 known Texas Red Snapper anglers that had been identified in their long-term data set along with mailings to charter captains and other participants in the *iSnapper* v1.0 pilot. The private anglers were encountered by TPWD creel surveys who had captured at least one Red Snapper on their trip during the last five years. A second mailing was also sent near the season opening to alert anglers that *iSnapper* was available for download and use. To further increase the number of *iSnapper* users, we also advertised in several state and local magazines, radio, and television news and public service segments. We also created an informative webpage (www.iSnapper.org) separate from the data collection portal, and these information sites were pushed extensively through social media avenues (e.g., Facebook® and Twitter®) by both HRI and TPWD and created an account on two of the most popular saltwater fishing forums to inform anglers about the app. In addition, TPWD produced flyers, wallet cards, and laminated signs that were distributed to bait shops and anglers several weeks prior to the opening of the federal Red Snapper season. Before the start of the season coordinators of groups such as the Coastal Conservation Association Texas, Saltwater Enhancement Association, and Texas Sea Grant volunteered to educate their members and contact them about the app and encouraged them to submit their catch using *iSnapper*.

Validation

Validation was performed at boat ramps by both TPWD surveyors and HRI staff by creeling as many boats as possible to intercept private recreational anglers using *iSnapper* after a fishing trip. During the private recreational federal Red Snapper season (June 1 – June 10) 7 additional TPWD creel surveys were conducted with the intent to “recapture” as many *iSnapper* users as possible. Additionally, to augment creeling effort HRI staff conducted 5 surveys at high use marinas and boat ramps during the federal Red Snapper recreational season, and TPWD increased the number of random creel surveys throughout the ‘high use’ season (May 15 – November 20, 2015) from 764 in 2014 to 832 in 2015. Despite these targeted creels at high use sites, the anglers were still randomly intercepted, because interviewers did not know if any *iSnapper* trips had been started or submitted prior to the creel survey. During the interviews, one angler (typically the captain or designee) from the boat was asked to report the catch and effort data for the entire boat, including how many Red Snapper were harvested, the number of anglers on the boat (angler-trips), depth fished, and if they had reported their catch using *iSnapper*. For example, if 4 boats came in and there were 4 anglers fishing on each boat that would be considered 4 trips and 16 angler-trips (i.e., there was only one report for each vessel to avoid duplicate reporting). The accuracy of data submitted with *iSnapper* was validated by crossreferencing the creel surveys using vessel registration numbers to determine if their reported catch was the same as what was recorded dock-side. Certainly, by maximizing the numbers of validations that could be performed, the most accurate catch estimates could be determined. Anglers that were encountered not using the app were also surveyed, and they were informed about *iSnapper*, the value of using it, and were highly encouraged to download and use it for the duration of the federal and state Red Snapper seasons.

Catch Estimation

The traditional method for estimating recreational catch for most species and locations uses two complementary surveys of anglers, one to measure “effort” (number of fishing trips and/or angler-trips) done by phone or mail, and one to measure mean catch per trip, done face-to-face with dockside interviews. Use of electronic reporting allows effort estimates to be reported by the anglers on the day when the fishing actually took place, reducing problems with inaccurate estimates due to recall bias. However, using this data requires a validation process to monitor how accurate the reporting is. Anglers were encouraged to submit their trip data prior to arriving back in case they were intercepted by TPWD or HRI at the dock to prevent any bias and ensure independence in the self-reported data and validation process. With both the self-reported data and the validation, the population and sample data can be broken down into four categories (Figure 3) to calculate the number of trips and Red Snapper harvested. All of the categories, aside from the ‘not reported or creeled,’ are used in a new estimator developed specifically for this project to calculate the reporting and error rates to estimate the total Red Snapper harvest for private recreational anglers using self-reported data. The development of a new estimator was required, since these types of data have yet to be included in catch and effort estimates, and there are several assumptions which must be met for it to be valid. The fundamental assumption is that the event of being captured does not change the probability of being recaptured (and/or vice versa), at least within each identifiable subset of the population. The only other assumption is that anglers are selecting the correct species harvested, in this case Red Snapper. To calculate the total harvest or total number of anglers, the following equation was used:

$$t_{y2} = t_{y*} + n_1/\hat{n}_1(t_y - t_{y*}) = t_{y*} + n_1$$

where is an estimator of = $(t_y - t_{y*})/n_1$ which is the total population underreport averaged over the units in the reporting domain (i.e. the *iSnapper* reports). In the formula, t_{y*} is the reported removals of Red Snapper (or reported number of angler-trips) based on the *iSnapper* app. n_1 is the number of vessel trips which reported their Red Snapper catch using *iSnapper*. t_y and t_y are the estimated Red Snapper catch (or number of angler-trips) of the whole population and the reporting domain, estimated from the validation sample only. The equation above is the best estimator (t_{y2}) for these data, and while details on the derivation can be found in Liu et al. (2017; Appendix 1). The estimator uses the catch report from *iSnapper* from the angler only as auxiliary data from which the angler-trip and harvest estimates are calculated. With the validation of the creel survey, any errors in the self-reported data are then corrected and those error rates are then applied to the total estimates.

		Sample Validated	
		Yes	No
Reported	Yes	Reported and Validated	Reported, not creeled
	No	Not reported, creeled	Not reported or creeled

Figure 3. Illustration of the population and sample data.

<u>Registration</u>

Anglers were able to download *iSnapper* at the App store (iOS) and Google Play (Android). Once downloaded, the first step in the data submission process involved anglers registering to set-up their *iSnapper* account (Figure 1A). At registration, participants provided their vessel registration numbers, giving a unique identifier critical for validation. Also, contact information was collected to allow administrators to contact anglers to resolve any observed errors. Once registered, the angler was able to immediately enter and submit their catch information from fishing trips. The process to submit a trip involved 3 simple steps (Figure 3). All of these steps could be done in less than five minutes, and typically within two minutes depending on the number of different species caught during the trip. Steps:

(1) Open the app and provide the date, time, marina/boat ramp the boat was launching from, and the number of anglers;
(2) Fill in the catch data by selecting the species caught and entering number of fish harvested and released; and
(3) End the trip by selecting a general fishing location on a map and the primary depth fished for the trip and submitting.

There were several features that made the process easy, streamlined, and as user friendly as possible. For example, the date and time were automatically populated for the current time both when starting a new trip and closing a trip, but could be adjusted if the angler forgot to create a trip before leaving the dock. Additionally, when anglers harvested Red Snapper the app divided the catch into two categories: fish harvested in state waters and fish harvested in federal waters. Additionally, Red Snapper anglers were required to report the primary fishing depth on the catch screen. All species commonly captured throughout the Gulf of Mexico were included in the app, allowing anglers to submit their entire catch not only Red Snapper. Once the trip was finalized and submitted anglers could not edit nor delete their entry in the app or on the webportal. The only way to change trip information was to email HRI and have one of the researchers log in and adjust the trip. This was very important because to calculate an accurate error estimate, we compared the number of fish reported to the number of fish counted during dockside interviews. If anglers were capable of changing their catch information after submission, calculating the error rate would not be valid because anglers would have the ability to change the number of fish submitted on *iSnapper* if they were interviewed at the boat ramps by TPWD or HRI staff. Finally, all required app updates were “pushed” the user’s phone as needed.

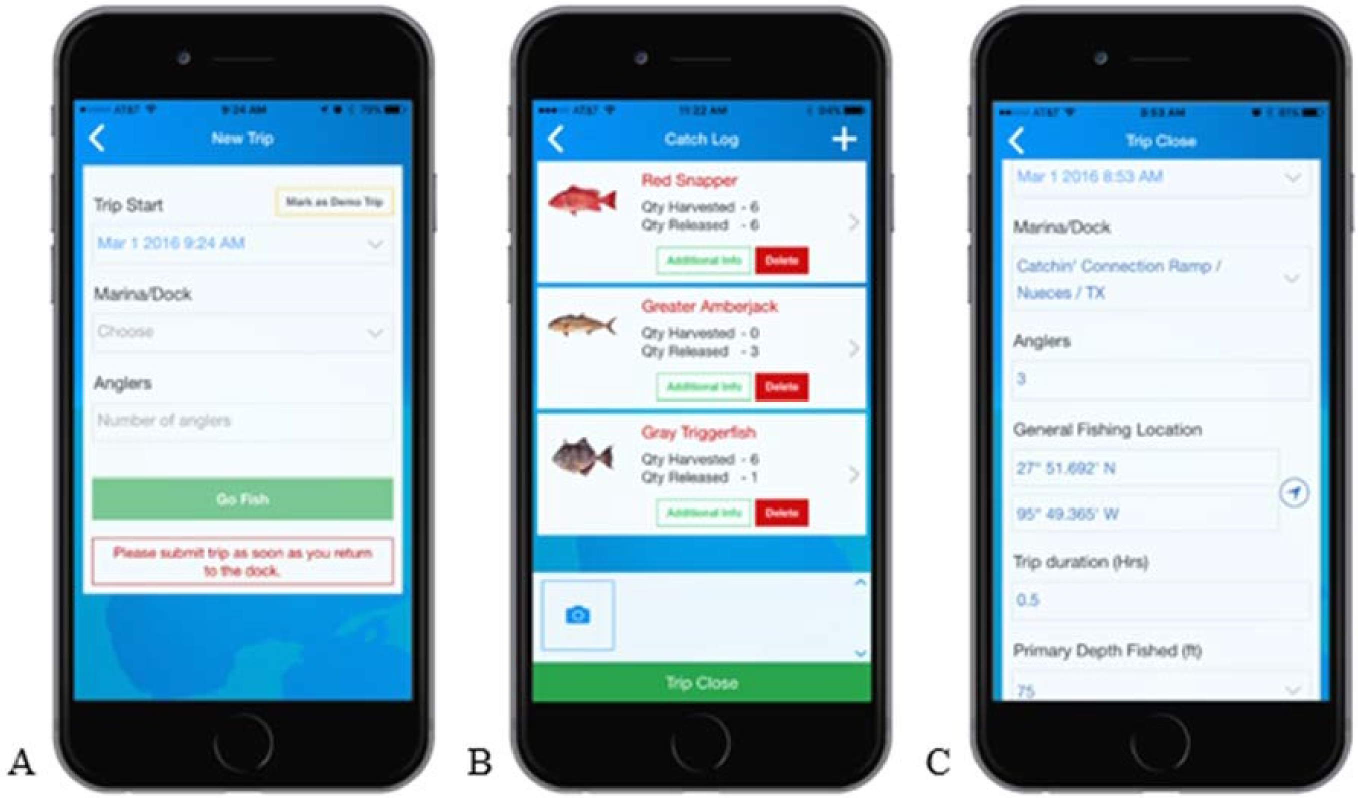

Figure 4. Screenshots of the new trip screen (A), catch log (B), and trip close screen (C).

Socioeconomics Survey

While the previous three steps were required of all trips, there were other optional features that anglers could choose to enter. One of the features that proved to be very beneficial is the availability to collect additional socioeconomic information. The socioeconomic survey was a separate optional feature in the app and available on the home screen. Questions in the survey were similar to those used in the previous version of *iSnapper*; however, after receiving feedback from NOAA, we included additional questions to get more refined information regarding the cost of trips and distances traveled (Table 1). Even with these additions, the survey was brief and took less than five minutes.

Table 1. Socioeconomic questions included in *iSnapper*.

How many people in total, including yourself, live in your household? Please include those people who fish and who don't fish.
How many people in your household, including children and adults, have been recreational saltwater fishing in the last 12 months anywhere in the Gulf of Mexico region including inshore and offshore?
How many days did you spend saltwater fishing in the last 12 months?
How many of these days were spent offshore?
If this fishing trip is part of a longer trip in which you will spend at least one night away from your permanent residence, how many days will this trip last?
What is your primary and secondary (if applicable) zip code? (Enter zip codes separated by comma Ex: 12345, 12346)

Gender
What is the total distance traveled by boat during this trip? (Miles)
Do you keep your boat at a marina or trailered?
What is the estimated bait and tackle expenses for this trip?
What is the horsepower of your boat?
What is the estimated fuel consumption used for this trip? (Gallons)
Which of the following best describes your household's annual income, before taxes? (US$)

For-Hire provision

Given the successful *iSnapper* pilot study in the for-hire sector and groundswell of interest by others, many groups routinely inquired as to the availability of its use. Thus, we redesigned *iSnapper* to include fields specifically for the for-hire captains so they could continue to submit trips, and the registration process enabled us to distinguish between private and for-hire trips. If a user selected the for-hire option during registration, state and/or federal permit numbers were required to complete the process. For-hire reporting followed the same format as with private anglers, and these trips were also validated at boat ramps, but for the majority of the analyses were not included due to the small number of for-hire trips reported.

Data from subsequent years changes to reduce variability

As the utility of *iSnapper* was shown during the 2015 study, subsequent funding for the continuation was secured for 2016 through Texas Parks and Wildlife and in 2017-2018 through the National Fish and Wildlife Foundation. While not the direct focus of this study, the authors and MRIP felt it was important to include here for continuity and for comparison purposes, although MRIP funding was not used for these portions of the study. Nevertheless, valuable information was gained, and we felt it was important to include here as a way to improve electronic data collection. The results from 2015 were characterized by a high degree of variability, and we were certainly interested in reducing this variability to improve the precision of our harvest estimates. Thus, when working with MRIP statisticians, it was recommended we modify the sampling methodology to incorporate more randomness to improve accurate estimates. The issue for 2015 was that the encounter rate for trips using *iSnapper* was unknown, so frequenting high-use boat ramps was essential to “recapture” anglers on these trips. However, we have no a priori idea of what this recapture rate would be in 2015, and since identifying these groups was essential for validation this was the preferred methodology for that particular year in terms of maximizing encounter rate. In 2016, HRI further increased sampling effort by sending additional staff to the most heavily used Red Snapper boat ramps for the majority of the season. Similarly to 2015, the goal was to encounter as many anglers fishing for Red Snapper as possible, and encourage them to download and submit their catch and effort information. The 2016 sites were selected based on previous Red Snapper catch data (as in 2015) with no randomization of boat ramp selection. However, even with this increase creel effort, harvest estimates did not improve (discussed later; see below). Therefore in 2017, our sampling methodology was modified to expand the Texas Parks and Wildlife proportional random sampling of Red Snapper sites and increase the creel effort in as many randomly selected locations as possible. This change was successful and dramatically lowered the variability. The TPWD determines pressure estimates based on roving counts of empty trailers and empty wet slips at inventoried boat-access sites. Survey sites are selected randomly but selection is weighted according to mean rove counts adjusted by the percentage of target-area fishing activity. For the initial 3 day season, 6 or 7 (depending on staff availability) sites were randomly drawn per day, excluding sites that TPWD had already drawn to sample. During the 39-day extended season, 21 days were randomly selected, two-thirds were weekend days and one-third were weekdays. Then, sites were selected the same way as the initial season, with two sites selected per day.

5. Results

<u>Development and implementation of *iSnapper* as a data collection app:</u>
The adaptation of *iSnapper* from a mobile application targeting for-hire captains to one that could be universally used by all recreational anglers was very successful. Since the release on May 15, 2015 the app was downloaded on 945 different devices through the end of 2015. The majority of the users (71%, 672 downloads) were operating a device with iOS (Apple®) platform, the remainder were Android-based users. A total of 393 individuals registered to use the app, with 199 users providing valid vessel registration numbers. During the initial development stages, the Windows® platform was an appealing operating systems, and we had anticipated high number of users. However, options for app development and the subsequent phase out of this platform by most developers lead to little interest, and we delayed implementing this platform to focus on the other two more popular formats. Moreover, with only 3% of all cell phone users listing Windows® as their phone type (Pew Research Center 2015), and that number rapidly declining, we do not recommend development of this platform for future data collection.

During the 10 day Federal Red Snapper season (June 1 – 10, 2015) there were 171 trips submitted using the app or the online web portal, with 163 trips from Texas private recreational anglers (Table 2). Red Snapper was the most dominant species captured, with all trips reporting a harvest of at least one Red Snapper. Other species commonly captured were King Mackerel (*Scomberomorus cavalla*), Cobia (*Rachycentron canadum*), and Dolphinfish (*Coryphaena hippurus*) (Table 3). A total of 2,012 fish were harvested during the federal season, with 75.5% of the harvested fish being Red Snapper. The next most prominent species harvested were Dolphinfish, King Mackerel, Vermilion Snapper (*Rhomboplites aurorubens*), and Blackfin Tuna (*Thunnus atlanticus*). Most trips (private and for-hire) using *iSnapper* were located within the continental shelf, generally within 100 nm offshore Texas (Figure 4A). Most vessels harvested 7 Red Snapper from their selected fishing locations, while some harvested 25 – 56 Red Snapper in these general areas throughout the season (Figure 4B). Despite a federal bag limit of 2 Red Snapper per angler, anglers were able to keep a maximum of 4 fish if 2 were from state waters. To calculate the estimated pounds of Red Snapper harvested the average length of fish recorded by TPWD was converted using the TPWD length/weight conversion chart (http://txmarspecies.tamug.edu/length-weight.cfm) and multiplying the weight by the total number of fish harvested. The TPWD methodology for measuring and estimating fish length is robust, and at each creel, they measure up to 6 fish of each species. Thus, our staff at HRI concentrated more on validating the number of fish harvested, which for our purposes was the most important parameter. We did not measure fish, and the additional time saved was instead focused on interviewing more boats. To use the additional data collected (*iSnapper* and HRI creels) we used estimates of the average size for state and federal water fish based on TPWD creels, and then applied that to estimate the total harvest including *iSnapper* and HRI creels.

An additional 22 trips were started by Texas anglers during the federal season, but not completed. Despite including a feature to alert anglers if they had a trip open longer than 24 hours, as well as reaching out to these anglers via email on multiple occasions, these trips were never submitted.

Although it was an abbreviated 10-day season, the weather conditions were optimal for offshore fishing. Light winds and small seas enabled most vessels to get out to fish federal waters (> 9 nautical miles), especially for some of the smaller (< 25') boats. The National Weather Service issues small craft advisories starting at wind speeds greater than 12.9 m/s. Average wind speeds throughout the federal season were never greater than 4.6 m/s and average wave height did not exceed 0.7 m (Table 2). With the conditions being relatively similar throughout the season, we were able to use the creel and app data to determine what days of the week corresponded with increased angler activity during the limited season. A majority (60%) of anglers fishing for Red Snapper went out on one of three days during the season: opening day (Monday), or the following Friday and Saturday (Figure 5). Not surprisingly, the day with the highest fishing pressure (Saturday) corresponded to the highest estimated daily harvest of 5,314 lbs.

Table 2. Number of self-reported trips using *iSnapper* from private recreational anglers in Texas during the 2015 federal Red Snapper season (June 1st - June 10th). Trips refers to the number of users that submitted fish captured during the season. Angler-Trips includes all individuals on the boat that were targeting Red Snapper for at least a portion of their trip. Total Released is the number of Red Snapper captured in either state or federal waters but discarded either due to size or bag limits. Harvested State is the number of Red Snapper harvested from state waters. Harvested Federal is the number of Red Snapper harvested from federal waters. Daily Harvest is the combined number of Red Snapper harvested from both state and federal waters. The asterisk (*) indicates what would be considered the weekend for a typical Red Snapper angler (Friday – Sunday).

Day	Trips	Angler-Trips	Total Released	Harvested State	Harvested Federal	Daily Harvest	Average Wind Speed (m/s) (SE)	Average Wave Height (m) (SE)	
6/1 (Mon)	23	84	191	22	178	200	2.4 (0.11)	0.4	(0.02)
6/2 (Tue)	22	80	272	40	155	195	1.9 (0.18)	0.3	(0.01)
6/3 (Wed)	17	55	300	11	97	108	3.6 (0.19)	0.3	(0.01)
6/4 (Thu)	11	37	230	17	74	91	4.6 (0.16)	0.7	(0.2)
6/5 (Fri)*	23	105	259	24	206	230	4.2 (0.21)	0.4	(0.2)

6/6 (Sat)*	27	124	309	62	227	289	2.7 (0.25)	0.4	(0.1)
6/7 (Sun)*	19	89	148	25	174	199	4.0 (0.25)	0.4	(0.1)
6/8 (Mon)	10	40	115	12	79	91	4.1 (0.28)	0.5	(0.2)
6/9 (Tue)	7	30	96	10	60	70	3.2 (0.17)	0.4	(0.2)
6/10 (Wed)	4	19	20	8	38	46	4.2 (0.22)	0.5	(0.2)
Total	163	663	1940	231	1288	1519	3.5 (0.08)	0.4	(0.006)

Table 2b. Number of self-reported trips using *iSnapper* from private recreational anglers in Texas during the 2016 federal Red Snapper season (June 1st - June 11th).

Day	Trips	Angler-Trips	Total Released	Harvested State	Harvested Federal	Daily Harvest	Ave Wind Speed (m/s) (SE)	Ave Wave Height (m) (SE)
6/1 (Wed)	5	17	51	8	28	36	5.7 (0.88)	1.3 (0.06)
6/2 (Thu)	3	6	6	2	10	12	6.2 (0.42)	1.0 (0.02)
6/3 (Fri)*	8	30	209	14	50	64	5.5 (0.29)	1.0 (0.02)
6/4 (Sat)*	6	21	21	16	42	58	4.7 (0.67)	1.0 (0.04)
6/5 (Sun)*	12	56	58	0	92	92	6.7 (0.40)	0.9 (0.06)
6/6 (Mon)	12	46	154	4	82	86	4.9 (0.45)	0.8 (0.04)
6/7 (Tue)	13	49	153	3	77	80	2.8 (0.22)	0.7 (0.02)
6/8 (Wed)	14	53	159	12	98	110	4.3 (0.32)	0.7 (0.02)
6/9 (Thu)	8	34	81	16	55	71	3.8 (0.28)	0.7 (0.01)
6/10 (Fri)*	4	16	36	4	27	31	3.4 (0.22)	0.5 (0.01)
6/11 (Sat)*	10	58	97	28	116	144	6.1 (0.27)	0.7 (0.02)
Total	95	386	1025	107	677	784	4.9 (0.15)	0.8 (0.02)

Table 2c. Number of self-reported trips using *iSnapper* from private recreational anglers in

Texas during the 2017 federal Red Snapper season.

Date	Trips	Angler-Trips	Total Released	Daily Harvest	Ave Wind Speed (m/s) (SE)	Ave Wave Height (m) (SE)
6/1	8	30	68	40	6.14 (0.23)	1.07 (0.03)
6/2	14	65	98	106	4.38 (0.30)	0.72 (0.01)
6/3	11	55	122	96	3.74 (0.33)	0.91 (0.02)
6/16	2	6	8	6	7.90 (0.16)	1.13 (0.01)
6/17	0	0	0	0	9.27 (0.15)	1.41 (0.03)
6/18	0	0	0	0	9.59 (0.28)	1.79 (0.05)
6/23	0	0	0	0	9.83 (0.36)	2.10 (0.04)
6/24	1	4	0	8	6.93 (0.32)	1.60 (0.03)
6/25	2	9	6	6	5.65 (0.34)	1.08 (0.04)
6/30	0	0	0	0	9.27 (0.21)	1.64 (0.02)
7/1	1	2	10	8	9.25 (0.32)	1.78 (0.03)
7/2	2	4	5	5	7.20 (0.32)	1.24 (0.05)
7/3	2	13	30	26	7.77 (0.18)	1.05 (0.02)
7/4	1	4	6	8	8.59 (0.25)	1.29 (0.03)
7/7	2	10	16	20	4.53 (0.46)	0.71 (0.02)
7/8	13	57	185	107	2.08 (0.17)	0.54 (0.02)
7/9	9	46	160	94	3.58 (0.24)	0.43 (0.01)
7/14	3	15	40	44	4.81 (0.40)	0.66 (0.02)
7/15	11	50	49	91	2.85 (0.31)	0.41 (0.01)
7/16	6	24	18	29	4.54 (0.36)	0.36 (0.02)
7/21	4	21	18	42	5.82 (0.28)	0.66 (0.02)
7/22	2	11	33	21	7.41 (0.22)	0.91 (0.03)
7/23	0	0	0	0	8.89 (0.08)	1.40 (0.03)
7/28	0	0	0	0	5.53 (0.21)	0.64 (0.02)
7/29	3	12	1	23	6.30 (0.30)	0.73 (0.03)
7/30	2	13	12	26	5.82 (0.28)	0.62 (0.02)
8/4	3	12	12	24	5.48 (0.32)	0.72 (0.02)
8/5	0	0	0	0	7.45 (0.22)	0.92 (0.02)
8/6	0	0	0	0	9.00 (0.14)	1.24 (0.02)
8/11	0	0	0	0	4.27 (0.24)	1.43 (0.06)
8/12	1	2	5	8	6.58 (0.21)	1.00 (0.02)
8/13	0	0	0	0	7.89 (0.12)	1.04 (0.01)
8/18	1	4	6	16	5.53 (0.31)	1.11 (0.02)

8/19	3	10	19	22	4.97 (0.43)	0.83 (0.03)
8/20	7	22	67	27	5.28 (0.26)	0.63 (0.02)
8/25	Hurricane Harvey made landfall				16.54 (1.52)	4.84 (0.33)
8/26	No data submitted					
8/27						
9/1						
9/2						
9/3						
9/4						
Total	114	501	994	903	6.70 (0.30)	1.10 (0.03)

Table 3. Summary of the species captured and released as reported by private recreational anglers using *iSnapper* during the Red Snapper federal season (June 1 – June 10). Number captured includes the combined number of fish harvested and released. An asterisk (*) next to a species name indicates the species is considered a bait fish. A horizontal dash (-) in the discard rate column indicates all fish captured were harvested. Angler-Trips is not mutually exclusive by species, since several species are typically caught during one trip.

Species	Number Captured	Percent of Total Capture	Number Harvested	Discard Rate	Number Released	Angler-Trips
Red Snapper	3459	82.0%	1519	56.1%	1940	663
King Mackerel	139	2.4%	85	38.8%	54	148
Dolphinfish	119	1.0%	98	17.6%	21	99
Blue Runner*	57	0.7%	42	26.3%	15	21
Cobia	50	0.9%	30	40.0%	20	115
Gulf Menhaden*	50	<0.1%	50	-	0	2
Blackfin Tuna	48	0.5%	37	22.9%	11	20
Greater Amberjack	42	1.7%	4	90.5%	38	57
Vermilion Snapper	42	0.1%	40	4.8%	2	34
Great Barracuda	39	1.5%	5	87.2%	34	18
Atlantic Sharpnose Shark	23	0.8%	5	78.3%	18	14
Gray Triggerfish	18	0.6%	4	77.8%	14	30

Golden Tilefish	15	<0.1%	15	-	0	18
Atlantic Spadefish	14	0.2%	10	28.6%	4	10
Lane Snapper	10	<0.1%	9	10.0%	1	17
Crevalle Jack	9	0.3%	2	77.8%	7	14
Almaco Jack	7	0.1%	5	28.6%	2	8
Tripletail	7	<0.1%	6	14.3%	1	16
Rainbow Runner	6	0.1%	4	33.3%	2	9
Rock Hind	6	<0.1%	6	-	0	9
Bermuda Chub	5	0.1%	2	60.0%	3	7
Gray Snapper	5	<0.1%	4	20.0%	1	17
Little Tunny	5	<0.1%	5	-	0	10
Remora	5	0.1%	2	60.0%	3	11
Bonito	4	0.1%	2	50.0%	2	10
Blacktip Shark	3	<0.1%	2	33.3%	1	8
Sand Seatrout	3	<0.1%	3	-	0	6
Scamp	3	<0.1%	2	33.3%	1	10
Spanish Mackerel	3	<0.1%	3	-	0	6
Atlantic Bonito	2	<0.1%	1	50.0%	1	7
Dog Snapper	2	<0.1%	2	-	0	4
Lesser Amberjack	2	0.1%	0	100.0%	2	4
Spotted Seatrout	2	<0.1%	1	50.0%	1	2
Wahoo	2	<0.1%	1	50.0%	1	11
Yellowfin Tuna	2	<0.1%	2	-	0	4
Blue Marlin	1	<0.1%	0	100.0%	1	4
Bull Shark	1	<0.1%	0	100.0%	1	5

Cubera Snapper	1	<0.1%	1	-	0	4
Gag Grouper	1	<0.1%	0	100.0%	1	6
Great Hammerhead Shark	1	<0.1%	0	100.0%	1	3
Horse-eye Jack	1	<0.1%	1	-	0	3
Others	1	<0.1%	0	100.0%	1	5
Sailfish	1	<0.1%	0	100.0%	1	5
Sandbar Shark	1	<0.1%	0	100.0%	1	6
Scalloped Hammerhead	1	<0.1%	0	100.0%	1	4
Warsaw Grouper	1	<0.1%	1	-	0	4
Yellowedge grouper	1	<0.1%	1	-	0	4
Total	4220	100%	2012	52.3%	2208	1492

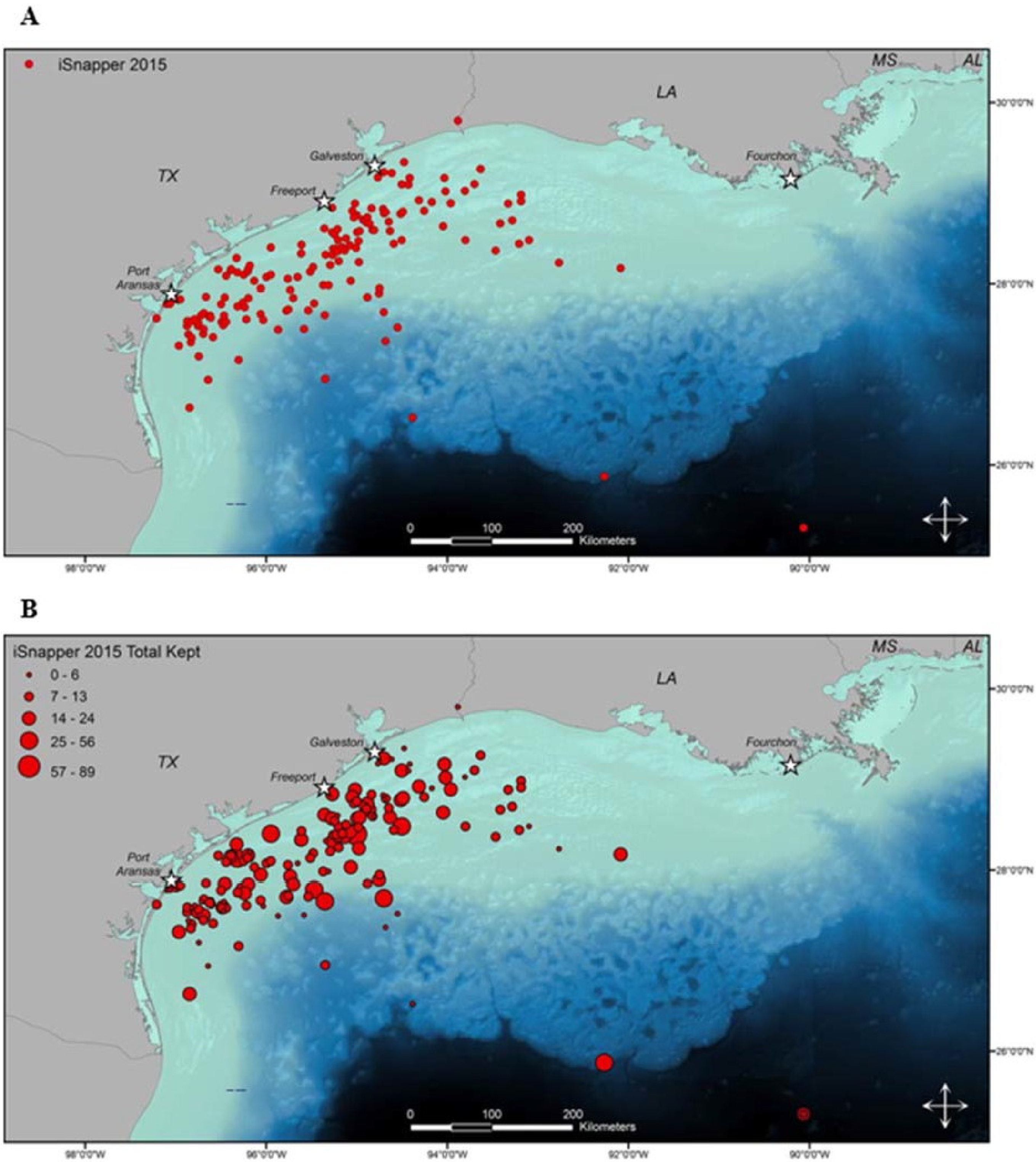

Figure 5. Approximate fishing locations (A) and total harvest by location (B) for *iSnapper* users during the 2015 Red Snapper season (June 1 – June 10). Locations on or outside of the continental shelf are likely errant entries, with the site on the shelf edge being the default location when the map loads on the app.

The majority of Red Snapper were caught at depths of 40 m or less (52.4%), with the most common reported depth of capture between 31 – 40 m (1,042 caught, 31.6%, Table 4). The overall estimated discard rate for Red Snapper was 56.1%, with the highest rate of 74.3%

occurring at depths between 21 – 30 m. A total of eight trips did not report their fishing depth, which accounts for the 100 Red Snapper released at unknown depths. The depth range where most fish were retained occurred between 71 – 80 m, although there were relatively few (n = 68) Red Snapper captured at these depths.

Table 4. Depth of capture for Red Snapper during the federal season (June 1 - 10) as reported using *iSnapper* by private recreational anglers in Texas. The discard rate by depth was calculated by dividing the number released by the total number captured. The overall discard rate was calculated by dividing the total number released by the total number captured. Unknown depth refers to trips where a depth of 0 ft was reported in the app.

Depth (m)	Number Released	Number Harvested	Discard Rate
1-10	10	8	55.6%
11-20	48	91	34.5%
21-30	399	138	74.3%
31-40	611	431	58.6%
41-50	596	600	49.8%
51-60	0	0	0.0%
61-70	112	82	57.7%
71-80	15	53	22.1%
81-90	0	0	0.0%
91+	49	67	42.2%
Unknown	100	49	67.1%
Total	1940	1519	56.1%

<u>Validation of self-reported data:</u>
During the 10 day federal season a total of 969 private recreational Red Snapper angler-trips were recorded at Texas boat ramps. Not all of these anglers were “unique” individuals, with some anglers going out on multiple days. These angler-trips represented 259 trips and harvested 2,268 Red Snapper (Table 5). Of the 163 trips submitted using *iSnapper*, 18 were interviewed at the boat ramp, generating an 11% trip validation rate and an overall reporting rate of 4.1% (see Appendix 1 for additional details on calculations). Two trips reported a higher number of fish harvested than were recorded during the creel survey and one trip reported fewer fish harvested, for a total reporting error of +5.1%. Most of the trips encountered at boat ramps occurred on June 5 and June 6 (Friday and Saturday; Figure 5). Close to half (45.2%) of the total Red Snapper harvest was recorded during these two days. While the app was created with the capability of capturing the trip start and end times, initially it did not capture the date and time when the trip was actually submitted. Unfortunately, it is difficult to know if anglers submitted trips prior to being interviewed. Following the federal season, we did update the app to begin collecting submission date and time, but there were not enough trips validated after the correction to calculate an adjusted error rate. Based on the data we have for 2016 and 2017, it appears as though anglers are submitting their trips after being creeled. This might be due to a lack of knowledge about the app, or the anglers simply not realizing the importance of logging

their trip prior to returning to the launch location. The harvest reported error rates in subsequent years were similar to that of 2015, +5.9% in 2016 and -6.7% in 2017. This is certainly a disadvantage when dealing with user-entered data, and this would be an important issue to overcome through outreach and education in order to ensure trips are entered prior to landing. Based on personal observations by creel teams, we do not suspect anglers are intentionally avoiding or only entering data after creeling. Rather than trying to circumvent the system, it is likely they are simply waiting to enter their catch data until a more convenient time (i.e. upon returning home). We believe they are entering the data at the same point of the trip, independent of whether they are creeled or not. Mandatory reporting could improve the timeliness of their reporting (see detailed discussion below).

Table 5. Creel survey summary of private recreational anglers intercepted during the 2015 federal Red Snapper season (June 1 - June 10) by TPWD and HRI scientists. The harvest reported by all *iSnapper* users is included for ease of comparison between the app and creel surveys. Validated trips were anglers that submitted a trip using the app and were also interviewed at their landing location. Reporting accuracy rate for validated trips was calculated by dividing the app harvest by the creel reported harvest and multiplying by 100.

Day	Trips	Angler-Trips	Creel Harvest	*iSnapper* Harvest	Validated Trips	Reporting Accuracy Rate
6/1/2015 (Mon)	33	107	278	200	2	100%
6/2/2015 (Tue)	13	42	113	195	2	100%
6/3/2015 (Wed)	22	87	206	108	0	-
6/4/2015 (Thu)	24	80	210	91	2	63%
6/5/2015 (Fri)	63	246	581	230	5	104%
6/6/2015 (Sat)	78	307	643	289	6	119%
6/7/2015 (Sun)	11	45	108	199	0	-
6/8/2015 (Mon)	7	33	65	91	1	100%
6/9/2015 (Tue)	5	13	38	70	0	-
6/10/2015 (Wed)	3	9	26	46	0	-
Total	259	969	2268	1519	18	105.1%

Table 5b. Creel survey summary of private recreational anglers intercepted during the 2016 federal Red Snapper season (June 1 – June 11) by TPWD and HRI scientists.

Day	Trips	Angler-Trips	Creel Harvest	*iSnapper* Harvest	Validated Trips	Reporting Accuracy Rate

6/1/2016	5	21	56	36	1	100%
6/2/2016	2	8	12	12	0	-
6/3/2016	20	69	163	64	3	115%
6/4/2016	23	114	228	58	0	-
6/5/2016	46	208	394	92	3	114%
6/6/2016	6	23	56	86	0	-
6/7/2016	5	25	52	80	2	115%
6/8/2016	11	42	84	110	1	100%
6/9/2016	10	39	104	71	1	100%
6/10/2016	31	115	212	31	0	-
6/11/2016	45	195	395	144	1	100%
Total	204	859	1756	784	12	105.9%

Table 5c. Creel survey summary of private recreational anglers intercepted during the 2017 federal Red Snapper season by TPWD and HRI scientists.

Day	Trips	Angler-Trips	Creel Harvest	*iSnapper* Harvest	Validated Trips	Reporting Accuracy Rate
6/1/2017	47	208	374	40	1	66.7%
6/2/2017	66	245	455	106	4	84.6%
6/3/2017	97	446	867	96	3	100%
6/16/2017	0	0	0	6	0	-
6/17/2017	0	0	0	-	0	-
6/18/2017	0	0	0	-	0	-
6/23/2017	0	0	0	-	0	-
6/24/2017	1	4	20	8	0	-
6/25/2017	0	0	0	6	0	-
6/30/2017	0	0	0	-	0	-
7/1/2017	0	0	0	8	0	-
7/2/2017	2	9	12	5	0	-
7/3/2017	3	17	34	26	0	-
7/4/2017	-	-	-	8	0	-
7/7/2017	3	9	17	20	0	-
7/8/2017	23	95	168	107	0	-
7/9/2017	54	198	356	94	2	100%
7/14/2017	3	15	21	44	0	-
7/15/2017	28	102	174	91	0	-

7/16/2017	-	-	-	29	0	-
7/21/2017	8	19	42	42	0	-
7/22/2017	13	65	125	21	1	100%
7/23/2017	1	3	1	-	0	-
7/28/2017	17	63	135	-	0	-
7/29/2017	21	109	222	23	0	-
7/30/2017	24	99	206	26	0	-
8/4/2017	7	32	86	24	1	100%
8/5/2017	11	50	93	-	0	-
8/6/2017	0	0	0	-	0	-
8/11/2017	-	-	-	-	0	-
8/12/2017	25	98	125	8	0	-
8/13/2017	6	28	45	-	1	100%
8/18/2017	5	19	28	16	0	-
8/19/2017	7	30	78	22	0	-
8/20/2017	20	82	165	27	1	100%
8/25/2017	0	0	0	-	0	-
8/26/2017	0	0	0	-	0	-
8/27/2017	0	0	0	-	0	-
9/1/2017	0	0	0	-	0	-
9/2/2017	15	71	201	-	0	-
9/3/2017	6	16	44	-	0	-
9/4/2017	-	-	-	-	0	-
Total	513	2132	4094	903	14	93.3%

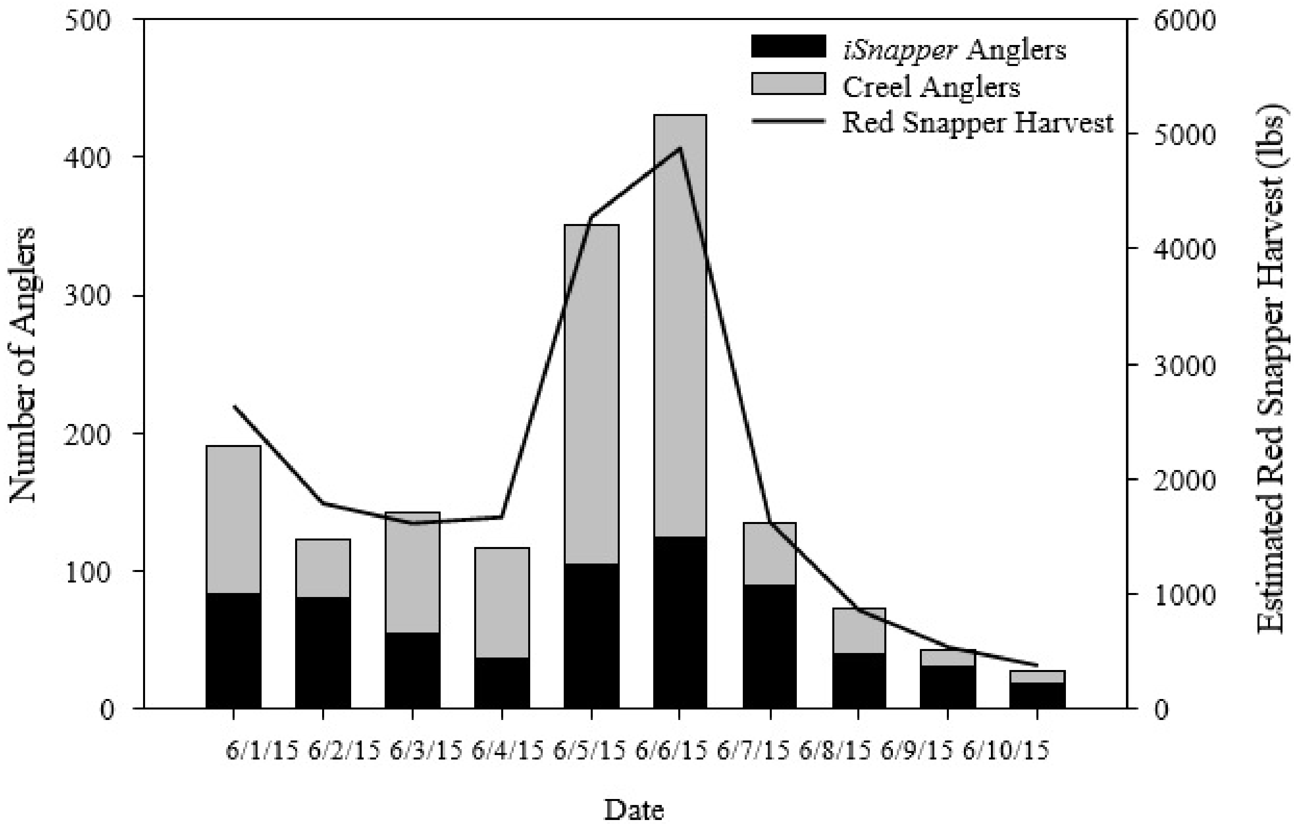

Figure 6. Total number of Red Snapper anglers reporting using *iSnapper* (black bar) or intercepted at boat ramps (gray bar) fishing for Red Snapper during the 2015 federal season. Estimated Red Snapper harvest was calculated by taking the average length of fish recorded by TPWD, converting the length to an estimated weight using the TPWD length/weight conversion chart (http://txmarspecies.tamug.edu/length-weight.cfm) and multiplying that weight by the total number of fish harvested.

After the federal season, anglers were encouraged to continue using *iSnapper* to report their state catch. The TPWD also continued their increased creel surveys to encounter Red Snapper anglers and validate additional trips. From June 11th until the last reported creel with Red Snapper on November 4th an additional 165 boats were interviewed with 1,734 Red Snapper harvested. Only 38 trips harvesting a total of 433 Red Snapper were reported using *iSnapper*, with one boat validated during this time and it was a charter for-hire vessel. The number of Red Snapper harvested was accurate (20), although with the app they reported one less angler than what was recorded by TPWD.

iSnapper total Red Snapper harvest estimate:

Texas is unique in that the harvest of Red Snapper occurs all year in state waters. Using the reporting rate and error estimates from the federal season, along with creel data provided by TPWD for the entire 2015 year, the estimated total number of Red Snapper captured by Texas private recreational anglers was 58,251 fish (SE = 25,344; see Appendix 1 for SE calculations). Using the reporting rates we calculated the total number of Red Snapper angler-trips was

23,358 (SE 6,660). To provide a landings estimate for private recreational anglers, the average size of fish captured from both federal and state waters was calculated and then the weight was estimated using a length/weight conversion website provided by TPWD (http://txmarspecies.tamug.edu/length-weight.cfm). In Federal waters the Red Snapper mean length harvested was 561 mm (22.1 in) and the mean estimated weight harvested was 2.55 kg (5.63 lbs). In State waters the mean size of 513mm (20.2 in), which is approximately 1.93 kg (4.25 lbs). Using the harvest estimates and average weights, the estimated landings of Red Snapper for Texas private anglers in 2015 was 125,986 kg (277,752 lbs, Table 6).

Table 6. Total Red Snapper harvest estimates for 2015-2017 using data from *iSnapper* and TPWD data. Angler-trips is the estimated total number of anglers fishing, which includes anglers fishing for multiple days. Season length varied by year (2015 = 10 days, 2016 = 11 days, 2017 = 3 day initial season, 39 day extended season). 2017 landings are preliminary and we anticipate having an annual estimate prior to June 1st, 2018.

Method	Number Harvested (SE)	Weight (lbs)	Angler-Trips
2017			
iSnapper - Entire 42 day Federal Season	43,992 (15,311)	237,889	24,364
iSnapper - 3 Day Initial Season*	123,121 (61,845)	707,730	63,583
TPWD	TBD	TBD	TBD
2016			
iSnapper - Federal Season	55,062 (31,610)	266,309	27,825
TPWD	18,743 (3,247)	89,415	11,158
2015			
iSnapper	58,251 (25,344)	277,752	23,358
TPWD	32,062 (4,409)	153,125	11,154

**Due to the limited derby nature of the 3-day 2017 season and very limited reporting, we are not confident in this estimate and therefore it should not be used for management advice. It is reported here for discussion purposes.*

Socioeconomic data:

Following the launch of *iSnapper* on May 11th, a total of 100 socioeconomic surveys were completed by 95 unique respondents. Most of these entries occurred after the opening of the federal season, with the most single day entries occurring on June 4th (Figure 6). A total of 98% of respondents were men and 93% of respondents were Texas residents (Figure 7). On average, participants went saltwater fishing 35 days over the past 12 months, with 14 days of those spent offshore. Most (70%) participants also trailered their boats. In terms of household income of Red Snapper anglers, approximately 25% had an annual household income of over $200,000 and approximately 61% had incomes over $100,000 (Figure 8). The average distance these anglers traveled per trip was 89.3 miles (7.09 SE), using an estimated 82 gallons of fuel (12.0 SE). The average cost of bait and tackle for the trip was $197 (34.22 SE).

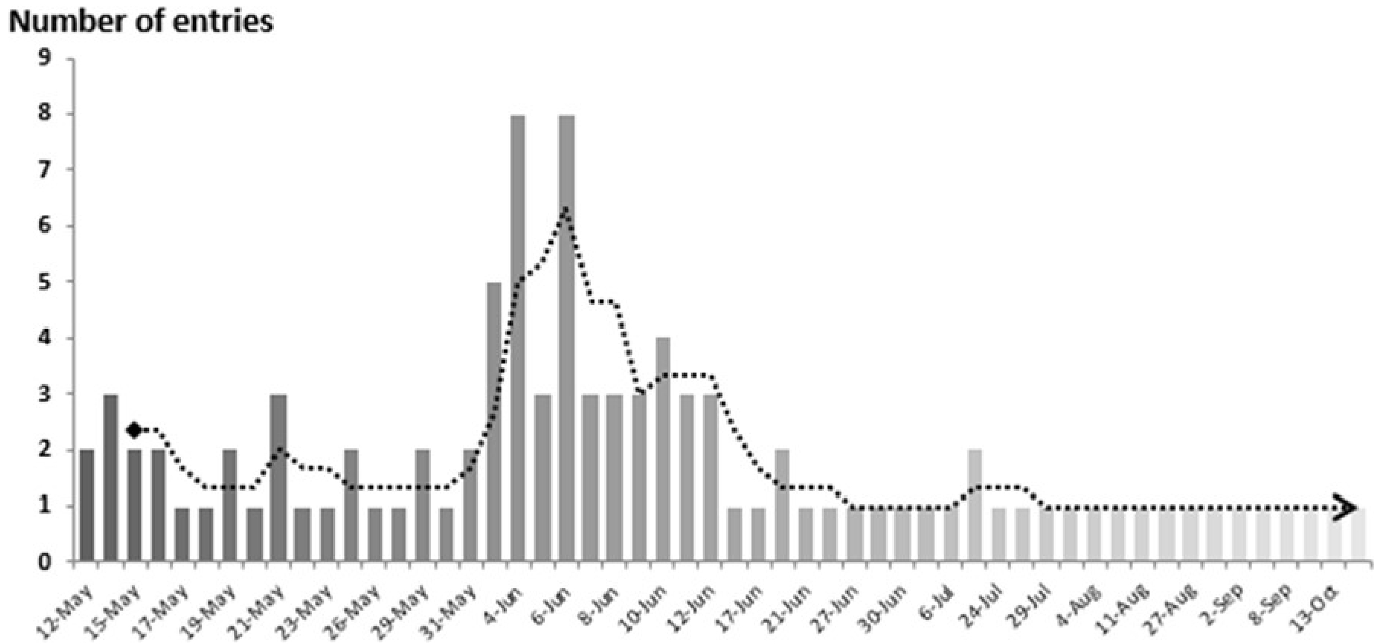

Figure 7. Number of socioeconomic surveys submitted each day following the release of *iSnapper* during the 2015 season. The dotted line indicates the two day lag average, where the number of entries from the previous two days were averaged.

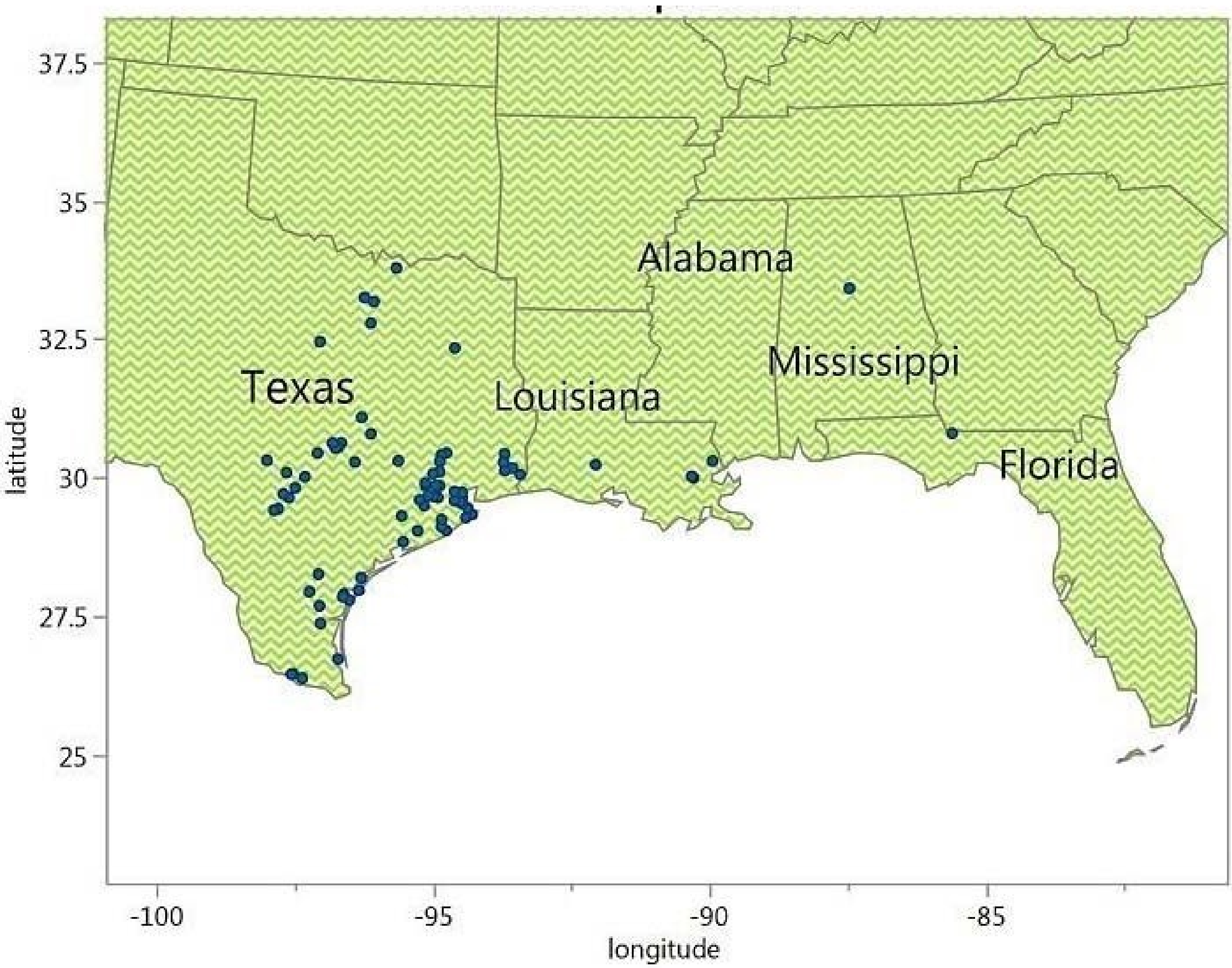

Figure 8. The primary residence locations based on zip code for anglers submitting a socioeconomic survey using *iSnapper* during the 2015 season.

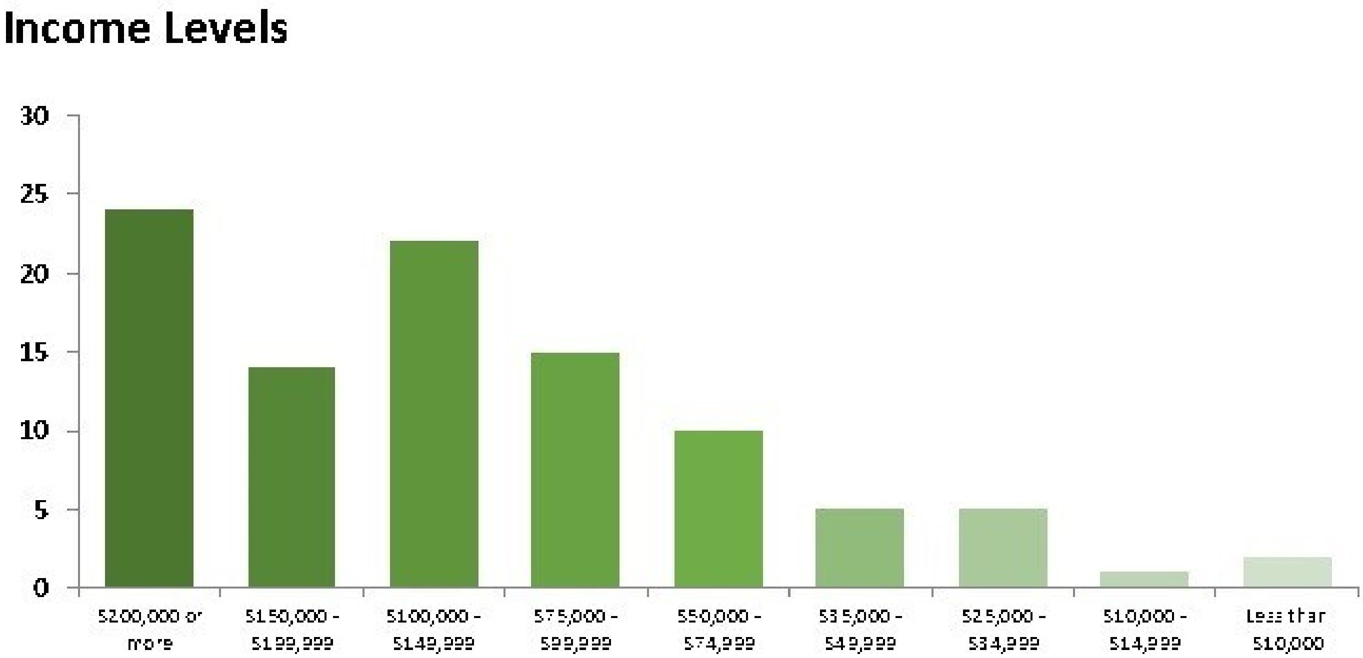

Figure 9. The average annual household income for respondents using *iSnapper*.

<u>*iSnapper* in the for-hire fishery and use in Exempted Fishing Permits:</u>
During development of the initial proposal, there were several Exempted Fishing Permits under consideration in the for-hire sector. *iSnapper* was proposed as an ideal data collection tool for these programs; however, none of these materialized to test the viability in these programs. Nevertheless, it still represented a valuable tool for data collection in the for-hire sectors, and *iSnapper* v1.0 was very successful and popular. In 2015 the for-hire captains had a 44 day season where they could harvest Red Snapper in federally managed waters from June 1st until July 15th. Curiously, there were few trips reported from this sector using this pilot - *iSnapper* v2.0, and this low reporting did not allow for catch estimates to be performed for this sector. Six for-hire captains submitted seven trips during their season and harvested a total of 76 Red Snapper, 25 Dolphinfish, 6 King Mackerel, 1 Warsaw Grouper, and 1 Yellowedge Grouper. Six additional state water trips were submitted by one user who harvested a total of 128 Red Snapper and did not report any other species captured or harvested. In terms of discards, these captains reported no Red Snapper released during the 44 day season in federal waters. However, when fishing in state waters, a total of 45 were released. Despite having so few trips reported, two of the for-hire trips were validated with creel surveys. Including these data with the validations from the private recreational data decreased the overall error rate from 5.1% to 4.1%. Nevertheless, these did not represent enough sample size to make valid harvest estimates for this sector.

6. Discussion/Conclusions/Recommendations

Discussion
This project demonstrates that smart device apps can successfully and efficiently be used as a data collection tool for private recreational anglers. The near-real time data collected can greatly aid in the timeliness of the data generated, which is especially important in the Red Snapper

federal fishery, which was only open for 10 days in 2015. This real-time data collection can allow for fishery managers to make more accurate estimates of the total harvest as well as better determine the fishing effort by allowing anglers launching from private ramps a way to report their catch. However, the utility of creel surveys cannot be overlooked. While an electronic application can provide an avenue for anglers to self-report their data, there still needs to be a high intercept rate so that these self-reported trips can be validated. For *iSnapper*, this involved collaborating with the local fisheries management agency (TPWD) to utilize their sampling effort to estimate certain parameters critical to the calculation of total harvest. Without this partnership, running the app could potentially become cost-prohibitive due to the robust validation rates needed to estimate the total effort and harvest. Overall, this study clearly demonstrates the high potential of using electronic reporting apps to aid in the private recreational fishing data collection and provide traditional and new forms of data in a very efficient and timely manner.

Collecting fisheries-dependent data is certainly key for effective fisheries management, and this process is typically characterized as being labor intensive, expensive, and often has relatively long time delays to produce final products. These inherent characteristics do not always mesh with certain management measures that create very short seasons, or afford the ability for managers to assess catch in relation to annual catch targets during the season, but often occur well-beyond the "projected" season closure. This makes adaptive management difficult and prone to risk if these targets are exceeded before the catch can be calculated, and it also contributes to inefficiencies in management. These issues could not be more apparent than in private recreational fisheries. Many of these problems have the potential to be allayed by supplementing (or even validating) traditional fishery dependent surveys with data from smart device application technology and data collection. For example, these devices capture portions of the fishery that are likely missed by traditional creel surveys (e.g., private dock and marinas) and collect a host of data that traditional surveys do not. These devices also allow for collection of spatial/locational fishing effort, refined effort estimates, discard rates, depth fished, socioeconomic data and a host of other parameters. Much of these data are collected in a seamless user-friendly fashion that is often automated.

The need for more robust, accurate, and timely data from the private recreational sector is making the use of smart devices as data reporting tools more of a reality. These devices allow anglers to self-report their catch and effort data without having to rely on being intercepted and interviewed by state agencies. It also enables anglers that have private docks to report their catch, which previously was not possible and often had to be estimated, since there is not a point of encounter such as at boat ramps. With high private angler reporting rates, managers will be able to more precisely estimate total harvest that likely would benefit the access to the fishery. Knowing how many private anglers need to report each year can be estimated, but greatly depends on what the desired or acceptable standard error rate is, which would need to be set by State and Federal managers. With high precision in the total harvest, the ability for these apps to collect near real-time data can be used to make in-season adjustments to season length based on the current harvest. This is particularly important for a species such as Red Snapper that has such a short season, where traditionally the harvest could not be estimated during the season, because it took months of data entry and analysis that were processed only after fishing had ended. While the user-submitted data is immediately available through the web portal, harvest estimates can only be calculated once the creel data has been entered and trips have been validated. A reasonable time frame to calculate a final harvest for the federal Red Snapper season is three weeks following the close of the season; however, "preliminary" estimates could be made sooner. The three-week time frame is due to the time it takes for TPWD to process creel data for validation. This amount of time allows for the creel data to be entered, trips validated, and then estimating the harvest, which is generally much faster than traditional methods.

With the prevalence of smart phones and the relative low cost of creating an app (Stunz et al., in prep), this type of data collection has a large potential to greatly benefit managers. If designed properly, the apps can provide catch and effort data as well as supplementary information such as depth of capture and release condition at relatively low cost once developed. Depending on what questions fisheries managers seek to understand, the app can be created with additional features and options to address specific concerns about a fishery in a particular region or focus area. For example, in *iSnapper* we provided a section where anglers could describe the release method and condition of fish that were released. With these data, the post-release mortality could be calculated for species in the Gulf that experience significant barotrauma (Lutjanids, Serranids), which could then be incorporated in overall fishing mortality estimates or estimates of depth fished. As a practical example, it has been shown that Red Snapper survival is highest when fish are caught in shallower waters (< 30 m) during the cooler months and are released using either rapid recompression (descender devices) or venting tools (Drumhiller et al. 2014, Curtis et al. 2015). Currently, there are no estimates of what methods private recreational anglers commonly use to release Red Snapper. Since all of those key parameters are collected in *iSnapper* it would be possible to gauge if anglers are releasing fish that have been vented, are descending them using various devices, or are simply discarding them without any barotrauma treatment. This information would help estimate survival rates based on our current knowledge of the likelihood fish survive based on the depth and season captured, which ultimately would provide information for more accurate discard mortality calculations. In addition, since the app is publicly available the potential to collect an expansive data set with relatively little effort would be invaluable to management agencies to better address barotrauma issues and how they impact fisheries. This is just one example of how ancillary data collected by these techniques go beyond traditional catch and effort estimates.

A key component of ensuring self-reported data are valid and of practical use in fisheries management is a strong validation process. Requiring users to provide their vessel registration numbers allowed scientists to compare trips submitted with the app reported data to dockside creel surveys to measure the accuracy of self-reported data. Partnering with TPWD proved to be an efficient way to validate self-reported data, since creel agents are already routinely interviewing boats for their coast-wide biological assessments. Ideally, expanding creeling effort to maximize validations, particularly future directed creel survey in areas of high use, would allow for greater harvest estimate accuracy, improve estimates and reduce variability. For example, the number of boat ramps and marinas that have access to the Gulf is relatively limited in Texas. Out of the 258 boat ramps that are sampled by TPWD Coastal Fisheries, from 2011 – 2014 only 31 ramps recorded Red Snapper harvest, and even fewer (13) recorded Red Snapper landings for more than ten trips. Due to the short federal season, it would be feasible to increase the number of creel surveys done at these highly used ramps to increase the validation rate, making the data more reliable. In addition, with this type of ‘mark and recapture’ method, managers can determine a priori the number of trips needing to be validated to obtain an ideal, or acceptable, standard error for their harvest estimate. Due to the high expense and time cost of creel surveys, knowing the minimum number of trips that must be validated to reach a certain standard error rate would improve efficiency while allowing accurate catch estimate calculations.

One of the unique aspects of this pilot was the ability to collect data from the entire recreational fishing universe and use that information to make harvest and effort estimates. This type of capture-recapture survey methodology allowed us to have a strong data validation component, while also ensuring a randomized sampling design. From the calculated reporting rate and the error estimates, we were able to estimate the total Red Snapper harvest for 2015, as well as the number of angler-trips fishing for this species. Understanding how many private anglers are targeting Red Snapper is critically important, as it has been difficult for TPWD to estimate the total number of Red Snapper anglers due to the high number of Texas residences that have private docks.

We calculated that the 2015 landings from *iSnapper* were 277,752 lbs compared to 153,525 lbs estimated by TPWD. Based on our validation studies, these appear to be reasonable estimates of the private recreational harvest for 2015, but they are also characterized by a high degree of error. This estimate is the first independent assessment from the traditional harvest methodology and represents only one season/data point. Clearly, more data collection across several seasons would need to be performed to increase the confidence in this estimate. Thus, some caution should be taken when evaluating these data for any type of management advice. While the estimates are somewhat larger here, it was also not unexpected, given the app likely captures a segment of the fishery not surveyed using the standard state methodology.

There are several plausible explanations for this difference. The estimate was characterized by a high standard error rate, and the lower bounds of our estimate fall within TPWD estimates. Also, we collected data from anglers that might not have the opportunity to be interviewed by TPWD, particularly people that have a private dock or arrive back to the boat ramps after 6 pm, when the creel assignments end. Thus, *iSnapper* allows these anglers an opportunity to have their catch included in the state harvest estimates. However, it is unclear as to the accuracy and reporting bias of anglers that submitted a trip from private docks, since there was no way to validate these trips. If anglers who return to private docks have a different reporting rate (for the capture period) than anglers who return to public docks, then combining the two data sets might result in errant harvest estimates. Our harvest estimates did combine all the data submitted using *iSnapper*. Currently, we do not know the impact this had on our estimates and the only way to determine the effect is to simultaneously collect data from anglers without requiring an intercept, such as a subsequent telephone or mail survey to a sample of random households. However, this has not been addressed in any of the current versions of *iSnapper*, but could be something to consider in future versions.

Finally, the 10-day federal season was characterized by ideal fishing conditions. Ideal weather led to high fishing effort, and our observations showed at certain times the boats arriving at the dock could easily overwhelm a creeling location leading to an underestimate. Overall, greater user buy-in would translate into additional submitted trips decreasing the standard error and ultimately calculating a more accurate harvest and angler estimate, making the data more useful for fisheries managers.

The method of data validation proved to be valuable in more than just estimating the total harvest of Red Snapper from the Texas private recreational fishery. Because this study was a capture-recapture design, it allows for flexibility and prioritization when determining how best to sample the population of private recreational anglers. For example, if managers seek a specific reporting rate, they would be able to calculate the number of trips that need to be validated during creel surveys. Additionally, to minimize the standard error of a harvest estimate, managers could also determine how many anglers need to report their catch for their desired estimate. Another major benefit is that the actual number of private recreational anglers in the fishery does not need to be known beforehand. These characteristics and benefits of a capture-recapture design show the feasibility of collecting statistically sound self-reported data that can and should be used by state and federal managers.

While the reporting rate was acceptable, and we were able to generate confident estimates, we anticipated more trip submissions. This was especially the case given the perceived interest in the fishery and extensive outreach and advertising campaigns undertaken. We would expect the reporting rate to increase through time as anglers become more aware and familiar with the app. Moreover, given the nature of "recreational" fishing, many anglers may not be willing to go the extra effort to enter data on a voluntary basis, given this is an activity to escape from these sort of tasks. Thus, future efforts should seek to maximize awareness in the private recreational community and simplify and streamline data entry as much as possible to ensure an enjoyable experience and that anglers will want to return and enter trips. Even though we had a very high number of downloads, we attribute some of the low reporting due to the app being voluntary. For example, Alabama has had mandatory Red Snapper reporting since 2014. This program has had a higher reporting rate - approximately 25% of Alabama private recreational vessels reported their catch (Alabama Department of Conservation and Natural Resources 2015). Because 2015 was the first season *iSnapper* was used for private recreational data collection, our reporting rate of 4.1% is encouraging. We believe there will be far greater participation in the future since it appears as though a majority of anglers were simply not aware of the app, and most interviewed seemed very willing to report their catch. This problem could be resolved by building on the current momentum from 2015 and continued advertising to inform anglers about the premise behind *iSnapper* and why private recreational data collection is critical. Nevertheless, we were still able to generate viable estimates and confidence intervals around those catches. Clearly, from these examples voluntary or mandatory reporting still does not ensure the majority of anglers report, and future studies should focus on the trade-offs associated with mandatory versus voluntary reporting and how these different collection frames influence the estimates and associated errors.

Despite our attempts to engage and encourage anglers to submit their trips in 2016 and 2017, it does not appear as though we are building the momentum we thought we would. It is likely that the shorter seasons have left anglers further frustrated and disenfranchised with both state and federal managers to be able to properly manage the species. For example, much of the feedback we received after several years of running the study was that they had been entering data for years, yet subsequent seasons keep getting shorter. After 3 years of doing creel surveys we have heard many complaints from anglers saying they do not want to report due to fears of their numbers "being used against them" to further limit their access to the fishery. This reluctance to report has the potential to be offset by mandatory reporting. This requirement has been implemented in some states with variable success (e.g., see Alabama and Mississippi reporting requirement). In small states with constrained fleets such as Mississippi, it appears to be successful. However, AL has a mandatory reporting requirement, yet they only achieved an estimated 31% in 2016 and 22% in 2017 (AL DNR Report to Gulf of Mexico Fishery Management Council). There are methods to improve this reporting rate, and they are detailed in our summary below. For example, requiring a report before an angler can engage/create the next trip is a solution that has been successfully implemented in MS.

In general, we discovered that using the app requires a behavioral change by private anglers outside of their normal routine. They have to start a trip using the app prior to leaving the dock and then also submit the trip once they return to port. It will take time for anglers to commit and remember to use *iSnapper* for every trip. Additionally, with the Red Snapper fishery being so contentious, there is often mistrust between the anglers and fisheries managers, which at least in discussion with some contributed to their unwillingness to provide fishery-dependent data. As is the case for any data collection endeavor whether it be in person or electronic, some anglers voiced opposition to any type of information transfer such as refusing creels, data entry, etc. Fortunately, the events were rare, and the capture-recapture statistical methodology is robust enough to account for these non-reports. Nevertheless, these findings point toward a need for additional outreach and education directed towards private recreational anglers and how reporting would benefit them, because generally we found most are unaware but willing to help and take what steps are necessary to provide better catch estimates.

iSnapper provides a convenient mechanism to collect socioeconomic data on the users of the resource. This allows managers to conduct equivalent socioeconomic 'monitoring' similar to standard bio-physical measurements instead of ad-hoc processes that are common whether it be commercial (Clay et al., 2014) or recreational (Carter, 2015) estimates. Furthermore, the socioeconomic data is collected from the angler at the time of the activity rather than the individual having to recall expenditures and other activity several months later when contacted for a phone, mail, or online survey. While it is not a stratified random sampling of anglers, the convenient sampling can provide valuable data to identify emerging trends that may require additional study as well as coupling it with traditional survey data. Even with our limited sample size (n=95) this pilot study demonstrated that anglers would voluntarily offer information on expenditures, as well as other data, when not prompted in a formal survey. Assessments of economic impacts and effort in recreational fisheries can be enhanced by knowing the distance of trips taken, consumption of fuel, and expenditures on bait and tackle.

Despite redesigning the app to be more marketable to the private recreational fishing sector, *iSnapper* also collected data with similar data fields for the for-hire sector. However, we were very astonished that only six *iSnapper* users were for-hire charter captains, despite extensive outreach to this community, particularly for Texas. We attribute the low numbers to Amendment 40 passing, which separated the for-hire boats from the private recreational boats into two separate sectors. With this change the for-hire boats were strongly encouraged to adopt a reporting system similar to that of the commercial vessels. There are also ongoing pilot programs using the vessel monitoring system (VMS). Other states were capturing the for-hire industry using their data collection program (e.g., Snapper Check in AL, and LACreel in LA) in their region. This may have led to confusion within the sector as to what reporting mechanism to use. Nevertheless, the current version of *iSnapper* was designed with a charter for-hire component and is readily available for use if desired.

Conclusions and Recommendations for Improvements

Our data collection pilot study demonstrated that smart device “apps” are viable fisheries management tools that effectively collect near real-time fisheries-dependent data from the recreational fishing sector. The app can also be customized to collect important and difficult to obtain data (e.g., discard and location information) that may be used to better estimate parameters such as fishing mortality. While this pilot was specifically targeting Red Snapper anglers, *iSnapper* has the potential to be used to help with management of other species, since all commonly caught species in the Gulf of Mexico are included in the app. The use of this new type of technology as a data collection tool has much potential when recreational anglers see the value in providing their catch data. Due to the short federal Red Snapper season, data collection tools like *iSnapper* in collaboration with federal (i.e., MRIP) and state creel surveys can be used to improve current harvest estimates as well as collection of ancillary fisheries and socio-economic data. In general, we had very positive feedback from anglers, and continuing the momentum created from 2015 would be beneficial to state and federal managers and result in additional data.

After examining the results from this study there are several key items that would improve any future data collection efforts using *iSnapper* and some lessons learned during this pilot. The majority of the issues uncovered related to problems when collecting self-reported data from private anglers; however, many have the ability to be improved:

- Mandatory reporting of all anglers targeting Red Snapper may be a logical next step in obtaining a larger dataset from the app. While mandatory reporting does not guarantee that all anglers will report, there would likely be more trips submitted which would allow for a more robust data set and a reduction in variability in catch estimates, thus helping with the accuracy of the total landings estimates for the season. Analysis of a voluntary versus mandatory program would be beneficial to assess the cost-benefit of each system. No changes would need to be implemented in the app for mandatory reporting. The only change would be to create an account for enforcement to log into so they can view current trips.
- Another unforeseen issue that occurred during this project was that in some circumstances users did not providing accurate vessel registration numbers. While users were required to provide a valid vessel format (ex. TX1234TX), cross referencing the numbers provided with an actual state registry list was not possible during this pilot. Ideally, the registration would be linked with TPWD Boater’s Registration office and users would query the database using their name and address to select their boat during the *iSnapper* registration process. Being able to confirm the vessel registration prior to the angler entering data would be very beneficial as it would improve validation at the marina. In the current study the trips with false registration numbers were excluded from our validated process and contributed to a lower validation rate. Requiring all registered users to find and select their boat would ensure that every *iSnapper* user encountered at the ramps would become a validated trip.
- One of the most difficult aspects of conducting a voluntary study is user involvement. Despite an extensive outreach campaign by both TPWD and HRI, a majority of anglers interviewed were not aware of the app. When anglers were informed about *iSnapper*, some suggested it was unlikely that they were going to download and report their 2015 catch since the federal season was only 10 days and many only fished one or two days of the entire season. These anglers expressed they did not believe their one day of fishing would have any impact on the data or it would be worth reporting, despite understanding the need for better data. This lack of knowledge and limited season were likely motivating factors influencing the number of trips submitted.
- Another challenge was having anglers change their behavior to use a smart device app. As

with any type of behavioral change it takes time to get people actively participating and anglers typically have their own fishing routines. Introducing an extra step of starting a trip on the app before they leave the dock and then submitting their catch prior to getting back to the dock takes additional effort. Not only do users have to be willing to participate, but they also need to remember to use *iSnapper* when typically their phones are stored throughout the trip and not used until after they are back at the dock. Although the app is user friendly and a trip can be filled out in typically less than five minutes, it is still an extra step anglers are not accustomed to doing and have to make a concerted effort to complete.

- The feedback from anglers while encountered at boat ramps was encouraging. Many of them were hopeful that a new way to collect harvest data could be the answer to providing more accurate and robust data which could lend itself to allowing for an increase in the quota. In addition, all anglers that were asked to provide feedback regarding the app were impressed by the features and ease of use, which is promising when considering the potential for *iSnapper* for the 2016 Red Snapper season.

7. References

Alabama Department of Conservation and Natural Resources. 2015. Summary of Alabama's 2015 Mandatory Red Snapper Reporting Program. Available at: http://www.outdooralabama.com/sites/default/files/Preliminary%20Results%20of%20Alabama%27s%202015%20Red%20Snapper%20Reporting.pdf

Baker, S. M. and I. Oeschger. 2009. Description and Initial Evaluation of a Text Message Based Reporting Method for Marine Recreational Anglers. Marine and Coastal Fisheries 1:143–154.

Carter, D. W., S. Crosson, C. Liese. 2015. Nowcasting Intraseasonal Recreational Fishing Harvest with Internet Search Volume. PLoS ONE 10(9): e0137752. doi:10.1371/journal.pone.0137752.

Clay, P. M., A. Kitts, P. P. da Silva. 2014. Measuring the social and economic performance of catch share programs: Definition of metrics and application to the U.S. Northeast Region groundfish fishery. Marine Policy. 44:27-36.

Cohn, J. P. 2008. Citizen Science: Can Volunteers do Real Research? BioScience 48:3, 192–197

Curtis, J. M., M. W. Johnson, S. L. Diamond, and G. W. Stunz. 2015. Quantifying Delayed Mortality from Barotrauma Impairment in Discarded Red Snapper Using Acoustic Telemetry. Marine and Coastal Fisheries 7:1, 434–449.

Drumhiller, K. L., M. W. Johnson, S. L. Diamond, M. R. Robillard, G. W. Stunz. 2014. Venting or Rapid Recompression Increase Survival and Improve Recovery of Red Snapper with Barotrauma. Marine and Coastal Fisheries 6:1, 190–199.

Griffiths, S. P., K. H. Pollack, J. M. Lyle, J. G. Pepperell, M. L. Tonks, and W. Sawynok. 2012. Following the chain to elusive anglers. Fish and Fisheries 11: 220–228

Laplace, P. S. 1786. Sur les Naissances, les Mariages et les Morts Histoire de L'Academic Royale des Sciences, Paris

Papenfuss, J. T., N. Phelps, D. Fulton, and P.A. Venturelli. 2015. Smartphones Reveal Angler Behavior: A Case Study of a Popular Mobile Fishing Application in Alberta, Canada. Fisheries 40:7, 318–327.

Pew Research Center Report. 2015. U.S. Smartphone Use in 2015. Available at: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015

8. Comments

The OT acknowledges that smartphone applications can collect timely recreational fishing data, including data elements not currently collected through ongoing catch and effort surveys such as specific fishing locations, depth fished, discard disposition, release method and socioeconomic data. In this respect, the project team has demonstrated that *iSnapper* is a viable platform for reporting recreational fishing data. However, challenges remain with estimating catch and effort from *iSnapper* data. Specifically, the OT is concerned about 1) angler participation and reporting rates, and 2) the assumption that reporting and validation sampling are independent.

Reporting rates for *iSnapper* decreased from 4.1% in 2015 to 3.5% in 2016 and 2.5% in 2017, despite extensive outreach and a relatively engaged population of anglers. These results suggest that voluntary reporting through smartphone apps may not be especially effective for collecting catch and effort data. As reporting rates decrease, it becomes increasingly difficult to intercept and validate reported trips (i.e. trips reported via *iSnapper* are less likely to be intercepted). Consequently, validation sample sizes may become insufficient to effectively validate reported information and support precision requirements. Alternatively, validation sampling could become cost prohibitive. The present study achieved annual validation sample sizes (validated trips) of 18, 12 and 14, despite supplemental sampling targeting high use red snapper sites. The observed reporting rates and validation sample sizes resulted in estimates with coefficients of variation (relative standard errors) that ranged from 35-57%.

The second concern relates to the observation that “…anglers are submitting their trips after being creeled.” The capture re-capture design requires independence between reporting and validation – if validated anglers are more or less likely to report, then estimates may be biased. The estimation design assumes that the data collection components are independent, but this assumption has not been tested. In fact, it seems likely that the assumption is violated as intercepted anglers “…were informed about *iSnapper*, the value of using it, and were highly encouraged to download and use it…”

The OT recommends approval of the project report. However, we encourage the ESC and NOAA Fisheries to consider the current limitations of voluntary, self reported data collection designs when establishing future program priorities.

9. Appendix

"Estimation of a Total from a Population of Unknown Size and Application to Estimating Recreational Red Snapper Catch in Texas", page 1

Journal of Survey Statistics and Methodology (2017) **5**, 350–371

ESTIMATION OF A TOTAL FROM A POPULATION OF UNKNOWN SIZE AND APPLICATION TO ESTIMATING RECREATIONAL RED SNAPPER CATCH IN TEXAS

BINGCHEN LIU
LYNNE STOKES*
TARA TOPPING
GREG STUNZ

This research was motivated by the desire to more efficiently estimate catch by recreational anglers than current methods do. The method illustrated here combines data from angler self-reports made by a smartphone app and dockside validation samples. The two data sources can be thought of as a capture and recapture, where the parameter of interest is the population total (catch) instead of the population size. We developed several estimators of the total and compared them to one that makes use only of catch observed in the validation sample but not self-reports of catch. All the proposed estimators allow measurement error in the self-reports and do not make any assumptions about their representativeness. The validation sample must be a probability sample for valid inference, and our estimators can accommodate a complex sample design. We provide recommendations about conditions under which one of the estimators discussed may be preferred to another. Finally, we illustrate the method with analysis of data from a pilot project to estimate recreational red snapper catch in the Gulf of Mexico off the coast of Texas.

1. INTRODUCTION

The catch of recreational anglers is more difficult for fisheries managers to estimate than that of commercial operations. The reasons include both wider

Bingchen Liu is an Associate Psychometrician at the Educational Testing Service. Lynne Stokes is Professor, Department of Statistical Sciences, Southern Methodist University. Tara Topping is a Research Specialist at the Harte Research Institute for Gulf of Mexico Studies, Texas A&M University - Corpus Christi. Greg Stunz is an Endowed Chair and Professor at the Harte Research Institute for Gulf of Mexico Studies, Texas A&M University - Corpus Christi
*Address correspondence to Lynne Stokes, E-mail: slstokes@smu.edu

doi: 10.1093/jssam/smx006 Advance access publication 25 July 2017

dispersal and less regulatory requirements for participants. Management of recreational fisheries in the United States is the responsibility of the National Oceanic and Atmospheric Administration (NOAA). NOAA's Marine Recreational Information Program (MRIP) produces estimates of recreational catch by species every two months. MRIP's usual estimator is the product of estimators from two complementary surveys of anglers, one by phone or mail to measure "effort" (number of trips) and one face to face at the dockside to measure mean catch per trip. Improvements to the designs of these surveys have been underway for about a decade, guided by input from a National Research Council report (National Research Council 2006). Recently, however, some stakeholders have proposed using angler self-reports to supplement, or perhaps even to replace, other data sources for estimating catch. Self-reporting has become increasingly simple and inexpensive, with the increasing penetration and ease of use of smartphones. The purpose of this paper is to present methodology that will allow direct use of this self-reported data for estimating recreational catch. The proposed estimators combine data from the nonprobability sample of self-reports with that from the dockside probability sample to produce scientifically defensible estimates of recreational catch. Our analysis shows that if angler self-reporting is sufficiently complete and accurate, these new data can make the effort survey unnecessary. This could save money and allow for more timely production of estimates. Thus our methodology is an example of Groves' prediction: "The combination of designed data with organic data is the ticket to the future" (Groves 2011).

Pilot projects in Alabama, Mississippi, and Texas are testing this approach for estimating red snapper catch. Red snapper is a favorite target of recreational anglers in the Gulf of Mexico, and one of its most economically valuable species. Scientists are concerned that the Deepwater Horizon oil spill has affected the health of this long-lived species (Tarnecki and Patterson III 2015) and slowed its recovery from overfishing. To manage the species properly, accurate and timely estimates of catch are needed. Thus, projects seeking to improve the precision and timeliness of data collection for red snapper have been funded by both NOAA and the Gulf Environmental Benefit Fund (GEBF), which distributes funds for mitigation of damages from the spill. The three state projects all combine their traditional dockside survey data with data from reports that anglers make via computer or via smartphone apps known as Snapper Check, Tails 'n Scales, and iSnapper in Alabama, Mississippi, and Texas, respectively. In all three states, anglers are requested to report their catch using their device before they remove the fish from the boat. In Alabama and Mississippi, reporting is mandatory and enforced by fines if the angler fails to comply, while in Texas it is voluntary. Anglers are also incentivized to provide their data by allowing them to access records of their own angling activities maintained by the system.

Survey researchers recently have shown interest in methods for making statistical inferences using data from nonprobability samples (Baker, Brick, Bates, Battaglia, Couper, et al. 2013). The most common approaches involve weighting, using various methods to produce weights reflecting the number of population members each sample member "represents." One method for doing this is to estimate by modeling a pseudo-probability of selection, which typically relies on knowledge of common auxiliary information in the sample and the population to which inference will be made. The population information can come from summaries of the entire population (such as from Census data) or from a valid probability (reference) sample. Pseudo-probability estimation methods include ones first developed for reducing coverage and nonresponse biases in probability samples, such as calibration (e.g., poststratification) and propensity score adjustments (Lee and Valliant 2009). The bias reduction from these methods depends on availability of auxiliary information that is highly correlated with response variables and on an assumption that the units in the sample are representative of those not present within analyst-defined classes. These methods have shown mixed results (Tourangeau, Conrad, and Couper 2013).

Though our data do include a reference sample containing auxiliary data in common with the nonprobability sample, we take a different approach. Both our probability and nonprobability samples contain measures of catch, but the ones from the nonprobability sample are considered less accurate than those reported by dockside samplers. Thus our proposed estimators are based on the probability sample data and use information about reported catch from the nonprobability sample only as auxiliary data. In fact, all the estimators we propose are either ratio estimators themselves or modifications of ratio estimators. Thus neither correction of the nonprobability sample by weighting nor assumptions about its representativeness are required for validity.

An alternative way to view our data collection and methodology is as a generalization of a Capture-Recapture experiment. This point of view provides an intuitive explanation for fisheries scientists who are typically more familiar with this methodology and the assumptions it requires than with ratio estimation. In a Capture-Recapture experiment, n_1 fish are caught, marked, and released. These fish are not required to be a probability sample of fish. Then a second sample of n_2 fish is captured, and m of them are noted as previously marked. This sample is required to be selected in such a way that all fish (including marked and unmarked) are equally likely to be included. These data can be used to produce an estimator of the unknown population size N by equating the fraction of marked fish in the population and sample:

$$\frac{n_1}{N} = \frac{m}{n_2}.$$

This yields the classical estimator, called the Lincoln-Petersen index, due to the pioneering work of two ecologists (Le Cren 1965):

$$\hat{N} = \frac{n_1 n_2}{m}. \tag{1}$$

Unlike the goal of the usual Capture-Recapture experiment, our goal is to estimate the total t_y of an attribute y (catch) over a population (angler trips), whose size N is unknown. We examine several estimators that generalize the approach of the Lincoln-Petersen index for this purpose. In our application, the sample of self-reports in each of the three red snapper programs plays the role of the capture sample. The dockside sample is viewed as the recapture sample. This validation sample has a complex probability sample design in each of the three programs. The frame of each consists of gulf access points crossed with time blocks, such as four- or six-hour time shifts, so that the sample includes clusters of trips ending at the sampled access point and shift. The clusters are selected into the samples with probabilities that vary, with active fishing times/places more likely to be selected than those with less activity. The catch for all trips sampled is observed and counted by agency personnel. The Lincoln-Petersen Index is based on an assumption that recaptures are a simple random sample of the population, so an adaptation is necessary for valid estimation when the recapture sample design is complex, as ours is. There have been previous applications of the usual Capture-Recapture methodology to estimate N where one or both of the samples are considered to have features of a complex design and weighting is used (Wolter 1986; Alho, Mulry, Wurdeman, and Kim 1993). We adopt a similar approach for our new estimators of total for a population of unknown size.

Pollock, Turner, and Brown (1994) previously considered the problem of estimating a total over a population of unknown size. This application differed from ours in that y was observable only for the units selected in the second sample; that is, there were no angler self-reports of catch for those n_1 trips that were included in the capture sample. They proposed as an estimator of total catch

$$\hat{t}_{yp} = \hat{N}\hat{\bar{y}}, \tag{2}$$

with $\hat{N}$ defined in (1) and $\hat{\bar{y}}$ as the sample average of the ys from the n_2 recaptured units. In our application, the self-reported catch is also available from the capture sample, so it seems reasonable to expect that this information should improve estimation if angler compliance and accuracy are high. Our application also differs from Pollock's scenario in that the validation sample is selected according to a complex design, so that generalizations of expressions (1) and (2) are needed.

In section 2, we introduce three new estimators that use the self-reported data together with that from the validation samples to estimate catch. We develop expressions of the large sample variances of the estimators under a simple random sample (SRS) design for the validation sample. In section 3, we compare these variances with each other and with that of $\hat{t}_{yp}$ under various assumptions about the self-reports. Specifically we consider cases in which reported catch is subject to measurement error, when reporters are not representative but rather are those with extremely high or low catch, and for a range of reporting rates. Section 4 examines estimator performance from simulation studies that are designed to examine the small sample bias of the estimators, as well as the effects of a complex design for the validation sample. One of the simulation settings we use is designed to mimic some features of the data from the Texas iSnapper program. Then we illustrate the method with estimates of red snapper from these data. Discussion follows in section 6.

2. ESTIMATORS OF POPULATION TOTALS USING CAPTURE-RECAPTURE METHODS

Let d_1 denote the subset of the N population units that self-report their trip and catch. d_1 is not assumed to be representative of the population, nor is it a probability sample, but rather is regarded as a domain. Each of the n_1 domain units reports a value for y, but the ith unit's report is denoted by y_i^* to distinguish it from the truth, y_i. No assumptions are made about the relationship between y and y^*. A validation sample s_2 is selected according to a probability design, and y_i is observed for each unit in s_2. A subset of s_2 will match self-reported trips; these units have both y and y^* available. The goal is to estimate $t_y = \sum_{i=1}^{N} y_i$ using the data from d_1 and s_2.

The population and sample data can be visualized as shown in Figure 1. The first row represents the reporting domain d_1 and includes the trips with y^* available, while the second row contains trips without y^*. The first column contains trips in the validation sample, for which y is available; the second column contains the trips without observable y. The upper left cell represents the m matched units with observable y and y^*; the upper right cell represents the $n_1 - m$ reported (y^* known) but unvalidated trips; the lower left cell represents the $n_2 - m$ validated (y known) but unreported trips. The lower right cell contains units with no data available to the analyst.

We denote the reporting rate, defined as the fraction of trips reported by anglers, as

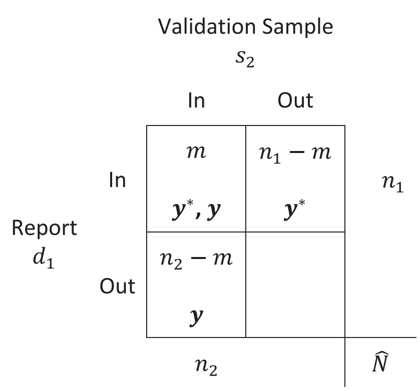

Figure 1. An Illustration of the Population and Sample Data.

$$p_1 = n_1/N, \tag{3}$$

and the population mean as $\bar{y} = t_y/N$. Then Pollock's estimator (2) can be generalized for a complex design to

$$\hat{t}_{yp} = \frac{n_1}{\hat{p}_1}\hat{\bar{y}} = n_1\frac{\hat{t}_y}{\hat{n}_1}, \tag{4}$$

where the $r_i's$ are reporting indicators ($r_i = 1$ if the i^{th} unit is included in d_1; 0 otherwise); the $w_i's$ are sampling weights for units in s_2; $\hat{n}_1 = \sum_{i \in s_2} w_i r_i$; $\hat{p}_1 = (\sum_{i \in s_2} w_i r_i / \sum_{i \in s_2} w_i)$; and $\hat{\bar{y}} = (\sum_{i \in s_2} w_i y_i / \sum_{i \in s_2} w_i)$. Thus $\hat{t}_{yp}$ is recognizable as a ratio estimator with auxiliary variable r_i and ratio

$$B_p = t_y/n_1.$$

Now we propose a new estimator that is also a ratio estimator, but with auxiliary variable $r_i y_i^*$ and ratio denoted by

$$B_c = t_y/\sum_{i=1}^{N} r_i y_i^* = t_y/t_{y^*},$$

where $t_{y*} = \sum_{i \in d_1} y_i^* = \sum_{i=1}^{N} r_i y_i^*$ is the total reported catch. This yields the estimator

$$\hat{t}_{yc} = t_{y*}\frac{\sum_{i \in s_2} w_i y_i}{\sum_{i \in s_2} w_i r_i y_i^*} = t_{y^*}\frac{\hat{t}_y}{\hat{t}_{y^*}}. \tag{5}$$

This estimator can be regarded as the total reported catch inflated by the estimated reporting rate $(\sum_{i \in s_2} w_i r_i y_i / \sum_{i \in s_2} w_i y_i)$, and adjusted for reporting errors by a multiplicative correction factor $(\sum_{i \in s_2} w_i r_i y_i / \sum_{i \in s_2} w_i r_i y_i^*)$. It is a generalization of the Capture-Recapture estimator, where totals of y and $y*$ replace counts of units in the two data collection periods.

If $y*$ is accurate and the reporting rate is high, $\hat{t}_{yc}$ would be expected to be more precise than $\hat{t}_{yp}$ due to its highly correlated auxiliary information. If reporting is inaccurate and rare, the reverse would be true. To avoid making the choice of estimator in advance, one could form a linear combination of the two estimators, with weights chosen to minimize variance. This estimator is a special case of what Olkin (1958) called the multivariate ratio estimator, which we denote as

$$\hat{t}_{MR} = (1 - W)\hat{t}_{yp} + W\hat{t}_{yc}. \tag{6}$$

Olkin (1958) showed (eq. 3.1, p. 157) that the optimal weight W for a simple random sample (SRS) can be approximated to order $O(1/n)$ by

$$w_{SRS} = \frac{S_{dp}^2 - S_{dp,dc}}{S_{dp}^2 + S_{dc}^2 - 2S_{dp,dc}}, \tag{7}$$

where S_{dp}^2, S_{dc}^2, and $S_{dp,dc}$ denote the variances and covariance of the residuals from the ratio models.

In our application, these residuals are $d_{pi} = y_i - B_p r_i$ and $d_{ci} = y_i - B_c r_i y_i^*$, and their variances and covariances can be expressed as shown in the Appendix in (A.2), (A.3), and (A.4). Using these expressions, (7) simplifies to

$$w_{SRS} = \frac{t_{y^*}}{t_y} \frac{S_{1,yy^*}}{S_{1y^*}^2} = \frac{t_{y^*}}{t_y} \frac{S_{1y}}{S_{1y^*}} R_{1,yy^*}, \tag{8}$$

where R_{1,yy^*}, S_{1,yy^*}, S_{1y}, and S_{1y^*} are the correlation, covariance, and standard deviations of y and y^* in the reporting domain d_1. Thus the optimal estimator gives $\hat{t}_{yc}$ the majority of weight ($w_{SRS} > 1/2$) when

$$R_{1,yy^*} > \frac{CV_{1y}}{2p_1 CV_{1y^*}},$$

where p_1 is the reporting rate (from (3)), and CV_{1y} and CV_{1y^*} are the coefficients of variation of y and y^* in the reporting domain.

In practice, w_{SRS} must be estimated in order to use $\hat{t}_{MR}$. We consider two estimators. For the first, we replace the components of (8) with estimators calculated from the observed data, as suggested in Olkin (1958). One such estimator is

$$\hat{w}_{SRS,1} = \frac{t_{y^*}}{\hat{t}_{yc}} \frac{s_{1,yy^*}}{s_{1y^*}^2},$$

where $s_{1y^*}^2$ and s_{1,yy^*} are the estimated variance and covariance between y and y^* in the reporting domain, made from the matched sample. (Alternatively, one could use $\hat{t}_{yp}$ or implicitly define an estimator by substituting $\hat{t}_{MR}$ for t_y in (6), or use the observable value of $S_{1y^*}^2$ in the denominator of $\hat{w}_{SRS,1}$. Simulation showed little difference in performance among these alternatives.) We denote the resulting estimator by $\hat{t}_{y1}$. The second estimator we consider is simpler and near optimal when reporting errors are small. Note from (8) that when $y = y^*$, $w_{SRS} = t_{y^*}/t_y$. Thus we estimate w_{SRS} by

$$\hat{w}_{SRS,2} = \frac{t_{y^*}}{\hat{t}_{yc}}.$$

The resulting estimator can be simplified to

$$\hat{t}_{y2} = t_{y^*} + \frac{n_1}{\hat{n}_1}\left(\hat{t}_y - \hat{t}_{y^*}\right) = t_{y^*} + n_1 \hat{\bar{\delta}}, \tag{9}$$

where $\delta_i = y_i - r_i y_i^*$ and $\bar{\delta} = (t_y - t_{y^*})/n_1$ is the total population underreport averaged over reporters. In contrast to $\hat{t}_{yc}$, this estimator augments the reported catch by an additive rather than a multiplicative component.

When the validation sample has a complex design, it can be accounted for in $\hat{t}_{yp}$ and $\hat{t}_{yc}$ as shown in (4) and (5), and these estimators combined as in (6). Olkin (1958) generalized (7) to produce an appropriate expression for W when the sample has a stratified design. For a general complex design, however, it is useful to note that the optimal value for W is

$$W = \frac{V(\hat{t}_{yp}) - Cov(\hat{t}_{yp}, \hat{t}_{yc})}{V(\hat{t}_{yp}) + V(\hat{t}_{yc}) - 2Cov(\hat{t}_{yp}, \hat{t}_{yc})}, \tag{10}$$

which reduces to Olkin's expressions for SRS and stratified designs when Taylor series variance approximations are used. Modern survey software can provide estimates of the variances and covariance for most designs, so that explicit expressions are not needed for each design type. The optimal W will not be well approximated by (7) if design effects for the two estimators differ greatly. Then the simplified forms shown in $\hat{w}_{SRS,1}$ and $\hat{w}_{SRS,2}$ and the resulting estimators will no longer be nearly optimal.

In section 3, we focus only on SRS designs. We compare the variances of the estimators to help us understand their relative performance as the completeness, representativeness, and accuracy of reporting change. The goal is to provide guidance on which estimator may be best for different applications. In section 4, we extend the comparison to complex designs via simulation.

3. COMPARISON OF ESTIMATORS FOR SIMPLE RANDOM SAMPLES

We compare the approximate variances of $\hat{t}_{yp}$, $\hat{t}_{yc}$, and $\hat{t}_{y2}$, to that of $\hat{t}_{MR}$ under an SRS design for the validation sample and for a range of scenarios for the quality and completeness of the self-reported catch information. We do not consider $\hat{t}_{y1}$ separately since its large sample behavior is that of $\hat{t}_{MR}$. The ratios of the variance expressions in (A.9) through (A.12) are unaffected by the size of the validation sample n_2, the population size N, or the total itself, t_y. They do depend on the reporting rate p_1, the correlation and CV's of y and y^* in the reporting domain (R_{1,yy^*}, CV_{1y}, and CV_{1y^*}), and the ratios of the means of y and y^* in the reporting domain to y's mean in the population ($\bar{y}_1/\bar{y}$ and $\bar{y}_1^*/\bar{y}$). Therefore, we present comparisons of the variances of the three estimators to that of $\hat{t}_{MR}$ for the following three scenarios: (1) no errors in reporting and reporters are representative of the population; (2) errors in reporting, but reporters are representative of the population; and (3) no errors in reporting, but reporters are not representative of the population. We examine the loss of precision for $\hat{t}_{yp}$, $\hat{t}_{yc}$, and $\hat{t}_{y2}$, as compared to $\hat{t}_{MR}$.

In scenario 1, we assume that $y = y^*$ ($R_{1,yy^*} = 1$, $\bar{y}_1 = \bar{y}_1^*$, and $CV_{1y} = CV_{1y^*}$) and reporters are representative of the population (defined to mean that $\bar{y}_1 = \bar{y}$ and $CV_{1y} = CV_y$). This will occur on average if reporting is "at random," though randomness is not required for representativeness. Figure 2 displays the ratio of the large sample variance of $\hat{t}_{MR}$ to that of each estimator as a function of reporting rate p_1, defined in (3). Panels (A) and (B) show how CV_y (set to 0.32 and 0.55) affects performance. When $CV_y = 0$, $\hat{t}_{yp}$ and $\hat{t}_{yc}$ are equivalent. When $CV_y > 0$, $\hat{t}_{yp}$ is more efficient than $\hat{t}_{yc}$ for small reporting rate but grows less efficient as p_1 increases. The crossover point occurs at $p_1 = 1/2$ regardless of CV_y, but the advantage for $\hat{t}_{yc}$ grows with CV_y. $\hat{t}_{y2}$ is uniformly optimal in this case since $y = y^*$.

Next we examine the performance of the estimators when self-reports are not accurate ($y \neq y^*$), but reporters are representative ($\bar{y}_1 = \bar{y}$ and $CV_{1y} = CV_y$). $\hat{t}_{yp}$ is unaffected by measurement error since it does not use y^*. (A.11) shows that errors increase the variance of $\hat{t}_{MR}$ by decreasing R_{1,yy^*}, while (A.10) shows that they affect the performance of $\hat{t}_{yc}$ through both R_{1,yy^*} and CV_{1y^*}. Since CV_{1y^*} can either increase or decrease when $y \neq y^*$, the effect of measurement error on the variance of $\hat{t}_{yc}$ is not clear. Finally, (A.12) shows that the variance of $\hat{t}_{y2}$ is affected by errors through R_{1,yy^*}, CV_{1y^*}, and $\bar{y}_1^*/\bar{y}$. Thus we compared the estimators under two measurement error models that impact these parameters differently: the classical measurement error (CME) (Carroll et al. 2006, section 1.2) and the Berkson model (Berkson 1950).

The CME model specifies

$$y^* = y + e, \tag{11}$$

where $e \sim (0, \alpha S_y^2)$, $S_y^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2/(N-1)$ is the variance of y in the finite population, y and e independent. When (11) holds, $R_{1,yy^*} = 1/\sqrt{1+\alpha}$, $CV_{1y^*} = CV_{1y}\sqrt{1+\alpha}$, and $\bar{y}_1^*/\bar{y} = \bar{y}_1/\bar{y}$. The Berkson model reverses the role of y and y^* and specifies

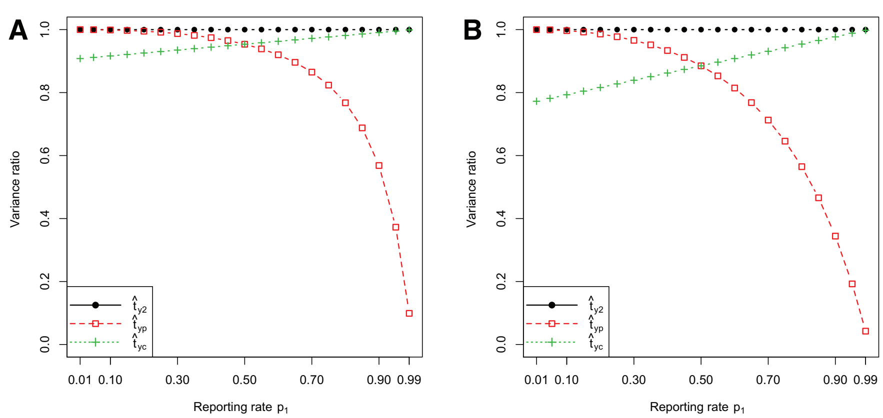

Figure 2. Relative Precision of Three Estimators to $\hat{t}_{MR}$ as a Function of p_1, When There are No Errors, and Representative Reporting. (A) Variance ratios when CV_y=0.32. (B) Variance ratios when CV_y=0.55.

$$y = y^* + e, \tag{12}$$

with $e \sim (0, \beta S^2_{y^*})$, y^* and e independent. A Berkson error model is plausible in our application if the trip reporter provides the catch value as the bag limit (maximum legal catch) * # of anglers aboard as an estimation strategy. Under (12), the domain parameters would be $R_{1,yy^*} = 1/\sqrt{1+\beta}$ and $\bar{y}^*_1/\bar{y} = \bar{y}_1/\bar{y}$. Berkson error causes CV_{1y^*} to decrease; $CV_{1y^*} = CV_{1y}/\sqrt{1+\beta}$.

Figure 3 shows the variance ratios of $\hat{t}_{MR}$ to $\hat{t}_{yp}$, $\hat{t}_{yc}$, and $\hat{t}_{y2}$ as functions of R_{1,yy^*}, where $CV_y = 0.32$ and $p_1 = 0.7$. The two panels show how their performance changes when the measurement error structure differs; CME is assumed for panel (A) and Berkson error for panel (B). Recall from Figure 2A that $\hat{t}_{yc}$ and $\hat{t}_{y2}$ outperform $\hat{t}_{yp}$ for these settings when $y = y$*. Figure 3A shows that this advantage is lost when CME afflicts y*, unless the correlation between y and y* is substantial (about 0.7 for $\hat{t}_{y2}$ and 0.84 for $\hat{t}_{yc}$). When y* has Berkson error, the relative performance returns to its no-error order. Berkson error does not increase the variance of $\hat{t}_{yc}$ and $\hat{t}_{y2}$ as much as CME does for the same correlation, while both models affect the variance of $\hat{t}_{MR}$ equally. In fact, the large sample variance of $\hat{t}_{y2}$ is identical to that of the optimal estimator in this case, even though errors occur. The point here is that the structure of the measurement error matters for determining which estimator is best, and the preference depends on more than R_{1,yy^*}.

Finally, we again assume that $y = y$*, but reporters are not representative. Then the mean and/or variance of y are different for reporters than for the population, affecting estimator variances through $\bar{y}_1/\bar{y}$ and CV_{1y}. To see how much, we must specify a mechanism for determining who reports. We examined two extremes: that reporters have the largest or smallest catch; that is, reporters are assumed to be those in the top (bottom) $100p_1\%$ of y's

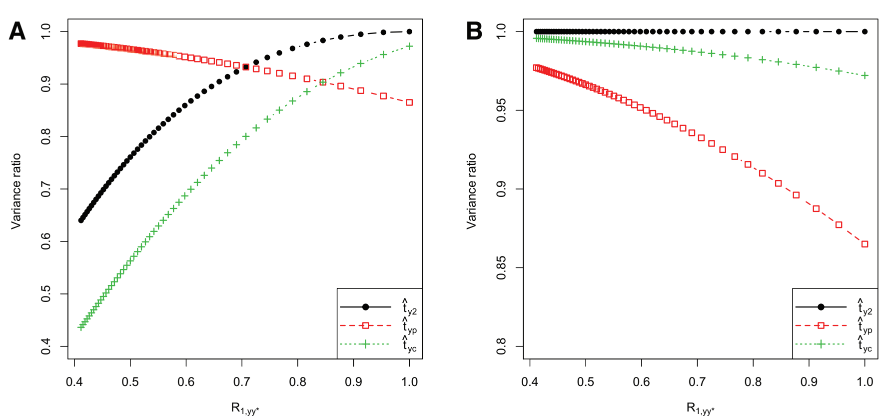

Figure 3. Ratio of Variance of $\hat{t}_{MR}$ to the Three Estimators as a Function of R_{1yy^*}, When p_1=0.7, and Reporters are Representative. (A) Classical measurement error model for y*. (B) Berkson measurement model for y*.

distribution. The effect of these mechanisms on $\bar{y}_1/\bar{y}$ and CV_{1y} depends on the distribution of y.

We considered two distributions for y, one continuous (zero-truncated normal) and one discrete (zero-truncated Poisson). When y is normal, the distribution of y in the high catch reporting domain is that of an upper tail truncated normal, with truncation point $A = \bar{y} + S_y\Phi^{-1}((1-p_1)(1-\Phi(-\bar{y}/S_y))$, where Φ is the standard normal CDF. The low catch domain was defined similarly. Thus the moments of y in the reporting domain are easily calculated (e.g., Johnson and Kotz 1970, pp. 81–3). When y is zero-truncated Poisson, its distribution in the reporting domain is also truncated Poisson, but at a value larger than 0. The moments of the k-truncated Poisson are also easily calculated (Johnson and Kotz 1970). Because of the discreteness of this distribution, only some values of p_1 are possible for this model.

Figure 4 shows the variance ratios as functions of p_1 when the domain contains high removal reporters, where panel (A) shows results for truncated normal y (with $CV_y = 0.32$) and panel (B) for truncated Poisson y (with $\lambda = 1.79$, which yields $CV_y \approx 0.55$). Thus the differences in Figure 2 and Figure 4 illustrate the impact of nonrepresentative reporting only. A comparison shows that high-removal nonrepresentative reporting improves the relative performance of $\hat{t}_{yc}$, especially for small p_1. $\hat{t}_{yp}$ alone declines in performance compared to the best estimator as the reporting rate increases.

Table 1 is designed to show how much nonrepresentative reporting affects the absolute and not just the relative variance of the estimators under largest and smallest catch reporting. It displays the ratio of each estimator's variance when reporters are representative and when they are not (larger for the upper and smaller for the lower half of the table) for normal y and two reporting rates:

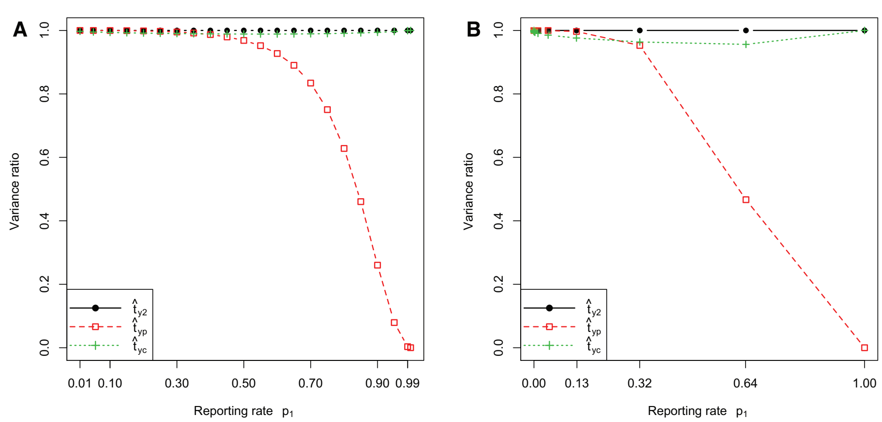

Figure 4. Ratio of Variance of $\hat{t}_{MR}$ to the 3 Estimators as a Function of p_1, When There are No Errors, and Max Catch Reporting. (A) For zero-truncated Normal distribution with CV_y=0.32. (B) For zero-truncated Poisson distribution with CV_y=0.55.

Table 1. Ratios of Variances When Reporters are Representative and Not Representative when $CV_y = .32$

	p_1	$\hat{t}_{yp}$	$\hat{t}_{yc}$	$\hat{t}_{y2}$
$\frac{Var(\hat{t}\mid\text{representative reporting})}{Var(\hat{t}\mid\text{large removal reporting})}$	0.15	1.21	1.30	1.21
	0.70	2.47	2.60	2.56
$\frac{Var(\hat{t}\mid\text{representative reporting})}{Var(\hat{t}\mid\text{small removal reporting})}$	0.15	0.85	0.87	0.85
	0.70	0.63	0.57	0.57

$p_1 = 0.15$ and $p_1 = 0.70$. The table shows that the estimators improve substantially if reporters are those with high values of y, and are damaged if only low y units report. This is easy to explain intuitively for $\hat{t}_{yc}$ and $\hat{t}_{y2}$, since they use the reported y values directly, so when large values of y are reported, less uncertainty remains for the unseen domain. It is less obvious for $\hat{t}_{yp}$, since no reported values of y are used in the estimator. The explanation is that when reporters ($r_i = 1$) have large catch the correlation between r_i and y_i in the population increases, reducing the variance of the ratio estimator $\hat{t}_{yp}$. This shows that reporting should be encouraged, especially for the most avid anglers.

4. SIMULATION STUDIES TO EXAMINE INFERENCE FOR POPULATION TOTALS

The bias in $\hat{N}$ and its standard error estimate are known to be substantial when the number of matches between the two capture periods is small. Since all the proposed estimators are related to $\hat{N}$, we wanted to examine their bias and confidence interval coverage, especially for low reporting rates. We were also interested in the performance of the estimators for complex designs as these are common in dockside samples used for validation. Therefore, we conducted simulation studies to examine these issues.

In the first study, we investigated the performance of the estimators for SRS designs under several settings for (n_2, p_1), some of which resulted in a small number of matches. For each estimator, a 95 percent normal theory-based confidence interval was calculated from each simulated sample and its coverage noted. The standard errors were calculated with off-the-shelf survey software (R's Survey package) (Lumley 2004), using the Taylor series-based variance estimates for ratio estimators. The variance estimator for $\hat{t}_{y1}$ was that proposed in Olkin (1958), which ignores the variability in the estimated W, that is, we calculated an estimate of $V(\hat{t}_{MR})$ (in A.11). The details of the simulation study and results mentioned in Section 4 are contained in the supplementary materials. The simulation study was conducted to investigate the performance for the proposed estimators under several settings for (n2, p1) when the validation sample is SRS. Both the report has no errors and errors were considered in the simulation study. In general the results showed that confidence interval coverage was near nominal (coverage rate estimates between 93.3 percent and 95.3 percent) for all but one of the 12 settings considered. The one exceptional setting was $(n_2, p_1) = (800, 0.99)$ in a population that showed no errors in reporting. For that case, the coverage of the confidence intervals based on $\hat{t}_{yc}$ and $\hat{t}_{y2}$ was about 90 percent. So the problem with coverage was not caused by the low number of matches in this case.

A second simulation was designed to study performance of the estimators for complex designs. The population and sample designs tested were chosen to mimic some of the features of the data from the Texas iSnapper project. We created the finite population structure by replicating each primary sampling unit (PSU) in the Texas validation sample a number of times that was proportional to its weight to obtain a population of 20,590 trips. The average number of trips per PSU was 12. Then we simulated the "catch" data y for each trip from a zero-truncated Poisson distribution with mean parameter 10. The simulated population total is 205,583. We examined estimators for two forms of "reported" data. For the first, we assumed perfect reporting $(y = y^*)$. For the second, erroneous reports were constructed by first computing $y^* = y + \epsilon$, where ϵ was simulated as a mean 0 normal random variable, and then y^* rounded to an integer (or to 0 if negative). The variance of the normal random variable was set (by trial and error) so that the correlation between y and y^* in the reporting domain was 0.66. In both cases, the reporting units were simulated to be nonrepresentative, by selecting them randomly from among the units in the largest 70 percent of the y values.

The validation sample was chosen according to a stratified cluster design with PSUs selected with probability proportional to size, where the size measures were those associated with the PSUs in our application data. The strata (weekday and weekend time periods) were defined as in the original data. The fraction of PSUs in the sample that were chosen from the two strata (0.56 from weekday, 0.44 from weekend) match the Texas dockside sample design. Two levels for the number of PSUs sampled (27 and 60) and the reporting rate (0.04, 0.80) were selected for the simulations, and estimates were calculated based on both the perfect and erroneous reports. Sampling was replicated 30,000 times. Then $\hat{t}_{yp}$ and $\hat{t}_{yc}$ were calculated for each sample, along with two hybrid estimators. The first was the complex sample analog of $\hat{t}_{y1}$ which takes the form of $\hat{t}_{MR}$, but with W estimated from (10). The second estimator, which we denote by $\hat{t}_{y2}$, was computed by simply substituting weighted estimates $\hat{n}_1$, $\hat{t}_y, \hat{t}_{y^*}$ in (9). This estimator is not necessarily optimal even if there are no reporting errors since the design effects for the two estimators may differ, but it is still approximately unbiased and is simple to compute. For each simulation setting and replicated sample, the estimator variances were estimated using both the

Taylor series and the jackknife standard error options in *R*'s *Survey* package. Ninety-five percent normal theory-based confidence intervals were computed.

A summary of the results is shown in Table 2. For each variance estimation method/estimator/setting, three statistics describing the results of the 30,000 replicates were computed. First is the proportion of confidence intervals that covered t_y, which is reported in the column labeled "Coverage." Next is the average standard error of the estimate, reported in the column labeled "SE." The variance of each estimator was also computed over the replicates of the simulation, as was the average of its replicate variance estimates. The relative bias in the variance estimate was computed as the difference between the estimated and simulated variance divided by the simulated variance. It is reported as a percentage in the table as "RelBias." A negative relative bias means that the variance estimator is biased downward.

The results show that the Taylor series variance estimates do underestimate the true variance for all estimators when the number of PSUs is small ($n_2 = 27$), resulting in confidence interval coverage that is less than nominal. The jackknife estimate of variance performs better and provides closer-to-nominal coverage of the confidence intervals. $\hat{t}_{y1}$ has especially low coverage when n_2 is small because it is slightly biased. As predicted from the SRS analysis of section 3, $\hat{t}_{yp}$ outperforms $\hat{t}_{yc}$ for the small reporting rate, and the reverse is true for the large reporting rate. The presence of reporting errors does degrade the precision of all the estimators except $\hat{t}_{yp}$ (which does not use y^*), but $\hat{t}_{yc}$ still maintains its advantage. The hybrid estimators show mixed results. When the number of matches is very small (small p_1 and n_2), $\hat{t}_{y1}$ does not perform well and $\hat{t}_{y2}$ is virtually identical to $\hat{t}_{yp}$ as expected. When the number of matches is large (large p_1 and n_2), they outperform both $\hat{t}_{yp}$ and $\hat{t}_{yc}$.

5. EXAMPLE

Red snapper is one of the most highly targeted species in the northern Gulf of Mexico. Since the 1980s, this fishery has been listed as overfished, and drastic reductions in both season and bag limits have been implemented to help this population fully recover. Despite the stock being classified as overfished, the population is recovering rapidly, and anglers are seeing more red snapper than in previous years. Anecdotal and stock assessment reports both indicate higher abundances of snapper, but these reports continue to be met with reductions in federal season length, which was only 10 days long in 2015. This unique enigma has led to heated conflicts regarding allocation among user groups in the fishery. Much of the concern could be allayed with better data, specifically addressing the uncertainty of recreational catch estimates. With several states trying new management strategies, it is an ideal species to test the feasibility of smartphone "app" technology for private recreational anglers to report their catch. Texas was one of three states with a pilot program using the new technology for estimating

Table 2. Coverage Rate, Standard Error of the Estimate, and Bias of Variance Estimate for Each Estimator Based on 30,000 Replicates

		No errors in report									Errors are present in report with $R_{1,yy^*} = 0.66$					
		$p_1 = 0.04, n_2 = 27$			$p_1 = 0.04, n_2 = 60$			$p_1 = 0.80, n_2 = 60$			$p_1 = 0.04, n_2 = 27$			$p_1 = 0.80, n_2 = 60$		
		Coverage	SE	RelBias	Coverage	SE	RelBias	Coverage	SE	RelBias	Coverage	SE	RelBias	Coverage	SE	RelBias
$\hat{t}_{yp}$	Taylor	0.936	59,554	−0.17	0.944	36,853	−0.05	0.948	3,307	0.00	0.936	59,554	−0.17	0.948	3,307	0.00
	Jackknife	0.948	70,336	−0.02	0.952	39,023	0.00	0.951	3,355	0.01	0.948	70,336	−0.02	0.951	3,355	0.01
$\hat{t}_{yc}$	Taylor	0.935	61,922	−0.19	0.946	37,562	−0.06	0.945	2,806	0.00	0.934	64,071	−0.19	0.945	3,526	0.00
	Jackknife	0.948	74,666	−0.02	0.955	39,913	0.00	0.948	2,841	0.01	0.948	78,313	−0.01	0.947	3,571	0.02
$\hat{t}_{y1}$	Taylor	0.885	54,606	−0.15	0.924	35,696	−0.07	0.940	2,748	−0.02	0.885	54,660	−0.17	0.940	3,029	−0.02
	Jackknife	0.914	72,361	0.12	0.935	38,026	0.00	0.943	2,783	−0.01	0.910	69,456	0.06	0.938	3,072	−0.01
$\hat{t}_{y2}$	Taylor	0.936	59,532	−0.17	0.947	36,844	−0.05	0.945	2,795	0.00	0.936	59,532	−0.17	0.945	3,346	0.00
	Jackknife	0.948	70,336	−0.02	0.952	38,981	0.00	0.948	2,829	0.01	0.948	70,378	−0.02	0.948	3,394	0.02

red snapper catch in 2015. The Harte Research Institute (HRI) created iSnapper to collect catch and effort data from recreational anglers targeting red snapper in state and federal waters off the Texas coast. Here we apply the methods of this paper to estimate the total catch of red snapper in 2015 in Texas by recreational anglers in private boats using data from iSnapper.

Unlike Alabama and Mississippi, self-reporting of catch in Texas was voluntary in 2015. To validate the self-reports, HRI partnered with the Texas Parks and Wildlife Department (TPWD), which routinely samples anglers in dockside surveys using a probability sample of locations and time blocks to produce the mandated estimates of fishing effort by recreational anglers for NOAA. (See NOAA's description of its survey methods at www.st.nmfs.noaa.gov/recreational-fisheries/index.) The time blocks are stratified by weekday and weekend, while the locations have unequal selection probabilities that are proportional to a "pressure" measure, which is meant to capture the average number of anglers using a particular site in past years. To augment the TPWD sample, HRI also sampled in targeted high use marinas and boat ramps during the first six days of the federal red snapper recreational season (June 1–10, 2015). These PSUs were treated as take-all strata for estimation (i.e., given weights of 1). Catch counts were collected from every vessel intercepted during sampled shifts. Vessel registration numbers were recorded and used, along with day and time, to identify matches to trips submitted using the iSnapper app.

The number of intercepted trips in the validation sample was 421, which were clustered in 27 PSUs, with 15 and 12 in the weekday and weekend strata, respectively. The proportion of the trips previously self-reported was estimated to be only $\hat{p}_1 = 0.04$. The estimates of mean catch made from the validation sample for the population and for the self-reporters were 9 and 10, respectively. Thus the self-reporters are not representative, but rather have larger-than-average catch. The CV of catch was estimated to be 0.68. The design effect for the estimate of mean catch from the validation sample alone was about 1.4. The accuracy of reporting was high when measured as a total, with only about a 3.8 percent higher catch reported than observed in the matched sample. However, the correlation between y and y^* was only about 0.66 due to the fact that the erroneous self-reports were small in number but tended to be high outliers.

The estimates of catch from the four estimators, computed as described for the complex design in the previous section, are shown in Table 3. (For context,

Table 3. Estimated Total Landings of Red Snappers in 2015 Using Four Different Estimators and Two Standard Error Estimates

	$\hat{t}_{yc}$	$\hat{t}_{yp}$	$\hat{t}_{y1}$	$\hat{t}_{y2}$
Estimate	61,659	58,686	59,422	58,789
SE (Taylor)	17,793	17,005	16,907	16,952
SE (Jackknife)	21,723	21,646	21,462	21,573

the official 2015 estimate of red snapper catch from Texas is 32,062.) Because of the very small self-reporting rate and imperfect correlation, we would expect to find that $\hat{t}_{y1}$ and $\hat{t}_{y2}$ weight $\hat{t}_{yp}$ more heavily than $\hat{t}_{yc}$, and that is what did occur. The jackknife standard errors are larger than the Taylor standard errors and, based on simulation results, are likely to represent the true uncertainty more accurately. However, unlike the simulation results for the small number of matches case, the (jackknife) standard errors are all similar. This could be because of the larger-than-average catch of the self-reporters or some other feature of its distribution that was not captured in the simulated population.

6. DISCUSSION

In this paper we have developed a methodology that uses participant-provided data as a supplement to data collected by a probability sample. Such volunteer data are increasingly easy to obtain, but as we have demonstrated here, its utility depends on a variety of factors. For MRIP to consider this method of data collection as a replacement for their current catch and effort surveys for all species, they would need to find a way to ensure an adequate reporting rate. If it were possible to require reporting, as Alabama and Mississippi did for red snapper, this would be easier to accomplish. For example, Alabama's reporting rate was 25 percent for private vessels and 67 percent for charter vessels in 2015 (Alabama DCNR/Marine Resources Division 2015). Alternatively, if the most avid anglers could be incentivized to cooperate at a high rate, even if the less active anglers did not, this too could provide sufficient precision for catch estimates.

The best estimator will depend on the situation. Besides the reporting rate, the accuracy of reporting is likely to vary from one application to another. When a fish is easy to identify, like red snapper, the reporting errors are likely to be fewer than for less familiar species. For catch of those species, anglers are likely to misreport one species for another, causing low correlation between reported and observed catch for both species. Thus if self-reports were to be used for estimating catch of all species, the precision of estimates could vary widely and the best estimator to use could vary also. When the reporting rate is low, the number of matches is small, and the species is easy for anglers to identify, $\hat{t}_{y2}$ is our recommended estimator. Besides being easy to compute with standard survey software and having an intuitively attractive form, it would be near optimal among the estimators considered. When the number of matches is large, $\hat{t}_{y1}$ would be recommended. For species that anglers cannot identify easily, $\hat{t}_{yp}$ may be best. In all our simulations, the jackknife estimators of variance were at least as good, and sometimes better, than the Taylor series ones. Thus we recommend jackknife estimators of variance for any estimator.

As with any estimator, nonsampling error can be even more problematic than sampling error due to the bias it can introduce. The usual sources of nonsampling error, nonresponse, and measurement error are typically less problematic in dockside surveys of anglers than in other types, such as the effort surveys made retrospectively by phone or mail. For example, nonresponse for private anglers in the TPWD survey is about 4 percent. In addition, the interviews are conducted by biologists or technicians, who are well qualified and trained to identify species, so that measurement error is minimized.

However, there are other assumptions required in the implementation of the data collection operation that are in common with those of Capture-Recapture models. The first of these is that there is no matching error, meaning that matching the self-reports to access point encounters is error-free. This holds reasonably well in this application since red snapper angling requires a boat. Catch does not need to be associated with individual anglers, but rather just the boat trip, which can be identified with good accuracy by a unique registration number and date/time of arrival at shore. The time report is critical in order to avoid mismatches due to multiple trips per day, so anglers must be educated about the definition of return time. For species that may be caught without a boat, accurate matching will be more difficult, so no attempts to implement these methods have been made.

More problematic is the Capture-Recapture assumption that the population is closed, meaning that no members enter or leave during the sampling period. This holds only if no angler trips eligible to self-report become inaccessible to selection in the verification sample. But anglers who return from their trip to a private dock, such as one behind a home, are removed from the validation sampling frame since dockside surveys can only occur at publicly accessible locations. The access points in the frame are often referred to as public sites, though some private marinas do allow samplers to conduct dockside surveys on their properties. This is a vexing problem for all recreational angler data collection systems. In the estimation system in current use, no measure of catch is available for trips ending at private sites, though counts of these trips are obtained from the effort survey. The unverifiable assumption that catch per trip is identical for trips ending at public and private sites has to be made in order to obtain an estimate of catch. Though the Capture-Recapture approach does provide some information about catch for the private access point anglers via their self-reports (y^*), we still have no source of data for y for these trips. Thus the estimators will not incur bias only under the unverifiable assumption that reporting rate and accuracy is the same for public and private trips. One possible alternative that has been considered is that the verification sample could add some intercepts that occur before landing, such as at fueling sites or on-the-water encounters. This would add its own problems of assessing probabilities of selection into the verification sample, so that further work is needed on this issue.

The Capture-Recapture model also requires assumptions we describe collectively as independence. This encompasses both independence of selection in the two capture periods (selection into the 2nd sample does not depend on capture in the first), as well as homogeneity of selection probabilities in the verification sample. We have generalized the estimation method so that it accounts for known differences in selection probabilities due to the probability sample. The problem occurs when those differences are not known. For example, if the decision to self-report is influenced by selection into the validation sample, this altered probability cannot be accounted for in estimation, and thus can bias the estimator. This could occur if a returning angler was more likely to report his catch if he could anticipate that he would be in the validation sample (what Capture-Recapture methodologists would refer to as “trap happy,” but with the two sampling periods having reversed labels). Care must be taken to prevent this problem by the way the sampling operation and collection of self-report data is implemented. One way to ensure angler reporting status is not influenced by being included in the validation sample would be to conduct the survey out of the view of returning anglers so that the decision to report, which must occur prior to removing fish from the boat in the mandated reporting states, cannot be influenced by the knowledge that they will be interviewed. However, this is difficult to accomplish because the most reliable access to returning anglers is at their landing point. An alternative approach is that used by Mississippi’s program, which makes prior authorization for a red snapper trip mandatory. Specifically, anglers must obtain a trip ID (“open” a trip) prior to departure via the website or app. Since this decision is made by the angler prior to the validation sample encounter, it cannot be influenced by the recapture event. The angler is incentivized to provide catch information after return (“close” a trip) by disallowing issuance of a new trip ID to the angler before he/she closes any open trip. Remaining open trips are followed up by the state agency to obtain catch. If neither of these approaches is used, there is a risk of bias due to differing reporting rates for verified and unverified angler trips.

Still, despite these problems, an estimation procedure that uses angler self-reports holds promise for improving the quality and timeliness of the estimates of catch over those currently available. The decreasing response rates for household surveys nationally have been shared by the effort surveys that are part of the current MRIP estimation methodology. Their dockside access point surveys, which are the ones used here as validation samples, enjoy a much higher response rate. Since all data collection is completed at the time that the trip is made, there is potential for a much faster production of estimates than the current MRIP system since the effort survey is conducted retrospectively. Finally, fisheries management agencies report that some angler advocacy groups are anxious to provide data to improve what they perceive as inadequately precise estimates. This methodology provides a valid way to make use of their shared data.

Supplementary Materials

Supplementary materials are available online at https://academic.oup.com/jssam. They include the details of the simulation study and results mentioned in

Section 4. The simulation study was conducted to investigate the performance for the proposed estimators under several settings for (n2, p1) when the validation sample is SRS.

APPENDIX

A.1 VARIANCES OF THE ESTIMATORS

As noted in (2) and (5), $\hat{t}_{yp}$ and $\hat{t}_{yc}$ are ratio estimators, so their variances can be approximated using Taylor linearization. Thus we see (e.g., Lohr 2009, eq. 4.11) that

$$V(\hat{t}_{yp}) = n_1^2 Var(\hat{B}_p) \approx \frac{N^2\left(1 - \frac{n_2}{N}\right)}{n_2} S_{dp}^2, \tag{A.1}$$

where $S_{dp}^2 = \sum_{i=1}^{N} (y_i - B_p r_i)^2/(N-1)$. This residual variance can be rewritten as

$$S_{dp}^2 = S_y^2 + \bar{y}^2\left(1 + \frac{1}{p_1}\right) - 2\bar{y}\bar{y}_1 \tag{A.2}$$

where $\bar{y} = t_y/N$ and $S_y^2 = \sum_{i=1}^{N} (y_i - \bar{y})^2/(N-1)$ are the mean and variance of y in the entire finite population, $p_1 = n_1/N$ is the fraction of the population in the reporting domain, and $\bar{y}_1 = \sum_{i=1}^{n_1} y_i/n_1$ is the mean of y in this domain. A similar computation yields the variance for $\hat{t}_{yc}$ to have a similar form to (A.1), but with residual variance

$$S_{dc}^2 = S_{dp}^2 + \frac{1}{p_1}\left(\bar{y}/\bar{y}_1^*\right)^2 S_{1y*}^2 - 2\left(\bar{y}/\bar{y}_1^*\right) S_{1,yy^*}, \tag{A.3}$$

where $\bar{y}_1^* = t_{y^*}/n_1$ is the mean of y^* in the reporting domain. The covariance of the two estimators also has the form shown in (A.1), but with residual covariance

$$S_{dp,dc} = S_{dp}^2 - (\bar{y}/\bar{y}_1^*) S_{1,yy^*}. \tag{A.4}$$

Next we consider the variance of the optimally weighted average of these two estimators as defined in (6). Its variance is (Cochran 2007 (6.100))

$$V(\hat{t}_{MR}) = \frac{V(\hat{t}_{yp})V(\hat{t}_{yc}) - Cov^2(\hat{t}_{yp}, \hat{t}_{yc})}{V(\hat{t}_{yp}) + V(\hat{t}_{yc}) - 2Cov(\hat{t}_{yp}, \hat{t}_{yc})}, \tag{A.5}$$

The covariance of the two ratio estimators is

$$Cov(\hat{t}_{yp}, \hat{t}_{yc}) \approx \frac{N^2\left(1-\frac{n_2}{N}\right)}{n_2} S_{dp,dc} = \frac{N^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{S_{dp}^2 - \left(\bar{y}/\bar{y}_1^*\right)S_{1,yy^*}\right\}, \quad (A.6)$$

where $S_{1,yy^*} = \sum_{i=1}^{n_1}(y_i - \bar{y}_1)(y_i^* - \bar{y}_1^*)/(n_1 - 1)$ is the covariance between y and y^* in the reporting domain. Then from (A.1) through (A.6), we have

$$V(\hat{t}_{MR}) \approx \frac{N^2(1-\frac{n_2}{N})}{n_2}(S_{dp}^2 - p_1 S_{1,yy^*}^2/S_{1y^*}^2), \quad (A.7)$$

where $S_{1y^*}^2 = \sum_{i=1}^{n_1}(y_i^* - \bar{y}_1^*)^2/(n_1 - 1)$ is the variance of y^* in the reporting domain.

Finally, it can be observed from (9) that $\hat{t}_{y2}$ can be written as a constant (t_{y^*}) plus a ratio estimator

$$\hat{t}_{y-ry^*} = n_1 \frac{\sum_{i \in S_2}\left(y_i - r_i y_i^*\right)}{\hat{n}_1}.$$

Therefore, the variance of $\hat{t}_{y2}$ can also be approximated using Taylor linearization, yielding

$$V(\hat{t}_{y2}) \approx \frac{N^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{S_{dp}^2 + p_1(S_{1y^*}^2 - 2S_{1,yy^*})\right\}. \quad (A.8)$$

In order to facilitate comparison of these variances, it is helpful to rewrite them in canonical form as follows:

$$V(\hat{t}_{yp}) = \frac{t_y^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{CV_y^2 + \left(1+\frac{1}{p_1}\right) - 2\left(\frac{\bar{y}_1}{\bar{y}}\right)\right\}; \quad (A.9)$$

$$V(\hat{t}_{yc}) \approx V(\hat{t}_{yp}) + \frac{t_y^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{\frac{CV_{1y^*}^2}{p_1} - 2\left(\frac{\bar{y}_1}{\bar{y}}\right)R_{1,yy^*}CV_{1y}CV_{1y^*}\right\}; \quad (A.10)$$

$$V(\hat{t}_{MR}) \approx V(\hat{t}_{yp}) - \frac{t_y^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{p_1\left(\frac{\bar{y}_1}{\bar{y}}\right)^2 R_{1,yy^*}^2 CV_{1y}^2\right\}; \quad (A.11)$$

$$V(\hat{t}_{y2}) \approx V(\hat{t}_{yp}) + \frac{t_y^2\left(1-\frac{n_2}{N}\right)}{n_2}\left\{p_1\frac{\bar{y}_1^*}{\bar{y}}CV_{1y^*}\left(\frac{\bar{y}_1^*}{\bar{y}}CV_{1y^*} - 2\frac{\bar{y}_1}{\bar{y}}R_{1,yy^*}CV_{1y}\right)\right\}. \quad (A.12)$$

REFERENCES

Alabama DCNR/Marine Resources Division (2015), "Summary of Alabama's 2015 Mandatory Red Snapper Reporting Program," Available at http://www.outdooralabama.com/sites/default/files/preliminary results of alabama's 2015 red snapper reporting.pdf.

Alho, J. M., M. H. Mulry, K. Wurdeman, and J. Kim (1993), "Estimating Heterogeneity in the Probabilities of Enumeration for Dual-system Estimation," *Journal of the American Statistical Association*, 88(423), 1130–1136.

Baker, R., J. M. Brick, N. A. Bates, M. Battaglia, M. P. Couper, J. A. Dever, K. J. Gile, and R. Tourangeau (2013), "Summary Report of the AAPOR Task Force on Non-probability Sampling," *Journal of Survey Statistics and Methodology*, 1(2), 90.

Berkson, J. (1950), "Are There Two Regressions?" *Journal of the American Statistical Association*, 45(250), 164–180.

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu (2006), *Measurement Error in Nonlinear Models: A Modern Perspective*, Boca Raton, FL: CRC Press.

Cochran, W. G. (2007), *Sampling Techniques*, New York: John Wiley & Sons.

Groves, R. M. (2011), "Three Eras of Survey Research," *Public Opinion Quarterly*, 75(5), 861–871.

Johnson, N. L., and S. Kotz (1970), *Distributions in Statistics: Continuous Univariate Distributions*, *Vol.* 1, New York: Houghton Mifflin.

Le Cren, E. D. (1965), "A Note on the History of Mark-Recapture Population Estimates," *The Journal of Animal Ecology*, 34, 453–454.

Lee, S., and R. Valliant (2009), "Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment," *Sociological Methods & Research*, 37(3), 319–343.

Lohr, S. (2009), *Sampling: Design and Analysis, Pacific Grove*, CA: Duxbury Press.

Lumley, T. (2004), "Analysis of Complex Survey Samples," *Journal of Statistical Software*, 9(1), 1–19.

National Research Council (2006), *Review of Recreational Fisheries Survey Methods*, Washington, DC: The National Academies Press.

Olkin, I. (1958), "Multivariate Ratio Estimation for Finite Populations," *Biometrika*, 154–165.

Pollock, K., S. Turner, and C. Brown (1994), "Use of Capture-Recapture Techniques to Estimate Population Size and Population Totals When a Complete Frame is Unavailable," *Survey Methodology*, 20, 117–124.

Tarnecki, J. H., and W. F. Patterson III (2015), "Changes in Red Snapper Diet and Trophic Ecology Following the Deepwater Horizon Oil Spill," *Marine and Coastal Fisheries*, 7(1), 135–147.

Tourangeau, R., F. G. Conrad, and M. P. Couper (2013), *The Science of Web Surveys*, New York: Oxford University Press.

Wolter, K. M. (1986), "Some Coverage Error Models for Census Data," *Journal of the American Statistical Association*, 81(394), 337–346.