Detection Accuracy in the Hawai‘i Longline Electronic Monitoring Program with Comparisons between Three Video Review Speeds

156th SSC

Jennifer P. Stahl[1,2] and Matthew J. Carnes[3]

This report includes results for a study comparing review speeds of electronic monitoring (EM) data from the Hawai‘i longline fisheries (deep-set and shallow-set fisheries) and was funded by the National Catch Share Program and Magnuson-Stevens Act Implementation Temp Funds from fiscal year 2019. The intent of the study was to determine accuracy of detection at three different speeds for EM video review, including review of video with known protected species interactions, and to outline best practices for EM data reviewers.

PIFSC Data Report DR-20-012. Issued 18 September 2020.
https://doi.org/10.25923/n1gq-m468

1 Joint Institute for Marine and Atmospheric Research, University of Hawai‘i at Manoa, 1000 Pope Road, Honolulu, HI 96822
2 Jennifer P. Stahl, corresponding author, Jennifer.stahl@noaa.gov
3 Matthew J. Carnes, matthewjcarnes@gmail.com

Contents

Tables

Figures

Introduction

Since 2017, a pilot program has been conducted in the Pacific Islands Region (PIR) to determine the efficacy of electronic monitoring (EM) in the Hawai‘i-based commercial longline deep- and shallow-set fisheries. Currently, these fisheries are monitored by at-sea observers that are deployed on 100% of shallow-set and 20% of deep-set trips. On average, these trips consist of 13 pelagic longline sets with 2,643 hooks per set (FRMD 2018) that are monitored by at-sea observers during gear retrieval to enumerate catch and protected species interactions. In 2018, detection rates were compared between these at-sea observers and EM systems for fish and protected species captured in the Hawai‘i longline fisheries (Carnes et al. 2019). A total of 238 fishing hauls (a haul is the retrieval of a set of longline gear that was baited and deployed to catch fish) were selected from 17 trips performed by vessels equipped with EM systems. Data derived through EM review during the 2018 comparisons were similar to those collected by at-sea observers with 89% of catch (in numbers) enumerated by EM (n = 15,180) compared to at-sea observer data (n = 17,052) for all retained and discarded fish for the shallow-set and deep-set fisheries. A data gap was identified for the detection of discarded sharks and non-marketable bycatch of longnose lancetfish (*Alepisaurus ferox*) and snake mackerel (*Gempylus serpens*); when these fish were removed from the analysis, 97% of retained and discarded species were identified by EM compared to the at-sea observer. For protected species—marine mammals, sea turtles, and seabirds—EM systems were able to detect the only false killer whale (*Pseudorca crassidens*) and the four hard shell sea turtles (olive ridley, *Lepidochelys olivacea*, and loggerhead sea turtles, *Caretta caretta*) identified by the at-sea observer. Eleven of the seventeen albatrosses (*Phoebastria nigripes, Phoebastria immutabilis*) reported by the at-sea observer were detected from the EM system. Future research is planned in 2020 to investigate catch handling and EM system configurations that improve detection of discarded fish, sharks, and seabirds.

The pilot study indicated that EM systems could be utilized for monitoring Hawai‘i longline fisheries; however, more research was needed to outline procedures for the review of EM video data. A study began in November 2018 to evaluate the preferred speed for video review for the most accurate detection rates while considering the time and cost of review. Prior to this study, no other research had been conducted to determine the optimal speed for EM review. This study tested the detection accuracy of speeding up video during the review process because speeds faster than real-time are likely to improve reviewer engagement and reduce costs. We tested speeds of video played at 4, 8, and 16 times (4×, 8×, and 16×) that of normal playback speed. Hawai‘i-based longline fisheries are well suited for EM review at high speed due to the relatively large size of fish and protected species in these fisheries, the large spacing between hooks, and the relatively slow rate of fishing (2,722 hooks per 12-hour deep set; FRMD 2018). The video review speed of 16× was chosen as the highest speed to investigate due to system performance restraints; at this playback speed a hook would still only be encountered approximately every second, similar to the speed at which at-sea observers enumerate fish brought aboard in Alaska’s groundfish longline fleet. At review speeds faster than 16×, the video skips, which results in catch events not appearing on the reviewer’s screen. The minimum speed in our study was 4×, which was chosen to improve review accuracy over review at real-time speed. In a typical week, at 4× speed, 10 hauls can be reviewed; however, at real-time speed, only 3 hauls would be completed. In addition, this study was conducted to determine best practices for structuring an EM program that would maintain reviewer engagement through countless hours of tedious review. Another focus of this study was to determine the ability to detect and identify protected

species using EM; consequently, trips with protected species interactions were selected *a priori* without the knowledge of the reviewer.

Methods

Detection of fish and protected species captured in the Hawai‘i longline fisheries was compared for EM video review speeds of 4×, 8×, and 16×. To ensure that reviewers had familiarity with the nuances of fishermen and fish behavior and were knowledgeable on species identifications, three reviewers were selected with previous experience as at-sea observers in the Hawai‘i longline fisheries through the Pacific Islands Regional Observer Program (PIROP). Protected species interactions in Hawai‘i longline fisheries are rare events; however, they are important to fisheries management. Consequently, fishing trips from both deep- and shallow-set fisheries were selected *a priori* with known protected species interactions by the project coordinator for review to increase protected species sample sizes and determine if detection rates would vary between individual reviewers and speed of review.

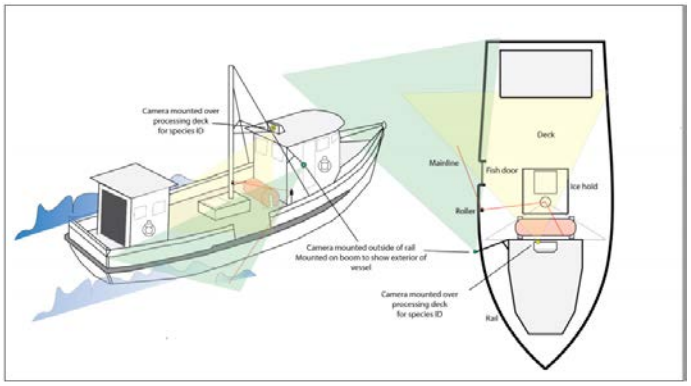

Figure 1. Preferred camera configuration on pelagic longline vessels.

For this study, EM systems deployed on Hawai‘i longline vessels consisted of a “deck” and a “rail” camera (Figure 1), as well as sensors for GPS, hydraulic, and magnetic rotation. Additional details on system configuration can be found in Carnes et al. (2019). The software used for EM data review was downloaded from the Chordata bitbucket repository.[4,5] JIMAR staff configured this software to contain separate fields for species name and species code to find inconsistencies during quality assurance examinations performed immediately after completing video review.

EM video and sensor data were reviewed together for a complete picture of fishing. The two EM cameras were reviewed simultaneously while a window with a timeline displayed sensor data. When a catch event was enumerated, a corresponding mark was created on the timeline. Sensor information, such as reel rotation and declines in hydraulic pressure and speed could be used to

[4] https://bitbucket.org/fisherieselectronicmonitoring/

[5] Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the NOAA Pacific Islands Fisheries Science Center of the Department of Commerce or any of their employees/contractors.

identify a possible catch. After all video footage was examined from a haul, the reviewer scanned the timeline for changes in sensor data with no matching catch event marks; these sections of video were reexamined to ensure no catch event was missed in the initial review.

EM data were examined for protected species, retained fish (kept catch), and bycatch (discarded catch) interactions with reviewers instructed to enumerate all catch events. Reviewers watched video at the specified speed (4×, 8×, or 16×) for any visible fish or protected species or any changes in crew behavior that could be an indication of a catch event. Changes in the hand position of the roller operator indicated an animal could be hooked and alerted the reviewer to watch for fish landed through the fish door, fish discarded at the hook boxes (where branchlines are coiled), or multiple crew working to pull in a line, an indication a large animal was caught on the line. Once a catch event was identified, the video was slowed down or paused to identify the species to the lowest taxonomic level possible. If the animal was not observed before the line broke or was cut, then the catch was enumerated as "unknown catch event". If a fish was observed being pulled to the vessel with relative ease but could not be identified, then the fish was enumerated as "unidentified bonyfish."

Reviewers were trained to use the same methods for identification and data entry for EM video review to improve accuracy and consistency. They were advised to limit review to six hours in a day and to not split review of hauls over multiple days to ensure the information was "fresh" and there was enough time in the day to complete quality assurance procedures. In addition, reviewers were instructed to review at speeds no higher than their specified speed and to not skip any video footage while gear was in the water. These protocols were instated to improve the likelihood that reviewers would not miss animals that were only briefly on camera or were associated with tangled gear.

To evaluate differences in reviewer detection rates, three EM reviewers were instructed to watch the same 30 hauls from three different fishing trips at 8× speed. These trips, reviewed by all three at-sea observers, were also used to ascertain differences in detection rates between trips. An additional 89 hauls were provided for review by two different reviewers with one using a 4× video playback speed and the other a 16× playback speed for comparisons of detection rates. The total number of hauls provided to reviewers for this study was equivalent to the EM pilot study (Carnes et al. 2019) and was within in the scope of the funding provided by the National Catch Share Program and Magnuson-Stevens Act Implementation Temp Funds.

This study was originally designed with the intention to determine the effect of speed on detection using paired reviews at 4× and 16× speed. A shift in the original study design occurred when the contractor delivered an incomplete data set. Instead, at-sea observer data were used as the basis for comparison of detections between EM video review speeds, reviewers, and fishing trips. At-sea observer detections were assumed to have no bias and were used to calculate EM detection accuracy using similar methods as Dressel and Herbert (2014):

Equation 1. $Crfs = Of/Erfs$

where Crf is the EM detection accuracy by reviewer (r), catch retention type (f), and speed (s). Of is the at-sea observer detections by catch retention type (f), and Erf is the EM review

detections by reviewer (*r*) and catch retention type (*f*). The *Crf* was first calculated by haul and then averaged across hauls to summarize for each reviewer, speed, and retention type. EM detection accuracy values less than 1 indicate more detections were performed by the EM reviewer compared to the at-sea observer; whereas, accuracy values greater than 1 indicate the at-sea observer enumerated a higher number of detections compared to the EM reviewer. EM detection accuracy was calculated separately for retained and discarded catch because we expected there may be differences in detection ability between these different dispositions based on results of a previous study (Carnes et al. 2019). Protected species were included within the discarded catch for the EM detection accuracy calculations; EM detection accuracy values were not calculated separately for protected species due to small sample sizes. To determine if differences existed in detection rates between trips, we calculated EM detection accuracy using equation 1 for the three trips processed by all reviewers at 8× playback speed; the *Crf* was calculated by haul and then averaged for each trip, reviewer, and retention type.

Results

Video reviewed by speed

EM data review was completed from a total of ten Hawai‘i longline fishing trips with three reviewers watching video from 240 fishing hauls at playback speeds of 4×, 8×, and 16× (Table 1). At a playback speed of 4×, six different fishing trips were watched with a total of 78 hauls (65 unique hauls), and at 8× speed, three different trips were watched with a total of 90 hauls (30 unique hauls). At 16× speed, six different trips were watched for a total of 72 hauls (72 unique hauls).

EM detections performed

During this study, a total of 13,708 detections were performed by reviewers from EM video. A total of 9,274 of these detections were from retained fish and 4,434 from discarded catch, which included 3,625 fish, 41 protected species, and 768 unknown catch events.

EM detection accuracy between reviewers

EM detection accuracy between reviewers was similar for retained catch but was variable for discarded catch when comparisons were performed from the three trips watched by all three reviewers at 8× speed (Table 2 and 3). For retained catch, the number of detections enumerated by reviewers was similar to the at-sea observer with the average EM detection accuracy for all three reviewers of 0.99 (Table 2). Generally, the at-sea observer detected more discarded fish per haul than the EM reviewer (Figure 3) with an average EM detection accuracy value greater than 1.00 for all reviewers. Some variability occurred between the average EM detection accuracy among reviewers for discarded catch with values ranging from 1.10 to 1.51 (Table 2).

EM detection accuracy between trips

Average EM detection accuracy between trips was similar for retained catch but more variable for discarded catch when comparisons were performed from the three trips watched by all three reviewers at 8× speed (Figures 4 and 5; Table 3). For retained fish, EM reviewers had similar average EM detection accuracy values for all three trips watched with slightly more fish detected for trip 2 for all reviewers compared to the at-sea observer (Figure 4; Table 3). Average EM detection accuracy was more variable by trip for discarded catch. The at-sea observer detected

more fish than the EM reviewer for all trips (Figure 5; Table 3). The fewest detections occurred for trip 1 and the most for trip 2 for the EM reviewer compared to the at-sea observer with detection accuracy values ranging from 1.16–1.48 for all three trips (Figure 5; Table 3).

EM detection accuracy between review speeds

For retained catch, the average EM detection accuracy is similar at all review speeds with little variability, and the number of detections was similar between EM reviewers and the at-sea observer data with average EM detection accuracy values close to 1 for all playback speeds (Table 2; Figure 2). For discarded catch, on average, reviewers have the best detection accuracy (closest to 1) at 8× speed. Reviewers 1 and 2 had the fewest detections compared to the at-sea observer at a review speed of 16×; whereas reviewer 3 had the fewest detections compared to the at-sea observer at 4× and 16× review speeds with similar detection accuracy values at both speeds (Table 1; Figure 3).

EM detection of protected species

EM reviewers were able to detect protected species interactions with the Hawai‘i longline fisheries (Figure 6; Table 4). A total of six marine mammals, 31 sea turtles, four albatrosses, and three protected shark[6] detections were performed by EM reviewers; some of these animals were detected by more than one reviewer that watched the same fishing trip for a total of three individual marine mammals, 18 different individual sea turtles, three unique albatross, and two protected shark detections performed. One sea turtle and one marine mammal were not detected by a reviewer; however, both these animals appeared in video footage based on subsequent review. One albatross (black-footed, *Phoebastria nigripes*) was missed that appeared in video footage, and three additional albatrosses (two black-footed and one Laysan albatross *Phoebastria immutabilis*) were missed that were not captured in video footage. Five protected oceanic whitetip sharks were missed by EM reviewers (Table 4). No other protected sharks or manta rays were enumerated by at-sea observers or EM reviewers in this study. We were unable to determine why the oceanic whitetip sharks were missed because these animals appear in limited video footage compared to other protected species due to the manner in which they are handled by fishermen. In addition, at-sea observer shark detections are difficult to match with EM video footage because at-sea observer data are not time stamped like EM data and instead are organized by approximate hook count.

Reviewers were able to detect protected species at all video review speeds, but with different accuracies (Figure 6; Table 4). Thirteen individual sea turtles were detected at 4× speed, one at 8× speed (by all three reviewers), and 15 at 16× speed. The only sea turtle not detected during video review was a leatherback sea turtle (*Dermochelys coriacea*) missed during a 4× playback speed. This missed leatherback sea turtle was in the video footage and detected by another reviewer. Two individual marine mammal detections were performed at 4× speed, one at 8× speed (by all three reviewers), and one at 16×. The only marine mammal that was not detected was a Risso’s dolphin (*Grampus griseus*), which was missed at 4× video speed. This animal was observed when the project manager, M. Carnes, reviewed video footage. Albatrosses were in video footage that was reviewed at 4× and 16× speeds with reviewers detecting three seabirds at 4× speed and one seabird at 16× speed. One of the albatrosses missed at 16× speed was observed

[6] Sharks protected under the Endangered Species Act include the oceanic whitetip, *Carcharhinus longimanus*, which was the only protected shark detected in this study and was designated as threatened in 2018.

by another reviewer (Table 4). The rest of the albatrosses missed by reviewers were not in the EM video footage as confirmed by the project manager, M. Carnes. Oceanic whitetip sharks were in video footage that was reviewed at 4× and 16× speeds. At 4× speed, one oceanic whitetip shark was detected; at 16× speed, two oceanic whitetip sharks were detected (Table 4).

EM identification of protected species

Protected species identifications were correctly performed from EM video to the species level for two of the three detected marine mammals, 16 of the 18 detected sea turtles, and all of the four albatross detections performed. All three reviewers were able to correctly identify the only false killer whale in the video footage during this study. One of the two Risso's dolphins detected in the video footage during the study was correctly identified; however, the other detected Risso's dolphin was called an unidentified dolphin by one reviewer and an unidentified dolphin or whale by the other reviewer. The majority of the sea turtles detected during the study were correctly identified to the species level. Of the 15 different loggerheads, 13 were identified to the species level and the other two were categorized as unidentified sea turtle. Both of the olive ridley sea turtles in the video footage were identified to the species level, although one of the three reviewers that detected a particular olive ridley categorized it as an unidentified sea turtle. The reviewer that detected the leatherback sea turtle correctly identified this sea turtle. All of the albatross detected by the reviewer were correctly identified as black-footed albatross.

Discussion

Study limitations

It is difficult to draw conclusions based on statistics for the optimal video review speed for retained and discarded fish due to limitations in the study design and high variability between fishing trips and reviewers. This study was originally designed with the intention to determine the effect of speed on detection. The between reviewer variance was to be accounted for by adjusting the detections at the speed trials of 4× and 16× to the reviewer's detection accuracy at the baseline speed of 8×. The review speed of 8× was selected as the baseline, because a previous EM study in the Pacific Islands conducted review at that speed and was able to detect 89% of all retained and discarded catch as determined when EM detections were compared to the at-sea observer detections (Carnes et al. 2019). Thus, for this study, the same three trips were watched by all three reviewers at 8× speed. All other trips were prescribed for review at both 4× and 16× speed by two different reviewers. However, the contractor did not complete review of all prescribed hauls, which limited the available data at those speeds. As a result, we shifted the methods for analysis to include hauls also reviewed at 8× speed and compared the EM data to observer data to assess detection accuracy for each review speed. The resulting analysis had a limited sample size with an unequal distribution of the number of hauls reviewed at each video review speed. Analyses of these limited data indicate that the variance between trips and reviewers may be greater than the variability between speeds.

EM detection and identification of protected species

Despite our limited ability to draw conclusions on the optimum review speed for retained and discarded catch, we did gain insight on issues that may occur with video review of protected

species at particular review speeds. At 4× speed, reviewers missed protected species possibly due to loss of focus when progress was slow towards completion of a haul. Although in this study no sea turtles or marine mammals were missed by the reviewers at a playback speed of 16×, we caution the adoption of this video review speed for an EM program that prioritizes the detection of protected species interactions. During this study, sea turtles and marine mammals were generally handled according to requirements to bring animals aboard or alongside the vessel to allow for identification and removal of gear. However, it is possible that protected species might not be handled as recommended or may break free from gear during handling and appear in limited video footage and, therefore, more likely missed at 16× playback speed. One black-footed albatross was missed at this speed but observed at 4× playback; smaller protected species, such as seabirds, may be more difficult to detect at the fastest speed.

Reviewers may be more engaged at faster speeds compared to 4× playback speed and, as a result, miss fewer catch events. The slowest speed was the only one at which reviewers missed sea turtles or marine mammals (one leatherback and one Risso's dolphin). In this study, reviewers missed the two protected species because they were likely rushing to process video footage at this slow speed. During the missed leatherback interaction, sensor data indicated that the reel was not rotating and the hydraulic pressure and vessel speed had decreased; consequently, the reviewer likely interpreted these data as no fishing activity was occurring (e.g., crew temporarily stopped fishing to deal with a gear tangle or to load fish into the freezer), decided to skip ahead, and began reviewing video again after this section. In addition, the reviewer that missed the Risso's dolphin probably skipped that bit of footage. A shark was captured on a hook near the Risso's dolphin, which was detected by the reviewer; this shark interaction may have been interpreted by the reviewer as the reason the crew had temporarily stopped fishing. Consequently, the reviewer skipped the rest of the time period when the fishing activity appeared to be stopped (i.e., the reel was not rotating, and the hydraulic pressure and vessel speed had dropped).

EM detection of retained and discarded fish

Higher variation exists in the detection accuracy of discarded fish compared to retained catch between review speeds. Retained fish in the Hawai‘i longline fisheries are generally large, easy to distinguish, and require handling time that results in sufficient video footage for detection; consequently, these fish had high detection accuracy at all playback speeds. However, variable EM detection accuracy (on average) occurred between the tested playback speeds for discarded catch. These animals are often not landed and may be difficult to detect. Reviewers must be diligent and focused to detect discarded catch and watch for crew behaviors that indicate a catch event, such as changes in hand movements by the fishermen at the roller (location where branchlines are removed from the mainline). This change in behavior alerts the reviewer to scan the video for discarded catch shaken or removed from the branchline, a larger animal being pulled in by the crew, and/or an animal cut from the branchline—a method often used to discard sharks. If an EM program is implemented in the Hawai‘i longline fisheries with the intent to collect information similar to the at-sea observer program, then collection of discard data would likely be prioritized, and a playback speed of 16× speed would be too fast to consistently detect crew behaviors associated with discarded catch that may appear in limited video footage.

EM detection accuracy between trips

Variability in average EM detection accuracy between trips was also low for retained catch but more variable for discarded catch. Vessels likely handle retained catch similarly with these animals all occurring in video footage for a sufficient time for detection. However, each vessel may handle discarded catch, especially sharks, differently. Sharks composed 24% of the discarded catch (by number)[7]; consequently, differences in the ability of the reviewer to detect sharks between vessels will result in overall differences in detections between trips. All three vessels compared for trip differences, handled the sharks by cutting the branchline once the crew verified a shark was captured. The location of the shark when this action occurred compared to the camera field of view likely differed between vessels, which would have affected detection ability between these vessels. The catch handling study that will be conducted in 2020 will endeavor to standardize shark detection ability between vessels.

EM detection accuracy between reviewers

Variability in average EM detection accuracy between reviewers was low for retained catch and moderate for discarded catch. Retained fish are consistently detected in video footage for all reviewers. However, discarded catch was often overlooked due to differences in reviewers' understanding of the nuances of the fishery or their attention span. Reviewers' understanding of the fishery varies because individuals have been deployed on different vessels with a variety of handling practices for bonyfish and sharks. Variation in review also occurs, because of differences in reviewer attention span that may lead to less thorough perusal of the video. For example, two different reviewers that watched video at 4× speed skipped footage during periods they assumed to only include fishermen working to fix gear issues which was discovered based on follow-up questions with reviewers. These periods actually included fishermen working to remove protected species captured on gear: a leatherback sea turtle and a Risso's dolphin.

At-sea observer data considerations

Within this study, at-sea observer data are assumed to be accurate and have no bias; however, in some cases, at-sea observers may miss catch or misidentify species. For retained catch, EM reviewers had similar or even better detection compared to at-sea observers (e.g., during one trip, all three reviewers detected at least 20 tuna more than the at-sea observer). Hauling gear may take an average of 12 hours; consequently, at-sea observers may require breaks. In rare cases, at-sea observers may not perform their job duties due to injury or illness, such as sea sickness, or other undetermined reasons. Consequently, EM provides a data stream for retained catch that maintains video coverage throughout the entire haul and is verifiable. For discarded catch, at-sea observers are currently able to collect more robust, accurate data with the ability to see farther than the EM camera; however, our future catch handling study will focus on improving discard detection ability using EM.

Further research

The high variability in the average EM detection accuracy between reviewers and trips and the low detection of EM compared to the at-sea observer for discards indicate there is a need to improve the ability of discard detection using EM. Consequently, a catch-handling study will be performed in 2020 to address shortcomings with detections of discarded fish and sharks. In this

[7] Based on unpublished PIROP database query for 2018 discard composition in the Hawai'i longline fisheries.

study, fishermen will be requested to bring sharks and other discarded fish alongside the vessel in the camera field of view before releasing them.

Detection of protected sharks was low; they appeared in limited video footage and remained out of the view of the cameras. Although regulations were in place during this study that required fishermen to bring protected sharks within the view of the observer for identification and to release them in a manner that minimizes injury, many sharks were likely brought to the back of the vessel out of the view of the cameras and cut from fishing line quickly. Consequently, the catch-handling study that will be performed for discarded species in 2020 will bring these protected sharks into the view of the cameras to help improve detection.

Catch and handling requirements for albatross and other birds may improve detection. The three albatross that were completely missed by the EM reviewers were not in video footage; modifications in catch and handling by fishermen, such as bringing animals on to the deck or by the fish door, may improve capture in video footage and thus detection. However, albatross that are temporarily hooked (e.g., on a wing) or wrapped in line may free themselves before they are caught on video. This type of interaction may have occurred with the three albatross that were completely missed in this study as they were recorded by the at-sea observer as released injured but did not appear in video footage.

EM systems provide an efficient way to detect protected species interactions in Hawai‘i longline fisheries; however, more research is needed to determine if EM systems are able to collect all protected species information needed for management. The EM systems were able to capture video footage of protected species and reviewers were generally able to detect interactions and identify to the species level. Future investigations will focus on determining if EM systems are able to collect data needed for bycatch assessments and criteria to perform serious injury determinations of marine mammals, i.e., length estimations for trailing branchline and animal size, hooking location and removal, and condition of the animal at capture and release. Modifications to EM systems and handling protocols will be explored to improve protected species data collection.

If an EM program was planned for catch monitoring and did not have the aforementioned concerns regarding detection of protected species and discards, then a future study may be warranted to ascertain the optimum video review speed. A more balanced study design would improve the ability to detect differences between video review speeds. To better assess trip and reviewer variability, we suggest that a total of 240 unique hauls be reviewed. At each review speed (4×, 8×, 16×), all three reviewers would watch the same 60 hauls for a total of 180 unique hauls. For example, reviewers 1, 2, and 3 would watch the same 60 hauls at 4×. An additional 60 hauls would be reviewed by all three reviewers at 8×, and 60 more by all reviewers at 16×. This sample size would allow an 80% power to detect differences between reviewer detection accuracy at each review speed. Because all reviewers would watch the same trips at a particular speed, between-trip variability could be assessed as well. To assess variability between review speeds, 60 additional hauls would be watched at all three review speeds. For example, reviewer 1 would watch all hauls from a particular trip at 4× while reviewer 2 would watch these same hauls at 8× and reviewer 3 at 16×. Review of these additional hauls would provide an 80% power to detect differences in detection accuracy between review speeds. This study design would remove the limitations of unequal sample size across review speeds and provide more insight on the

between-trip and reviewer variability in order to draw further conclusions on the optimum review speed.

EM program recommendations

To ensure accuracy and timeliness in the EM data stream, we suggest programmatic recommendations for video review that are applicable to other EM programs. It is recommended that EM reviewers have previous at-sea observer experience in the fishery of the particular implemented EM program so they understand the nuances of the fishery and fishers and to ensure accurate species identifications. To maintain knowledge in the fishery and reviewer engagement, we suggest reviewers participate as an at-sea observer at least once annually in the fishery of the implemented EM program. In addition, a prescribed number of hauls per week should be required by reviewers to ensure annual targets are completed on schedule; however, we recommend against a daily prescribed number of hauls. Video review requires a high level of attention; consequently, reviewers should only perform review when feeling well, focused, and able to commit to the rigors of review necessary to produce high data quality. To prevent fatigue, the amount of review per workday should not exceed six hours with productivity breaks taken during this time period to allow reviewers to return to review focused. Reviewers should have other work responsibilities in addition to video review that allows balance between duties.

Standardized operating procedures need to be established and followed to ensure video reviewers are processing all EM video properly and do not miss protected species or other animals that are captured in EM footage. Protected species capture events may require normal fishing activity to stop to allow time for the crew to handle animals according to defined protocols; consequently, reviewers need to review all video footage to ensure that a change in fishing activity is not due to a protected species capture event. The requirement to review all video footage while fishing gear is in the water during a haul needs to be outlined in EM review operating procedures, and audits of reviewers will need to be performed to ensure they are meeting review requirements.

We recommend EM review software that allows for verification of detections and identifications. Pause and rewind features accommodate a playback speed faster than real-time, because reviewers are able to stop footage when a possible interaction is observed. Reviewers are then able to carefully reexamine footage to confirm a detection or to identify characteristics for species identification. These features as well as the ability to capture images or video are essential to allow a secondary review of protected species by experts to confirm species identifications and collect supplemental information necessary to estimate mortality (e.g., trailing gear). Accurate information on protected species is crucial as these data may lead to management actions. The ability for EM software to capture images and video clips also provides documentation of protected species information without the need to save all video footage from a trip or a haul. These features may allow for a reduction in storage costs—a desire expressed by commercial fishermen for implemented EM programs.

References

Carnes MJ, Stahl JP, Bigelow KA. 2019. Evaluation of Electronic Monitoring Pre-implementation in the Hawaii-based Longline Fisheries. NOAA Technical Memorandum. NMFS-PIFSC -90, 38 p. doi:10.25923/82gg-jq77.

Dressel SC, Hebert KP. 2014. Underwater Calibration Sampling to Correct for Bias in Visual Estimates of Pacific Herring Egg Abundance on Macroalgae by Scuba Divers in Southeast Alaska. In Diving for Science 2014: Proceedings of the American Academy of Underwater Sciences 33rd Symposium. Eds. Eckert G., S. Keller and S. Tamone.

Fisheries Research and Monitoring Division [FRMD], Pacific Islands Fisheries Science Center, NOAA Fisheries. 2018. The Hawaii limited access longline logbook summary report, January to December 2017. Pacific Islands Fisheries Science Center, PIFSC Data Report, DR-18-007, 10 p. doi:V5/DR-PIFSC-18-007.

Pacific Islands Fisheries Science Center. 2020. Electronic Monitoring in the Pacific Islands Longline Fisheries Data, https://www.fisheries.noaa.gov/inport/item/62654

Pacific Islands Regional Office, 2020: Longline Observer Data System, https://www.fisheries.noaa.gov/inport/item/9027

Tables and Figures

Table 1. Number of fishing hauls, by Hawai‘i longline fishing trip, watched at three different video speeds by electronic monitoring (EM) reviewers.

Speed	Trip ID	Reviewer ID	Number of hauls reviewed
4×	4	3	9
4×	5	1	4
4×	6	2	11
4×	7	3	12
4×	8	1	13
4×	8	3	13
4×	9	2	16
4×	10	2	0
8×	1	1	10
8×	2	1	11
8×	3	1	9
8×	1	2	10
8×	2	2	11
8×	3	2	9
8×	1	3	10
8×	2	3	11
8×	3	3	9
16×	4	1	9
16×	5	2	0
16×	6	3	16
16×	7	1	4
16×	8	2	13
16×	9	3	16
16×	10	3	14

Table 2. Average electronic monitoring (EM) detection accuracy for three reviewers and at three playback speeds for electronic monitoring data collected from Hawai‘i longline vessels. EM detection accuracy was calculated per haul by comparing EM to at-sea observer data with detection accuracy calculated separately for retained and discarded catch.

		Average detection accuracy (number of at-sea observer detections, number of EM detections)			
Catch retention type	**Speed**	**Reviewer 1**	**Reviewer 2**	**Reviewer 3**	**Reviewer average**
Kept	**4×**	0.95 (676, 724)	0.99 (800, 803)	0.98 (1404, 1438)	0.97
Kept	**8×**	0.99 (1268, 1288)	0.99 (1268, 1293)	0.99 (1268, 1288)	0.99
Kept	**16×**	1.01 (581, 576)	0.95 (582, 619)	1.00 (1246, 1245)	0.98
Average Kept	**All speeds**	0.98	0.97	0.99	0.98
Discarded	**4×**	1.10 (367, 353)	1.68 (550, 328)	1.26 (822,699)	1.35
Discarded	**8×**	1.25 (789, 650)	1.51(789, 568)	1.10 (789, 729)	1.28
Discarded	**16×**	2.14 (448, 263)	3.20 (271, 105)	1.27 (888, 739)	2.20
Average Discarded	**All speeds**	1.50	2.13	1.21	1.61
Average Kept & Discarded	**All speeds**	1.24	1.55	1.10	1.30

Table 3. Average electronic monitoring (EM) detection accuracy for three reviewers for three different Hawai‘i longline fishing trips reviewed at 8× playback speed. EM detection accuracy was calculated per haul by comparing EM to at-sea observer data with detection accuracy calculated separately for retained and discarded catch.

		Average detection accuracy (number of at-sea observer detections, number of EM detections)			
Catch retention type	**Trip**	**Reviewer 1**	**Reviewer 2**	**Reviewer 3**	**Reviewer average**
Kept	**Trip 1**	1.00 (418, 417)	1.00(418, 416)	1.00 (418, 417)	1.00
Kept	**Trip 2**	0.96 (493, 522)	0.94 (493, 530)	0.96 (493, 521)	0.95
Kept	**Trip 3**	1.02 (357, 349)	1.03 (357, 347)	1.01 (357, 350)	1.02
Average Kept	**All trips**	0.99	0.99	0.99	0.99
Discarded	**Trip 1**	1.33 (254, 203)	1.95 (254, 143)	1.15 (254, 228)	1.48
Discarded	**Trip 2**	1.18 (252, 216)	1.22 (252, 212)	1.08 (252, 231)	1.16
Discarded	**Trip 3**	1.24 (283, 231)	1.36 (283, 213)	1.06 (283, 270)	1.22
Average Discarded	**All trips**	1.25	1.51	1.10	1.29
Average Kept & Discarded	**All trips**	1.12	1.25	1.04	1.14

Table 4. Reviewer detection of protected species interactions that occurred in selected video footage from the Hawai‘i longline fisheries. Interactions that were detected by the reviewer are indicated with the symbol ✓; whereas a × symbol indicates a missed interaction. All three reviewers reviewed the same hauls at 8× speed. All other hauls were watched at 4× and/or 16× speeds. Not applicable (N/A) indicates the interaction was not reviewed at a particular speed.

			Reviewer detection by speed					
Species	**Scientific name**	**Observation #**	**4×**	**8×**	**8×**	**8×**	**16×**	**Why not detected**
Black-footed albatross	*Phoebastria nigripes*	1	✓	N/A	N/A	N/A	✓	
		2	✓	N/A	N/A	N/A	N/A	
		3	×	N/A	N/A	N/A	×	Not in video footage.
		4	×	N/A	N/A	N/A	×	Not in video footage.
		5	✓	N/A	N/A	N/A	×	In limited video footage.
Laysan albatross	*Phoebastria immutabilis*	1	×	N/A	N/A	N/A	×	Not in video footage.
Unident. whale/dolphin		1	✓	N/A	N/A	N/A	✓	
Risso's dolphin	*Grampus griseus*	1	✓	N/A	N/A	N/A	N/A	
		2	×	N/A	N/A	N/A	N/A	Video footage skipped by reviewer.
False killer whale	*Pseudorca crassidens*	1	N/A	✓	✓	✓	N/A	
Leatherback sea turtle	*Dermochelys coriacea*	1	×	N/A	N/A	N/A	✓	Video footage skipped by reviewer.

Species	Scientific name	Observation #	Reviewer detection by speed 4×	8×	8×	8×	16×	Why not detected
Loggerhead sea turtle	*Caretta caretta*	1	✓	N/A	N/A	N/A	✓	
		2	✓	N/A	N/A	N/A	N/A	
		3	✓	N/A	N/A	N/A	N/A	
		4	✓	N/A	N/A	N/A	✓	
		5	✓	N/A	N/A	N/A	✓	
		6	✓	N/A	N/A	N/A	✓	
		7	✓	N/A	N/A	N/A	✓	
		8	✓	N/A	N/A	N/A	✓	
		9	✓	N/A	N/A	N/A	✓	
		10	✓	N/A	N/A	N/A	✓	
		11	✓	N/A	N/A	N/A	✓	
		12	N/A	N/A	N/A	N/A	✓	
		13	N/A	N/A	N/A	N/A	✓	
		14	✓	N/A	N/A	N/A	✓	
		15	✓	N/A	N/A	N/A	✓	
Olive ridley sea turtle 1	*Lepidochelys olivacea*	1	N/A	✓	✓	✓	N/A	
		2	N/A	N/A	N/A	N/A	✓	
Oceanic whitetip shark	*Carcharhinus longimanus*	1	×	N/A	N/A	N/A	×	Not possible to determine
		2	×	N/A	N/A	N/A	×	Not possible to determine
		3	×	N/A	N/A	N/A	×	Not possible to determine
		4	×	N/A	N/A	N/A	×	Not possible to determine
		5	✓	N/A	N/A	N/A	✓	Not possible to determine
		6	×	N/A	N/A	N/A	✓	Not possible to determine
		7	×	N/A	N/A	N/A	×	Not possible to determine

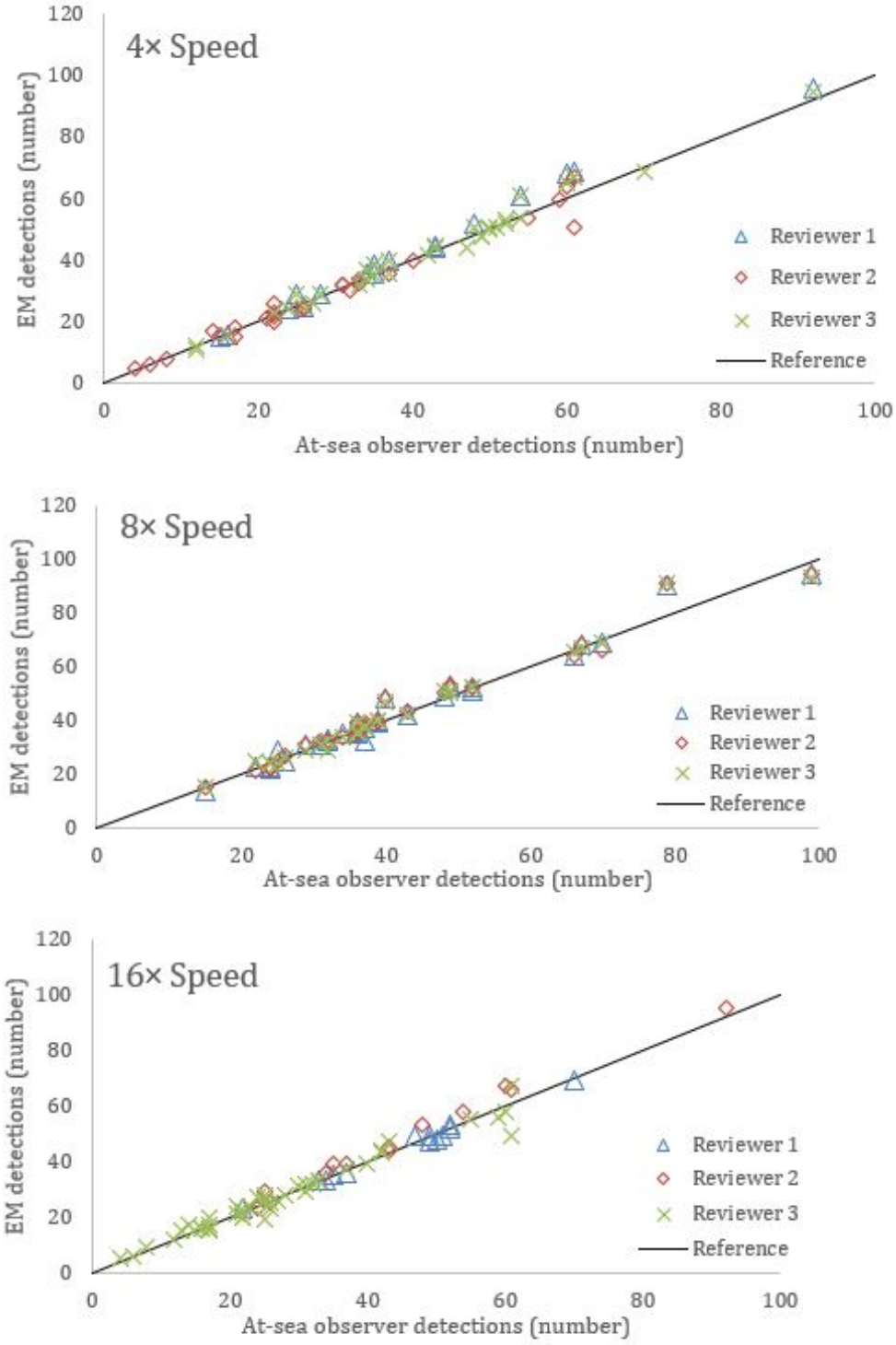

Figure 2. Comparison of retained catch detections by electronic monitoring (EM) reviewers and video review speeds. At-sea observer detections were assumed to have no bias and considered ground truth estimates to compare bias and precision by EM reviewers and speeds. Detections are shown by haul.

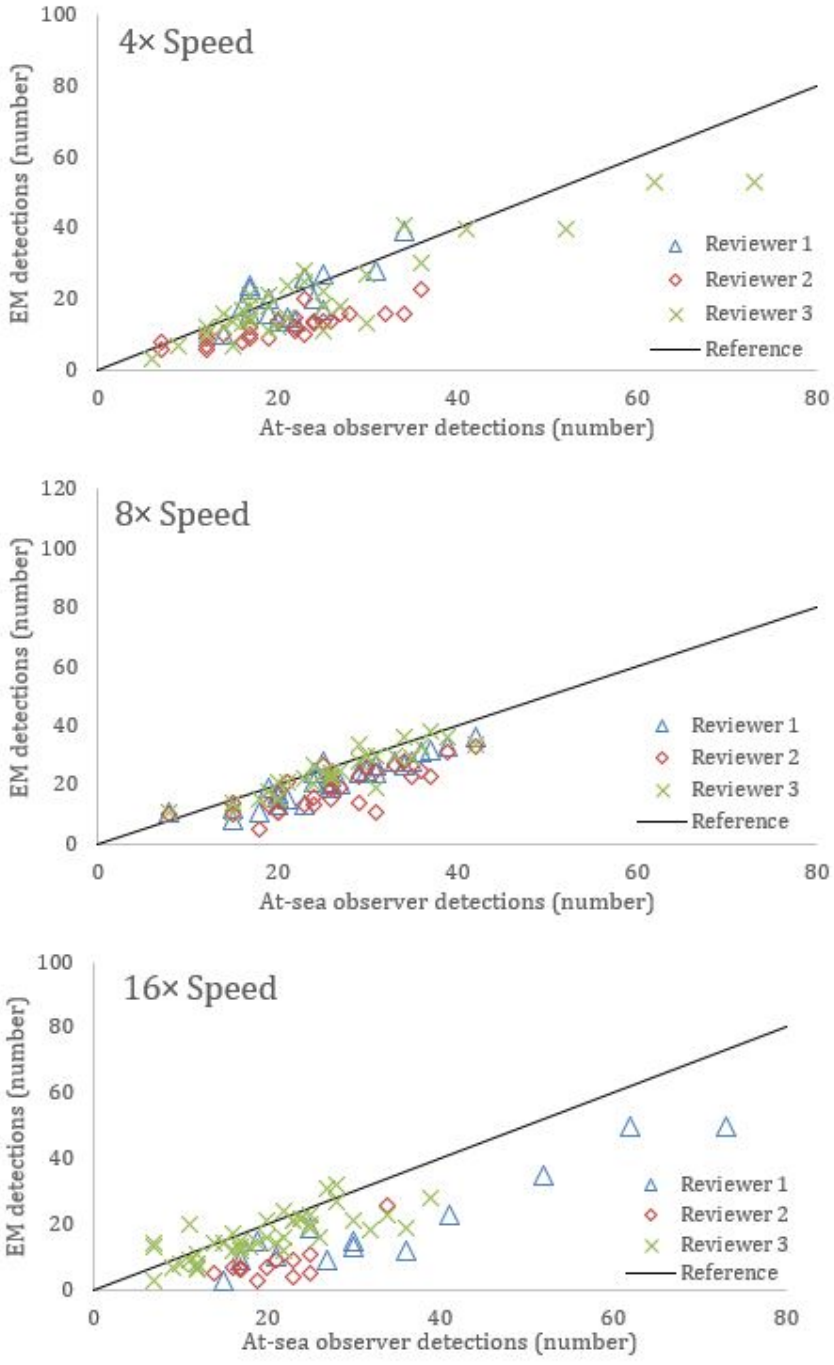

Figure 3. Comparison of discarded catch detections by electronic monitoring (EM) reviewers and video review speeds. At-sea observer detections were assumed to have no bias and considered ground truth estimates to compare bias and precision by EM reviewers and speeds. Detections are shown by haul.

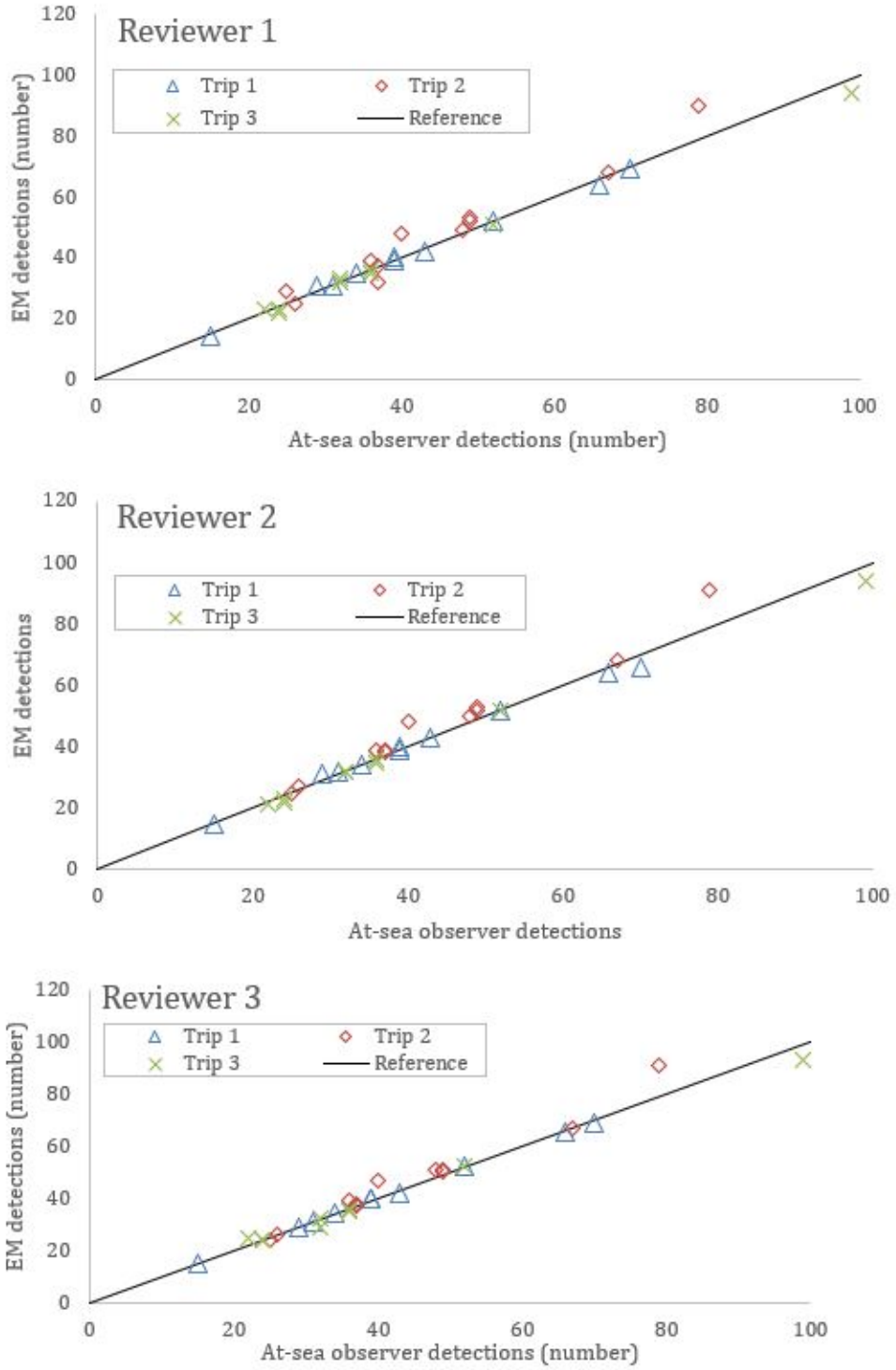

Figure 4. Comparison of retained catch detections by electronic monitoring (EM) trip for each reviewer at 8× video review speed. At-sea observer detections were assumed to have no bias and considered ground truth estimates to compare bias and precision by EM reviewers and speeds. Detections are shown by haul.

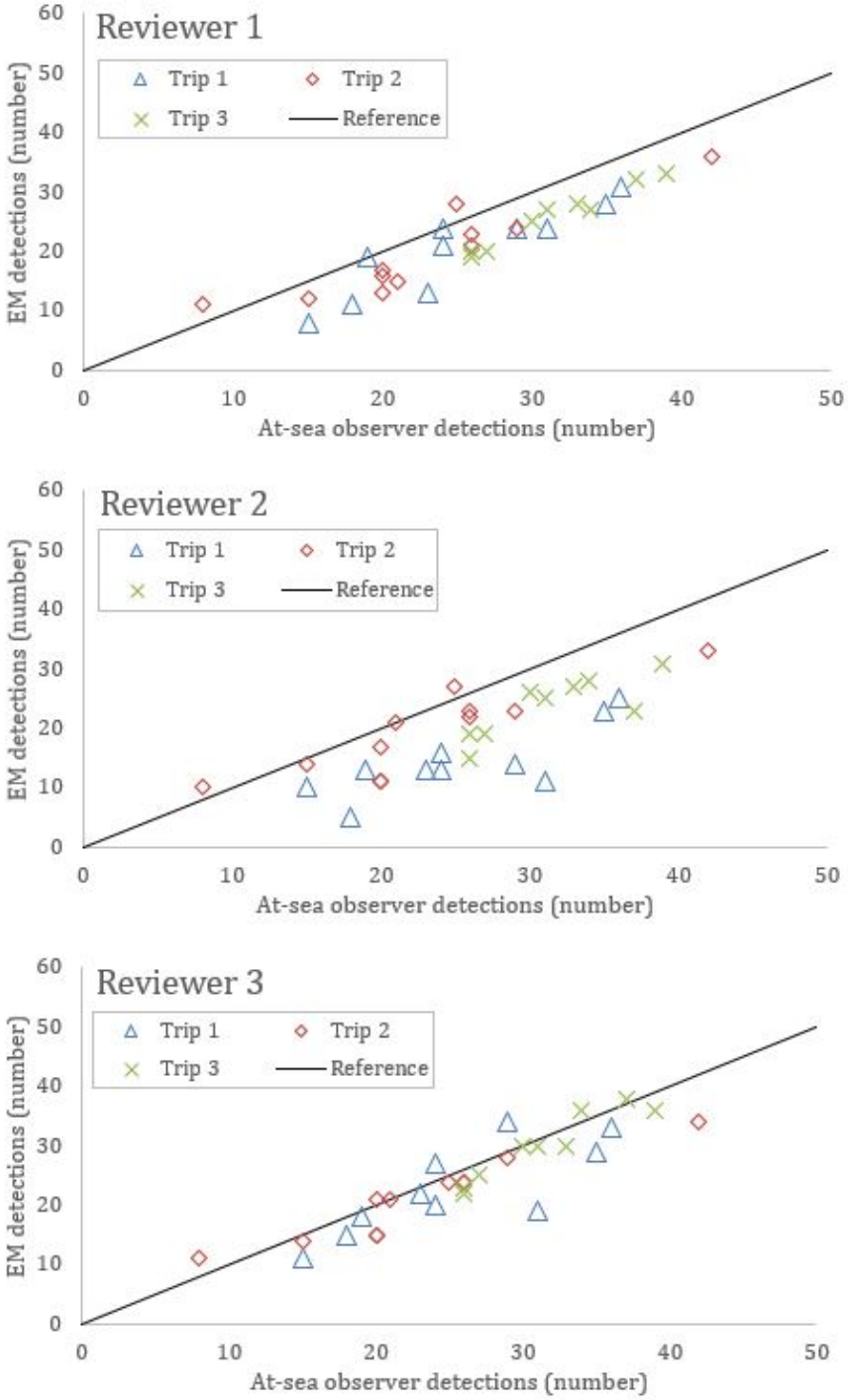

Figure 5. Comparison of discarded catch detections by electronic monitoring (EM) trip for each reviewer at 8× video review speed. At-sea observer detections were assumed to have no bias and considered ground estimates to compare bias and precision by EM reviewers and speeds. Detections are shown by haul.

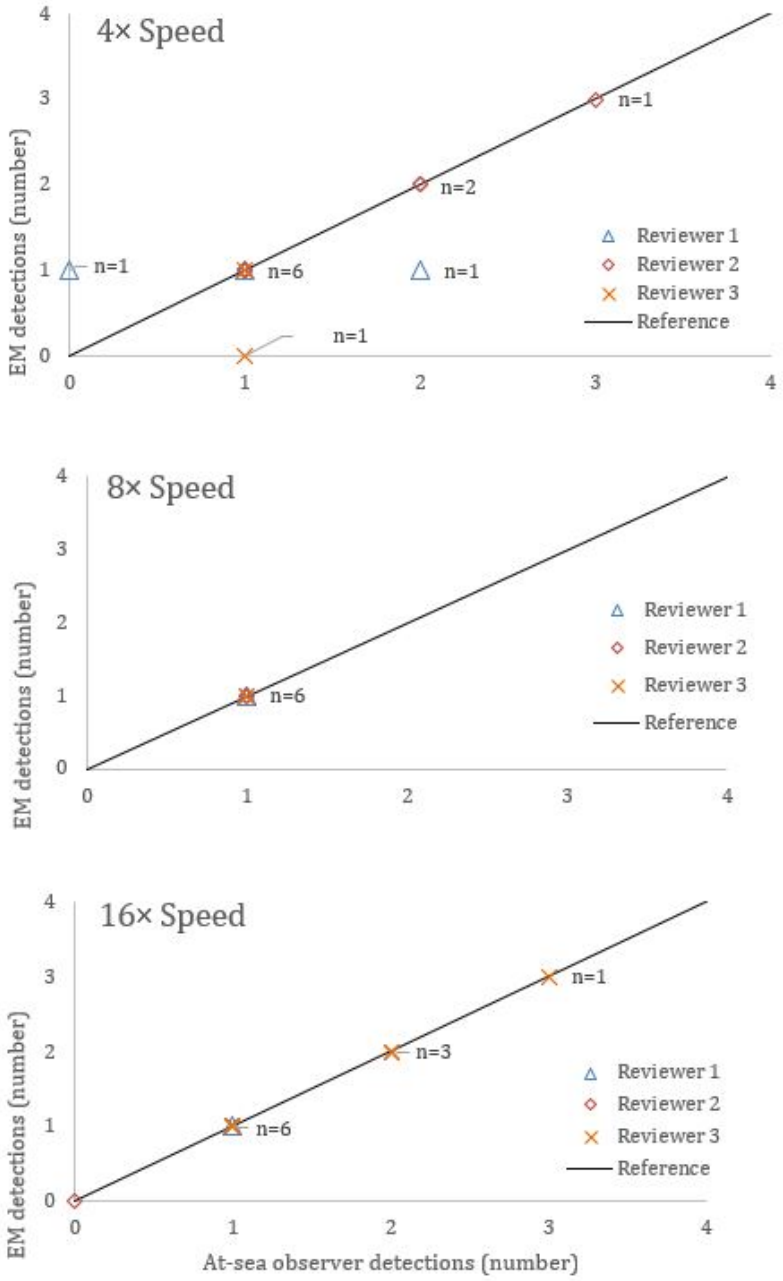

Figure 6. Comparison of protected species detections by electronic monitoring (EM) reviewers and video review speeds. At-sea observer detections were assumed to have no bias and considered ground truth estimates to compare bias and precision by EM reviewers and speeds. Observer detections were excluded if animal was not captured in video. Detections are shown by haul.