Fisheries New Zealand

Tini a Tangaroa

Video observation of the FMA 1 bottom longline fishery in 2020–21 and 2021–22

New Zealand Aquatic Environment and Biodiversity Report No. 340

D. A. J. Middleton,
E. R. Abraham

ISSN 1179-6480 (online)
ISBN 978-1-991308-37-5 (online)

August 2024

Te Kāwanatanga o Aotearoa
New Zealand Government

Requests for further copies should be directed to:

Fisheries Science Editor
Fisheries New Zealand
Ministry for Primary Industries
PO Box 2526
Wellington 6140
NEW ZEALAND

Email: Fisheries-Science.Editor@mpi.govt.nz
Telephone: 0800 00 83 33

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-and-resources/publications
http://fs.fish.govt.nz go to Document library/Research reports

Please cite this report as:

Middleton, D.A.J.; Abraham, E.R. (2024). Video observation of the FMA 1 bottom longline fishery in 2020–21 and 2021–22. *New Zealand Aquatic Environment and Biodiversity Report No. 340*. 49 p.

TABLE OF CONTENTS

Plain language summary

The 'petrel project' continued the use of on-board cameras to monitor seabird captures in the small-vessel bottom longline fishery off the north-east of New Zealand. Results demonstrated the value of the video observation data in providing data that were representative of the fishery, and the importance of assessing the variation between reviewers in their ability to make observations from the footage. Footage was reviewed to May 2022 and, for the first time, included reviewing footage from the winter months (June–October) in addition to the November–May period when black petrels and flesh-footed shearwaters are breeding in the area. Models using the video observation data estimated 40 black petrel captures, and 159 flesh-footed shearwater captures, in all bottom longline fishing in Fisheries Management Area 1 during the 2021–22 fishing year. The estimated captures of black petrels were lower than previous estimates, which were based on observed data only, but estimated captures of flesh-footed shearwaters were similar. Additional blind reviews were carried out of footage where captures occurred to allow estimation of reviewer skill. Of the five reviewers involved in the project, two reviewers detected over 90% of the seabird captures, but one reviewer detected fewer than 50% of the captures.

EXECUTIVE SUMMARY

Middleton, D.A.J.[1]**; Abraham, E.R.**[2] **(2024). Video observation of the FMA 1 bottom longline fishery in 2020–21 and 2021–22.**

New Zealand Aquatic Environment and Biodiversity Report No. 340**. 49 p.**

Video observation was used to monitor seabird captures in the Fisheries Management Area (FMA) 1 bottom-longline fishery. The observations were made during the 2020–21 and 2021–22 fishing years, from November 2020 to May 2022. This programme (typically referred to as the 'petrel project') extended previous video monitoring of this fishery that began during the 2016–17 fishing year. Existing camera deployments, on eight voluntarily participating vessels, were used to collect the footage. The participating vessels set 35% of the total hooks in FMA 1 bottom longline fisheries during 2020–21, and 40% of the total hooks in these fisheries during 2021–22. The statutory Electronic Reporting (ER) data were used to define haul periods, and hauls were randomly selected for video review. Any hauls that had fisher-reported captures were also selected for review. There were also a small number of hauls that were reviewed for other reasons. Overall, for the 2021 summer season (November 2020 to May 2021), 441 (12.1%) of bottom longline fishing events in FMA 1 were reviewed, and for the 2022 summer season 211 events (8%) were reviewed. Additionally, 252 (3.7%) of bottom longline fishing events occurring from June to October 2021 were reviewed.

If any seabird captures were found during the primary review, the capture events were passed to an expert reviewer for further review. There were a total of five reviewers reviewing footage from the 2020–21 fishing year. In addition, during the 2021 fishing year, any hauls with capture events were passed for secondary review from at least three other reviewers. This allowed variation in the skill of the reviewers at detecting seabird captures to be quantified. Based on expert review of the video footage, there were 176 seabird captures during 2021, and 21 during 2022. Around a third of the seabird captures were dead (26.7% of captures during 2021, and 66.7% of captures during 2022). The highest number of captures recorded from a single vessel was 71 captures. Flesh-footed shearwater (*Ardenna carneipes*) was the most frequently caught species (139 captures during 2021, and 19 captures during 2022), followed by black petrel (*Procellaria parkinsoni*; 35 captures during 2021, and 2 captures during 2022). A fluttering shearwater (*Puffinus gavia*) and a sooty shearwater (*Ardenna grisea*) were also captured, both in 2021. The secondary reviewing established that there was high variation in reviewer skill. Two reviewers detected over 90% of the seabird captures; however, one reviewer detected fewer than 50% of the captures.

A model was used to estimate the total captures of black petrels and flesh footed shearwaters using the video-review data. There were an estimated 40 (97.5% c.i.: 14 to 79) black petrel captures in all bottom longline fishing in FMA 1 during the 2021–22 fishing year, and an estimated 159 (97.5% c.i.: 56 to 392) flesh-footed shearwater captures in the same fisheries and period.

During the 2020–21 and 2021–22 fishing years, there were 257 seabird captures reported by fishers from bottom longline fishing within FMA 1. Of these captures, 218 were reported from vessels participating in the video monitoring trial and only 39 seabird captures were reported from other vessels. The rate of fisher reported seabird captures was close to ten times higher on vessels that were participating in the petrel project.

[1]Pisces Research, Wellington, New Zealand
[2]Dragonfly Data Science, Wellington, New Zealand

1. INTRODUCTION

Video observation has been used as a tool to monitor seabird captures in parts of the Fisheries Management Area (FMA) 1 bottom-longline fishery since the 2016–17 fishing year (Middleton & Guard 2021, Middleton & Abraham 2023). This programme (the 'petrel project') has aimed to gather more information on seabird bycatch in the area of high overlap between bottom longline fisheries and the distributions of black petrel (*Procellaria parkinsoni*) and flesh-footed shearwater (*Ardenna carneipes*). These two species are considered to be at risk of population impacts as a result of fisheries bycatch (Richard & Abraham 2020). Fisheries New Zealand projects PSB2020-10 and PRO2021-07 continued the petrel project in 2020–21 and 2021–22. This included two summer seasons (Nov–May), that have been the previous focus of video observation in FMA 1 because this approximates the breeding seasons of black petrel and flesh-footed shearwater, and additionally the winter (Jun–Oct) of 2021.

The winter season has previously received lower observational coverage because several of the seabird species of interest, including black petrels and flesh-footed shearwaters (see Peatman et al. 2023, tables 59 and 68), are absent from New Zealand at this time As a consequence, any interactions with species that are present during these months have been less well studied and the observational data from FMA 1 longline fisheries (both from video observation and human observers) has been unrepresentative of the annual effort.

This report summarises the coverage of the fishery by the video observation programme in 2020–21 and 2021–22, the results of reviewing the footage, and presents some initial modelling of the data using the full time-series of video observation data gathered since the 2016–17 fishing year. Middleton & Abraham (2023) noted that multiple reviews had identified between-reviewer variation in the detection of seabird captures in the video footage, and a more comprehensive approach to estimating between-reviewer differences was undertaken.

Two appendices consider specific aspects of video observation data: (i) the nature of the data and how it might best be managed, and (ii) frameworks for reusing the footage in future projects, in particular for the development of computer vision techniques for analysis of the footage.

2. METHODS

The Electronic Monitoring (EM) systems used for footage collection were installed at the start of the 2020 season, as described by Middleton & Abraham (2023). During 2020–21 and 2021–22 similar footage retrieval (via USB devices) and cataloguing approaches were employed. Footage was collected under the PSB2020-10 and PRO2021-07 projects from 1 November 2020 to 31 October 2022, with reviewing conducted for footage collected between 1 November 2020 and 31 May 2022. Some footage was also collected for winter 2020 (i.e., from May to October 2020, between the previous project, PSB2019-06, and the first of the current projects, PSB2020-10) and this is also reported here.

The key methodological change in the work reported here is that slices of footage defined for reviewing were based on hauling events, rather than complete port-port trips as previously implemented (Middleton & Abraham 2023). Hauling events were defined using the statutory Electronic Reporting (ER) data. Because the EM systems installed in 2020 were generally only operational during hauling (Middleton & Abraham 2023), this approach had little impact on the amount of footage from a trip that was reviewed (i.e., there was no footage collected other than within hauling periods).

The principal advantage of the haul-based approach to reviewing was that the randomised selection of slices for review was able to be implemented at the fishing event level, rather than the trip level. Hauls were selected for review using a vessel-month stratification; once all the footage from a vessel in a given month had been retrieved and catalogued, then that vessel-month's activity was processed for review.

A haul was considered eligible for review if footage was available for at least 60% of the haul period defined using the ER data. For the summer 2020–21 and winter 2021 seasons, a sampling fraction of 45% was employed; that is, a random selection of the hauls eligible for review in a vessel-month stratum was selected, with the aim of reviewing 45% of the hauls undertaken by the vessel that month. If footage was not available from sufficient hauls, then all eligible hauls were selected for review. The random sample of hauls selected for review was supplemented with any hauls where fishers reported seabird captures in their ER data and that were not previous selected (i.e., by the randomised selection). For the summer 2021–22 season, the sampling fraction was reduced to 35% of hauls undertaken, with the aim of better matching the review activity to the contracted levels of reviewing.

2.1 Primary and expert reviewing

Other than the use of haul-based slices, rather than trip-based slices, review protocols were the same as those used in previous seasons (Middleton & Abraham 2023). All slices selected for review were initially reviewed in their entirety by one of a team of primary reviewers, with the key task being the identification of any seabird captures. Where seabird captures were identified in a Primary review, new slices were defined for Expert review. The Expert review slices extended for five minutes before and after the capture events identified in the Primary reviews, after first amalgamating any capture events that were within five minutes of each other (to avoid Expert reviews with overlapping periods).

The footage for one haul selected for review from the summer of 2021 was found to be corrupt. A further sixteen hauls selected for review in winter 2021 and the summer 2022 season had their primary review abandoned due to a bug in the review software that prevented events being recorded if GPS data were unavailable. These hauls are treated here as if they were not reviewable.

2.2 Secondary reviewing

Middleton & Abraham (2023) noted that duplicate reviewing had revealed higher than anticipated differences in the detection of seabird captures by the Primary review process. In previous seasons, this duplicate (Secondary) reviewing was undertaken by randomly selecting from all reviewed slices – including those where no seabird captures were identified by the Primary review. While verification of a lack of captures in these slices was worthwhile, events without captures do not contribute data that are useful for assessing between-reviewer differences in their ability to detect seabird captures in the footage. As a result, Secondary reviewing for the summer 2020-21 season was implemented using a more comprehensive cross-over design. Any slice where seabird captures were detected during the Primary review was sent for review by the four other primary reviewers. Secondary reviews were blind; that is, the reviewers were generally unaware of the review scheduling process, and the review software did not explicitly identify whether the reviews available to a reviewer were Primary or Secondary reviews.

2.3 Fishery data

Consistent with previous seasons, coverage of the video observation programme was assessed within the sampling frame of the FMA 1 bottom longline fishery, defined as fishing effort reported to MPI using the bottom longline method and with a set start position within the FMA 1 area. Fishing events meeting these criteria, and occurring between 1 November 2020 and 31 May 2022, were included. Individual vessels are referred to in this report using anonymised vessel identifiers (three letter codes).

2.4 Analysis datasets

A linked dataset that included data generated from the current programme, and earlier seasons (Middleton & Guard 2021, Middleton & Abraham 2023), was generated. This spanned the period from 1 November 2016 to 31 May 2022, noting that any reviewing of effort between June and October was largely restricted to the 2021 fishing year. The data were compiled in a format intended to facilitate linking to the Protected Species Captures (PSC) database. Seabird capture data were based on the results of the Expert reviews.

A further dataset was compiled based on the Secondary reviewing. This contained information on all seabird captures identified in the summer 2020–21 season, whether or not they were identified by each of the five Primary reviewers, and whether they were confirmed as a capture by the Expert review.

2.5 Estimated captures of black petrel and flesh footed shearwater

In order to place the video observations in context, generalised linear models (GLMs) were used to estimate black petrel and flesh-footed shearwater captures in FMA 1 bottom longline fisheries. The models followed the same method used previously (Middleton & Abraham 2023), with the exception that a separate model was fitted to each of the two species, and more resolution was included in the seasonal effect. The input data included video observations from the 2017 to the 2022 fishing year (a 6 year period), and included winter observations from the 2021 fishing year.

For each of the two species, the number of captures identified by expert review, on each set, was assumed to be drawn from a negative-binomial distribution. The logarithm of the mean of the distribution was represented as a sum of covariates (Table 1), with the coefficients of the covariates being estimated through statistical model fitting. The model fitting was carried out within a Bayesian framework using the software BRMS (Bürkner 2017). The model formula (in BRMS notation) was:

```
expert_captures | rate(observed_hooks)  ~ target +  s(month, bs='cc', k=12) +
          (1|year) + (1|vessel_alias) + (1|stat_area)
```

The models were initialised with unit normal priors on the coefficients of the covariates and on the standard deviation of the random effects. The default Gamma(0.01, 0.01) prior was used for the shape of the negative binomial distribution, a Normal(-10, 3) distribution was used for the prior of the intercept, and a default Student-*t*(3, 0, 2.5) distribution was used for the prior of the spline variance parameter. The models were run with four chains, for 1000 warm-up iterations and 1000 sampling iterations, so providing a total of 4000 samples of the posterior distribution. Convergence was assessed using the Gelman-Rubin $\hat{R}$ statistic (Gelman & Rubin 1992), which compares within chain and between chain variance. An $\hat{R}$ value of close to one indicates convergence.

The model was applied to a linked dataset, which had each video observation associated with a fisher-reported fishing event. The model was fitted to the fishing events that had been video-reviewed (with the hooks observed being the total hooks on the fishing event, scaled by the proportion of the haul that was observed). In the 2021 and 2022 fishing years, the method used for selecting events for review was recorded (either as ‘at random’, ‘fisher reported’, or ‘ad hoc’). All video-observed events from 2017 to 2020 were included in the model, however the video-observations from 2021 and 2022 were restricted to events that had been selected at random (excluding 44 events selected to be reviewed because of fisher-reported captures, and 22 events selected for ad hoc reasons). This helped maintain the representativeness of the data modelled. The data set used for fitting the model included 3149 fishing events (with 8.59M video observed hooks).

The fitted model was applied to all fisher reported effort data from the 2017 to 2022 fishing years, to estimate total captures in FMA 1 bottom longline fisheries over that period. When estimating the total captures, the model was applied to the number of hooks on each set that were not observed, and the observed captures were then added to the estimates. The model was also used to estimate captures on the observed hooks to allow the model fit to be investigated.

Table 1: Parameters used in estimating seabird captures from video observation data. For each parameter the table gives the representation in BRMS (Bürkner 2017) notation, and a description of how the parameter was defined.

Parameter	Representation	Description
Captures	`expert_captures`	The number of captures on each set confirmed by expert review of the footage. Includes both live and dead captures. This was the response variable of the GLM.
Observed hooks	`rate(observed_hooks)`	The number of hooks reviewed on each set, derived from the fisher-reported total hooks set. If the number of hours of video observation was less than the number of hauling hours, then the number of observed hooks was the number of hooks set, multiplied by the ratio of the video reviewed hours to the total hauling hours. The number of observed hooks is included in the model to normalise the mean of the negative binomial distribution, on the assumption that the number of seabirds caught is proportional to the number of hooks set.
Target species	`target`	The fisher-declared target species of the set, either snapper (SNA) or other species (OTH).
Month	`s(month, bs='cc', k=12)`	The month of the set, as an integer. The month is represented in the model as a cyclic spline, with 12 knots. The spline smoothly joins December (month 12) and January (month 1), with some December months (value 12) set to have a value of 0.
Fishing year	`(1|year)`	The fishing year of the set (the fishing year runs from October 1 to September 30), included in the model as a random effect.
Vessel alias	`(1|vessel_alias)`	A unique code for each fishing vessel, included in the model as a random effect.
Statistical area	`(1|stat_area)`	The General Statistical Area of the start of the set, included in the model as a random effect.

3. RESULTS

3.1 The FMA 1 bottom longline fishery

The bottom longline fishery in FMA 1 operates year-round (Figure 1), with similar levels of effort in the summer and winter seasons. An overall reduction in effort in the fishery was apparent in 2022. Snapper target effort dominates the fishery (Figure 2). There have typically been 30 to 40 vessels active in the fishery (Figure 3), but the reduced effort in 2022 has been associated with a reduction in the fleet. Less than 30 vessels operated in most months in the 2022 fishing year.

The EM programme in 2020–21 involved eight vessels, with the same eight vessels also participating in 2021–22. Vessels in the project fleet undertook 25% of the bottom longline sets in FMA 1 in the November to May period in the 2021 fishing year and 28% in the 2022 fishing year (Figure 4a). Expressed in terms of number of hooks set, this represents 35% of effort in the 2021 season and 40% in 2022 (Figure 4b).

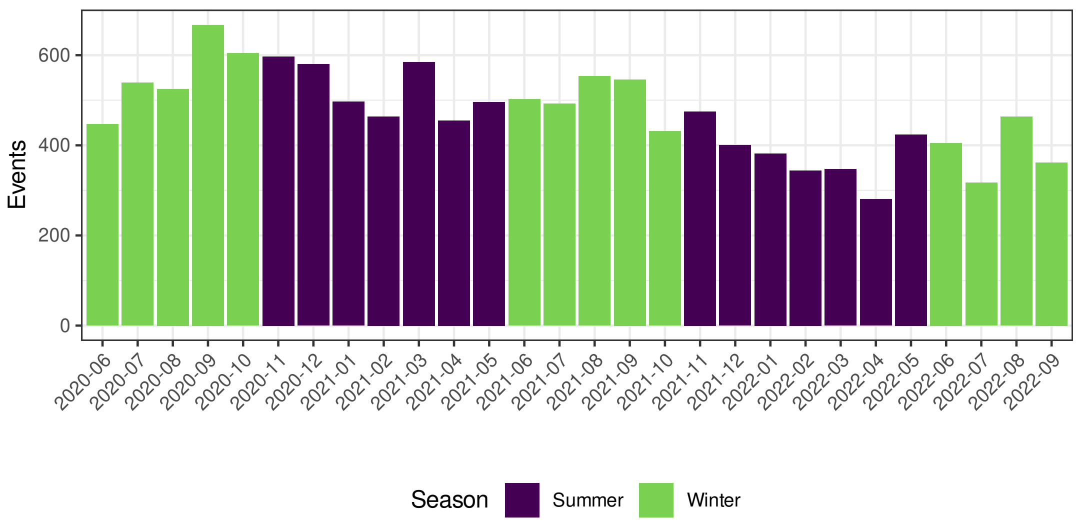


Figure 1: Bottom longline fishing events in FMA 1 in winter 2020, and the 2021 and 2022 fishing years, by month.

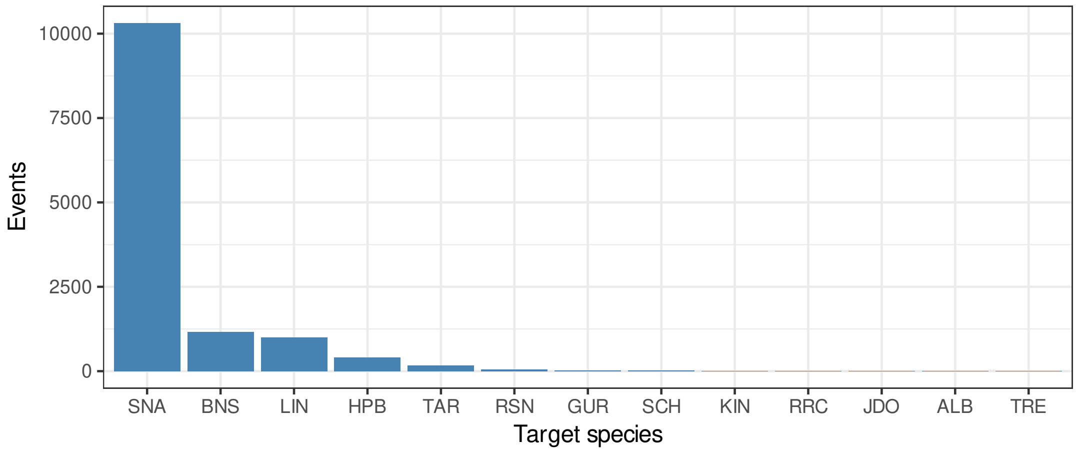


Figure 2: Bottom longline fishing events in FMA 1 in the 2021 and 2022 fishing years by target species: SNA (snapper, *Chrysophrys auratus*), BNS (bluenose, *Hyperoglyphe antarctica*), LIN (ling, *Genypterus blacodes*), HPB (hapuku and bass, *Polyprion oxygeneios*, *Polyprion americanus*), TAR (tarakihi, *Nemadactylus macropterus*, *Nemadactylus* sp.), RSN (red snapper, *Centroberyx affinis*), GUR (gurnard, *Chelidonichthys kumu*), SCH (school shark, *Galeorhinus galeus*), KIN (kingfish, *Seriola lalandi*), RRC (red scorpion fish, *Scorpaena cardinalis*, *Scorpaena papillosus*), JDO (John dory, *Zeus faber*), ALB (albacore, *Thunnus alalunga*), TRE (trevally, *Pseudocaranx georgianus*).

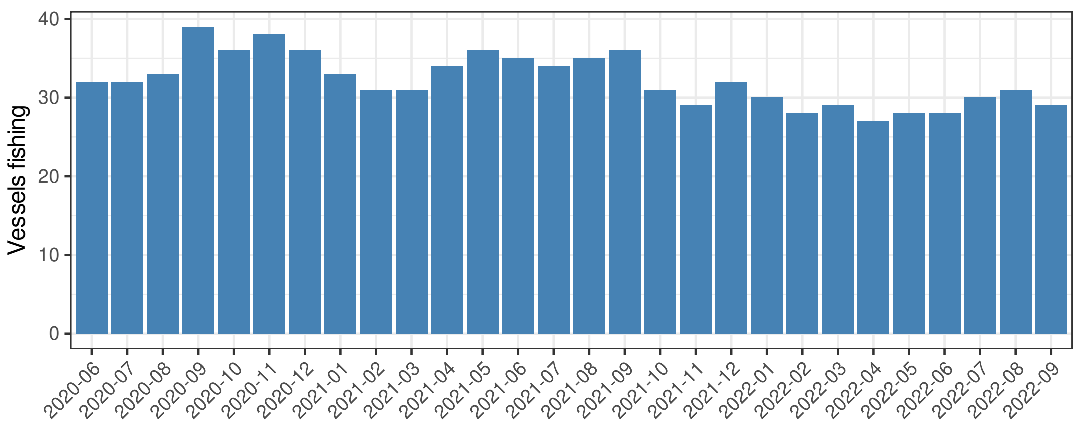


Figure 3: Bottom longline fishing vessels in FMA 1 in the 2021 and 2022 fishing years by month.

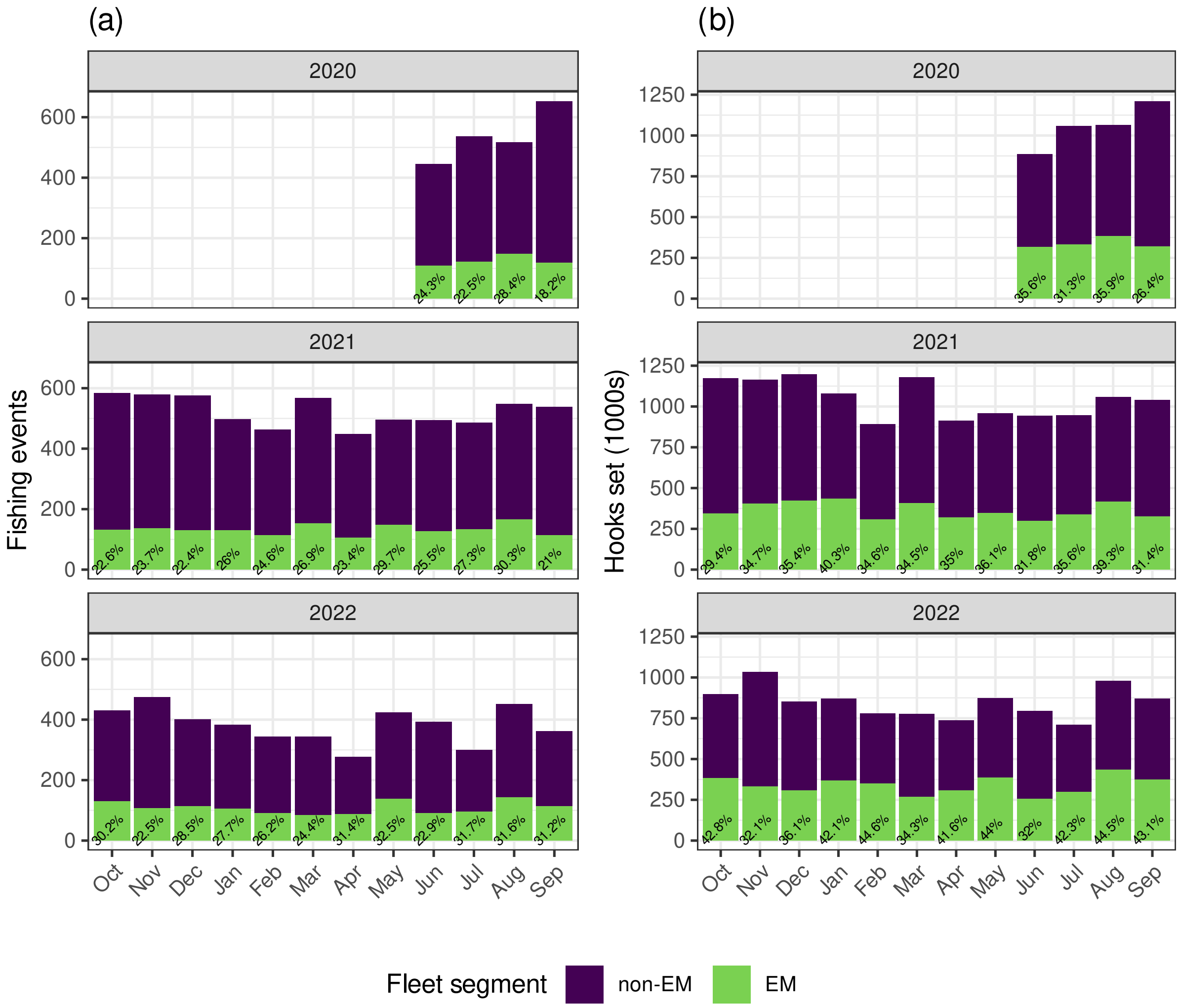


Figure 4: FMA 1 bottom longline effort in winter 2020, and the 2021 and 2022 fishing years, in terms of (a) fishing events and (b) hooks set, categorised according to whether vessels are part of the EM fleet. Note that this does not necessarily imply that EM footage was collected.

3.2 Footage collection

3.2.1 Operational aspects

Middleton & Abraham (2023) noted that the EM systems installed in 2020 had a number of operational issues. In particular, the domes covering the camera lenses developed hairline cracks. As a result, all cameras were replaced at the start of the 2021 season with cameras that had a flat cover over the lens. This resulted in a small reduction in the field of view. Additionally, to further address the fact that the noise generated by the EM systems required that they were only powered up during hauling, several of the vessels were fitted with switches to automatically power up the cameras, so reducing reliance on manual powering up of the EM systems by the crew.

Retrieval of footage using USB drives was generally successful, although on some occasions footage was not received in date order (i.e., oldest files first). There were also some technical issues with the process of cataloguing some footage, once received. The result of these issues was that footage from some vessel-months was not available for review as promptly as intended.

3.2.2 Footage collection achieved

For the vessels in the project's EM fleet, Figure 5 indicates the monthly effort undertaken and whether EM footage was collected. In this case, an event was considered to have footage if EM footage was recorded and catalogued for at least 80% of the haul period. This is consistent with previous reports but, as indicated in the Methods, a lower threshold of 60% was used to determine whether hauls were eligible for review. Occasions where no footage was collected on a vessel for a period were associated with problems with the EM systems; periods where there is a mix of events with and without footage generally relate to failures to power the EM system during the haul period.

The distributions of the proportion of the haul period for which footage was recorded (Figure 6) tend to indicate that footage either was or was not recorded from an event; i.e., there are few occasions where cameras were powered for only a part (25–75%) of the haul. For hauls with footage, some mismatch between the haul period recorded in the ER data and the footage available is expected because the ER data does not exclude non-hauling periods such as lunch breaks or vessel positioning, whereas the EM systems may power down in such breaks.

Overall footage availability for the FMA 1 bottom longline fleet (Table 2) during the Nov–May season indicates that the EM programme typically obtained footage from 15% to 27% of hauling hours. Monthly coverage of hauling hours by the EM fleet (Table 3) ranged from 50% to 85% of hauling hours with between-vessel and between-month variation apparent (Table 4).

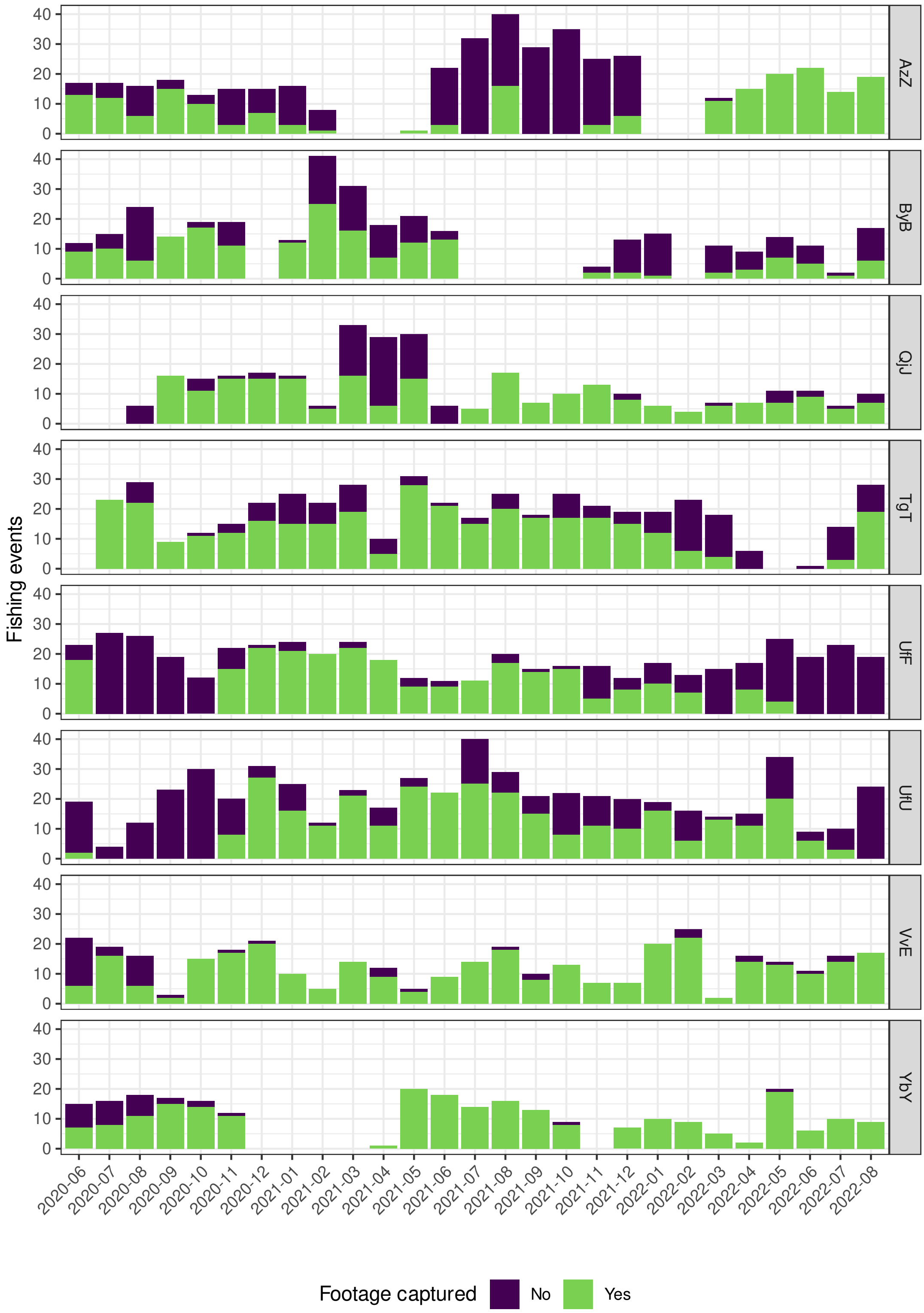


Figure 5: Fishing effort (fishing events in the statutory catch and effort data) by project vessel (anonymised identifier) and month, categorised according to whether footage was received from the vessel and catalogued in the review system database for more than 80% of the hauling period.

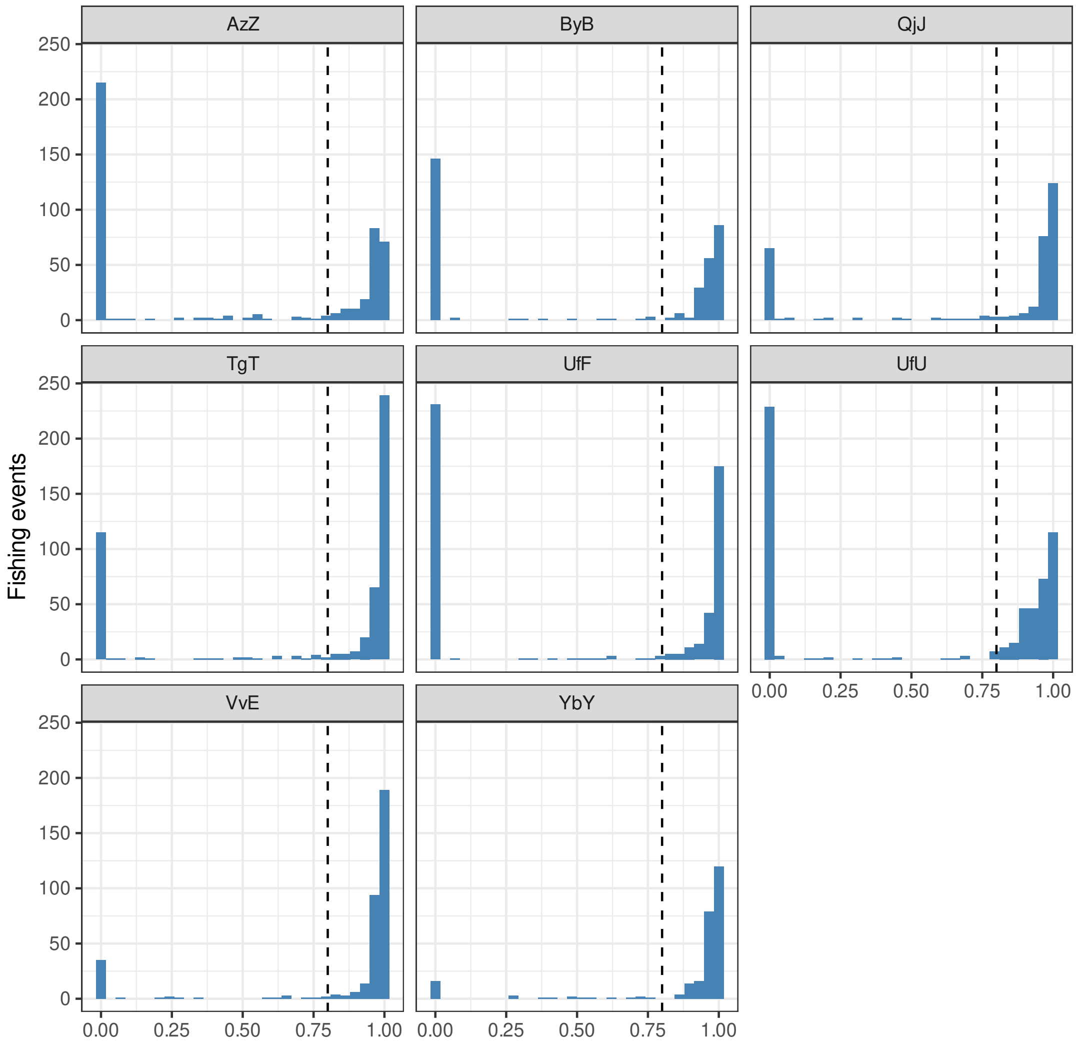


Figure 6: Proportion of the haul period with footage, by project vessel (anonymised identifier). The dashed vertical line indicates the proportion of the haul period above which an event was considered to have sufficient footage for the purposes of reporting coverage achieved.

Table 2: Monthly coverage (in terms of footage captured) for hauling periods of the FMA 1 BLL fleet in the 2021 and 2022 seasons.

		Hauling hours		
	Month	Total	With footage	Coverage (%)
2021				
	Nov	1993.6	352.0	17.7
	Dec	1963.2	493.2	25.1
	Jan	1841.7	434.0	23.6
	Feb	1539.8	397.0	25.8
	Mar	2087.9	488.8	23.4
	Apr	1655.9	287.3	17.4
	May	1532.1	410.3	26.8
2022				
	Nov	1681.9	265.2	15.8
	Dec	1414.9	301.9	21.3
	Jan	1452.3	383.1	26.4
	Feb	1356.2	319.2	23.5
	Mar	1327.9	205.4	15.5
	Apr	1229.7	332.7	27.1
	May	1444.3	374.5	25.9

Table 3: Monthly coverage (in terms of footage captured) for hauling periods of the EM fleet in the 2021 and 2022 seasons.

		Hauling hours		
	Month	Total	With footage	Coverage (%)
2021				
	Nov	535.9	352.0	65.7
	Dec	580.7	493.2	84.9
	Jan	600.8	434.0	72.2
	Feb	505.3	397.0	78.6
	Mar	628.7	488.8	77.7
	Apr	458.6	287.3	62.7
	May	508.1	410.3	80.8
2022				
	Nov	485.2	265.2	54.7
	Dec	499.0	301.9	60.5
	Jan	512.7	383.1	74.7
	Feb	518.1	319.2	61.6
	Mar	408.6	205.4	50.3
	Apr	466.9	332.7	71.3
	May	589.8	374.5	63.5

Table 4: Monthly coverage (percentage of hauling hours with footage) for individual vessels, referred to by anonymised identifiers, in the EM fleet in the 2021 and 2022 seasons.

	Vessel	Nov	Dec	Jan	Feb	Mar	Apr	May
2021								
	AzZ	20.8	53.5	19.0	27.4			
	ByB	58.3		96.0	68.4	53.6	30.6	72.9
	QjJ	94.0	85.8	93.0	89.5	60.7	40.8	42.1
	TgT	81.6	74.6	68.9	68.7	79.0	63.9	93.0
	UfF	76.5	99.5	91.3	98.9	94.3	98.2	87.9
	UfU	35.2	91.8	58.3	89.6	89.5	69.1	89.5
	VvE	93.5	91.8	97.2	96.1	95.4	74.9	88.3
	YbY	90.3					100.0	98.6
2022								
	AzZ	9.9	36.2			88.9	97.8	97.0
	ByB	43.3	17.7	5.3		25.7	33.6	51.6
	QjJ	99.9	83.3	99.8	99.8	90.9	96.3	72.5
	TgT	72.8	80.7	64.7	34.0	23.2	0.0	
	UfF	57.2	63.3	63.7	50.3	0.6	52.8	20.7
	UfU	49.1	52.3	85.5	38.0	87.5	72.1	57.5
	VvE	97.1	98.9	96.7	92.5	99.4	89.6	89.6
	YbY		95.8	94.6	96.5	94.5	98.6	96.5

3.3 Overall coverage

Figure 7 provides a graphical summary of fishing effort, fisher-reporting of seabird captures, and footage capture for the duration of the project.

3.3.1 Fleet level coverage

In 2021, 3647 bottom longline fishing events took place in FMA 1 during the November–May period, and 645 (17.7%) had footage for at least 80% of the hauling period.

For the November–May season in 2022, 2638 bottom longline fishing events took place in FMA 1 and 439 (16.6%) had footage.

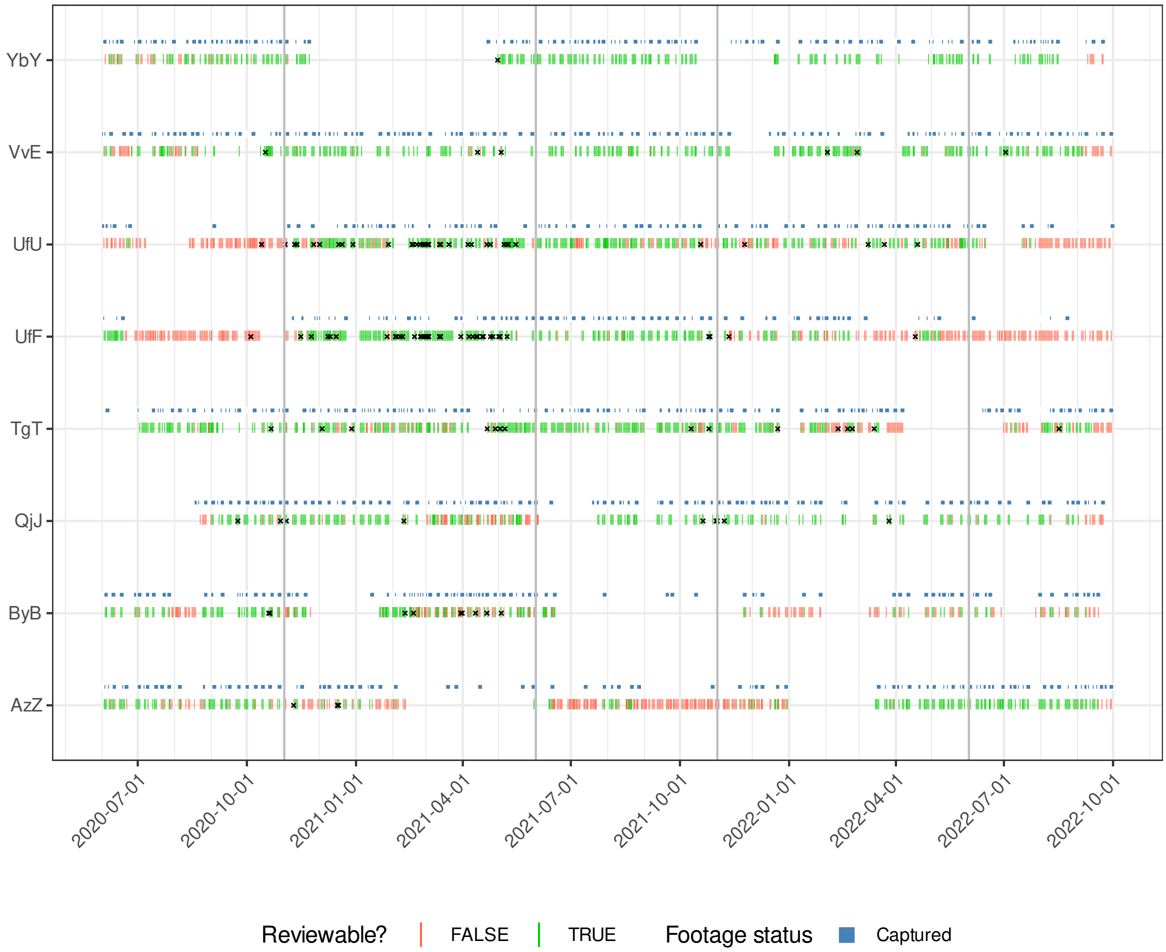


Figure 7: Bottom longline fishing events in FMA 1 (vertical lines), fisher-reported captures (crosses), and footage for the EM fleet, with anonymised vessel identifiers. The heavier vertical lines mark the boundaries of the period from 1 November to 31 May which has been the main focus of the programme. Events are considered reviewable if footage was captured for at least 60% of the haul period.

3.3.2 Video observation coverage

Actual coverage by the video observation process is dependent on the amount of footage that is reviewed to generate EM data; 904 hauling events were reviewed, from the subset of events that had footage available for at least 60% of the haul.

For the 2021 summer season, 441 (12.1%) of bottom longline fishing events in FMA 1 were reviewed; in the 2022 summer season 211 events (8%) were reviewed. Additionally, 252 (3.7%) of bottom longline fishing events occurring from Jun–Oct 2021 were reviewed.

Expressed in terms of coverage of hooks set to target snapper, video observation coverage in the 2021 fishing year was 16.2% of effort and in 2022 was 8.9% of effort.

3.4 Representativeness of coverage

In addition to assessing the volume of observational data available, it is useful to consider the extent to which this is representative of the fishery in terms of temporal and spatial patterns and the different target fishing activity.

3.4.1 Footage

The footage collected during the project was broadly representative of the FMA 1 bottom longline fishery in terms of seasonal pattern (Figure 8), but Statistical Areas 003, 005 and 009 were overrepresented in the spatial distribution while 002, 004, 007 and 008 were under-represented (Figure 9). The footage was dominated by snapper target fishing, with some minor target species also represented (Figure 10). However, no ling target fishing, which comprised the second most important bottom longline target species in FMA 1 during the period of the project, was represented.

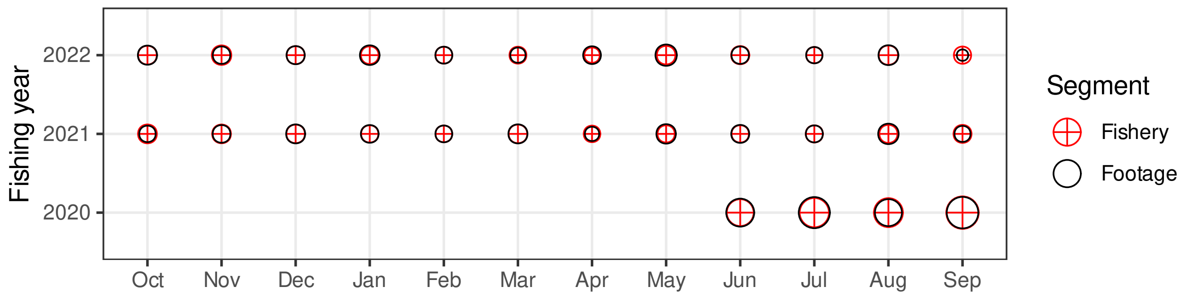


Figure 8: Footage collection representativeness by month and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and footage proportions in a month will be the same.

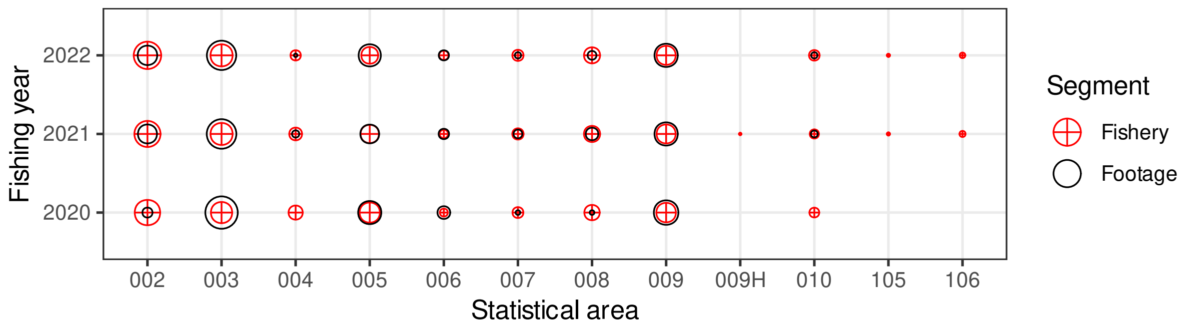


Figure 9: Footage collection representativeness by statistical area and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and footage proportions in an area will be the same.

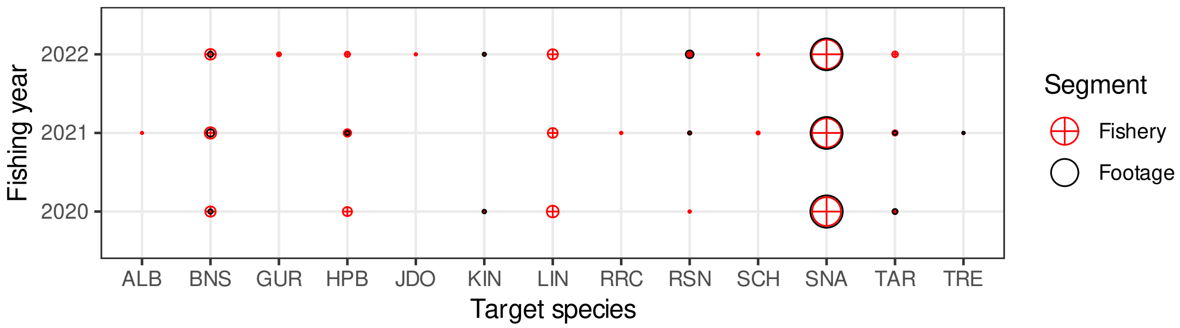


Figure 10: Footage collection representativeness by target species and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and footage proportions for each target species will be the same. Species codes are defined in Figure 2.

3.4.2 Reviewed footage

Footage reviewed was representative of monthly fishing patterns from November 2021 to June 2022 (Figure 11). Statistical Area 003 was over-represented in the reviewed footage, with the wider fishery having proportionally more effort in Statistical Area 002, in particular (Figure 12). As with footage collection, the reviewed footage was dominated by snapper target effort with some minor targets represented (Figure 13).

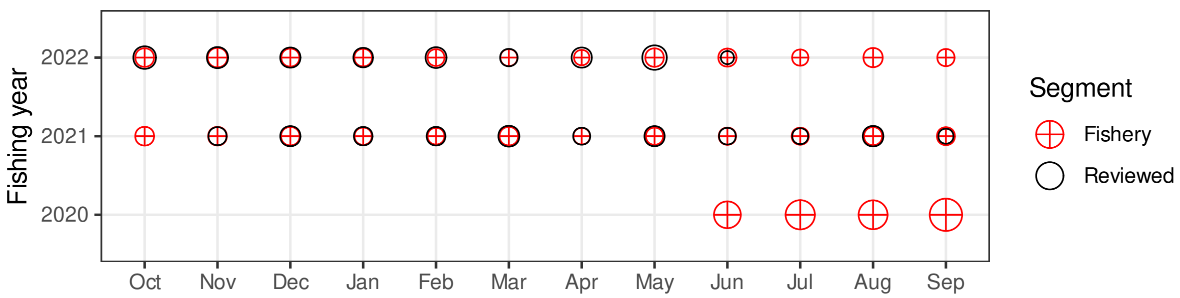


Figure 11: Reviewing representativeness by month and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and reviewing proportions in a month will be the same.

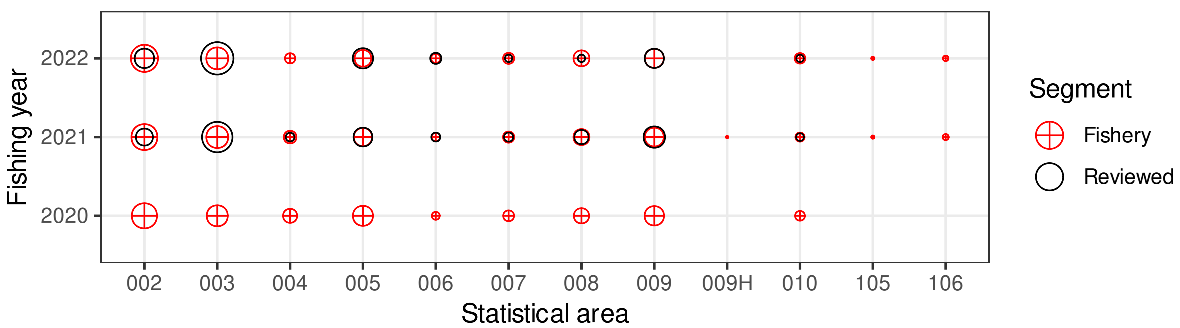


Figure 12: Reviewing representativeness by statistical area and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and reviewing proportions in each area will be the same.

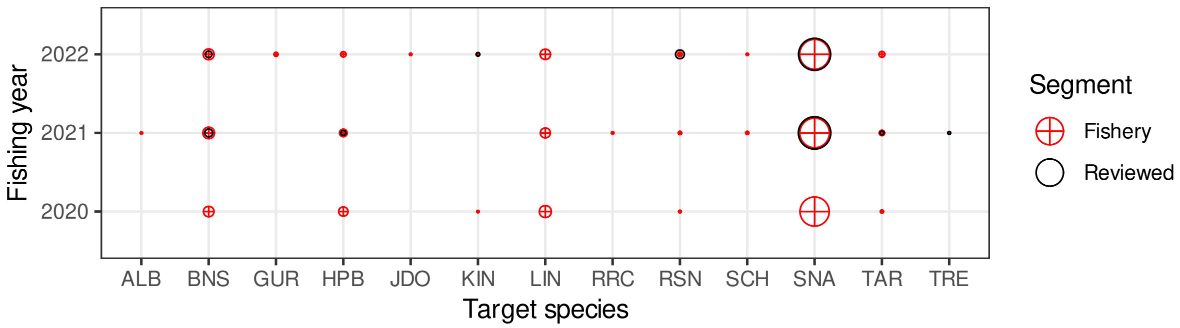


Figure 13: Reviewing representativeness by target species and fishing year. The area of the circles represents the proportion of the effort in the segment, by year. If coverage is completely representative, then the fishery and reviewing proportions for each target species will be the same. Species codes are defined in Figure 2.

3.5 Reviewing results

3.5.1 Assessment of footage quality

Assessment of camera deployment status (Figure 14) indicated that the vessels in the EM fleet were generally ensuring that the camera booms were deployed while fishing. Footage quality varied between vessels and over time (Figure 15, Table 5). Glare and water spots on the camera lens were the key reasons for reductions in footage quality (Table 6).

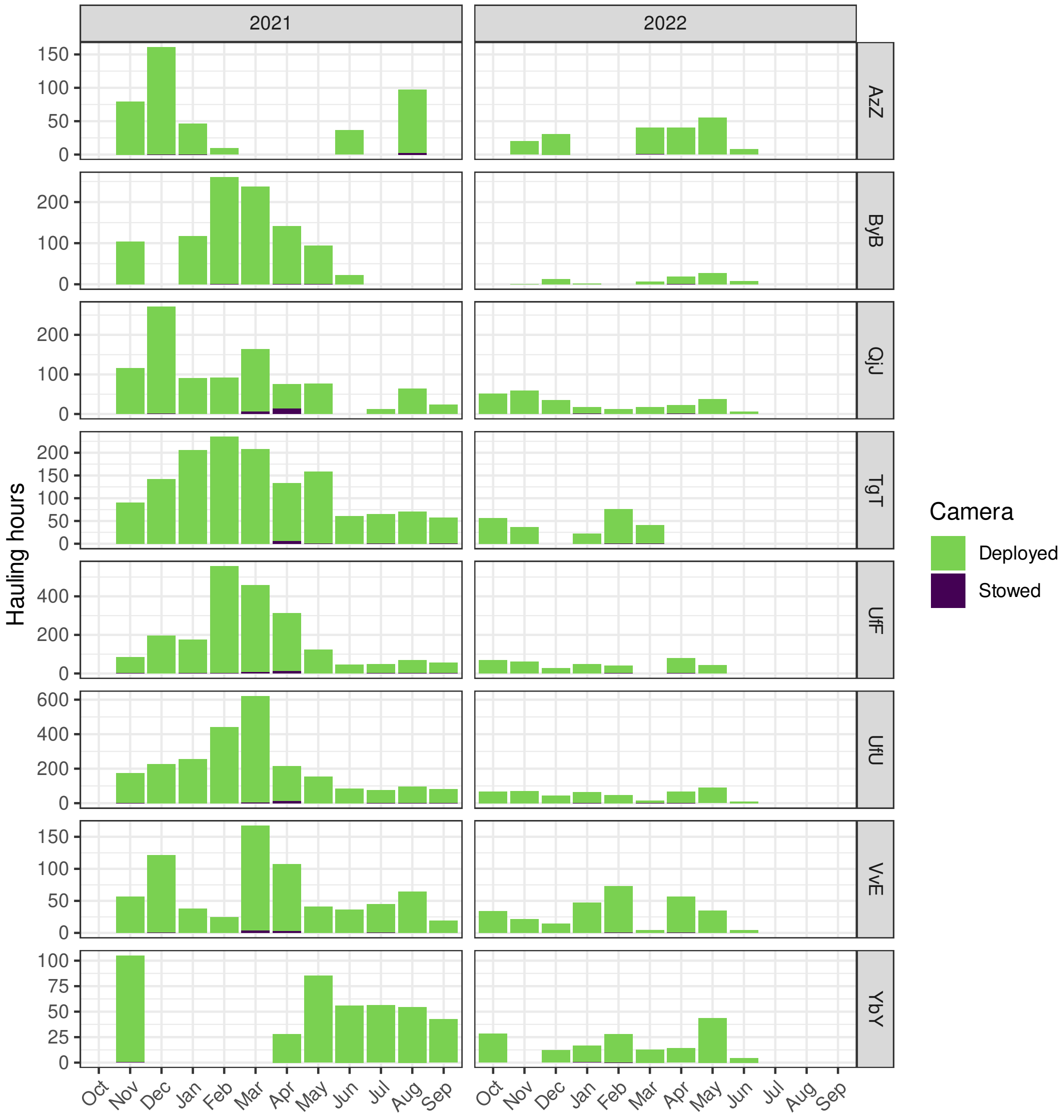


Figure 14: Hauling hours per vessel (anonymised identifier) per month classified according to whether the camera was stowed or deployed, based on the first review of a slice.

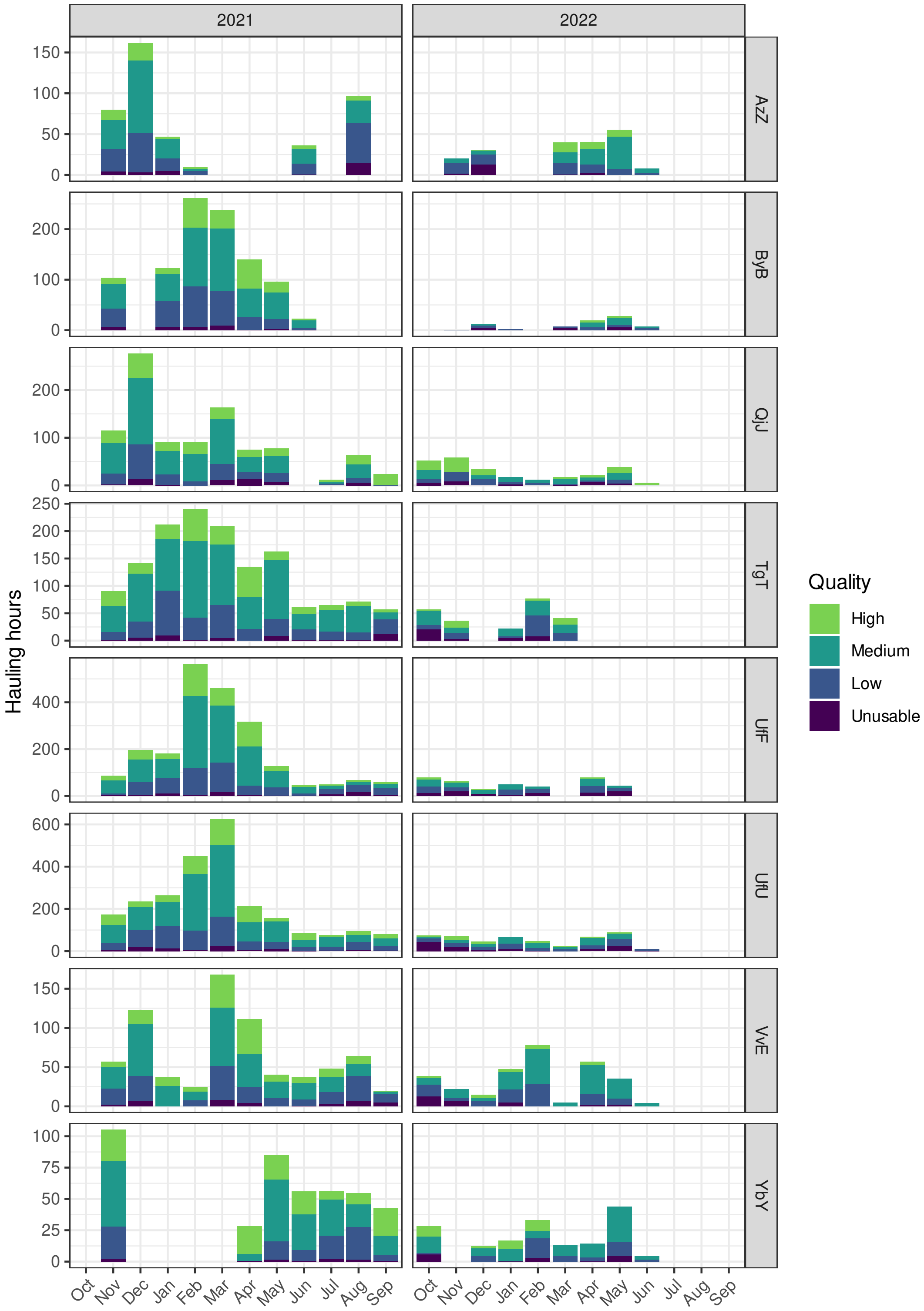


Figure 15: Hauling hours per vessel (anonymised identifier) per month classified by footage quality, based on the first review of a slice.

Table 5: Overall reviewer-assessed footage quality during hauling by vessel (anonymised identifier) and fishing year.

		Footage quality (% hours)			
Vessel	Fishing Year	High	Medium	Low	Unusable
AzZ	2021	11.8	44.9	37.3	6.0
AzZ	2022	15.8	44.7	31.2	8.4
ByB	2021	20.5	47.1	29.2	3.3
ByB	2022	9.9	36.5	34.8	18.9
QjJ	2021	22.8	50.6	21.4	5.2
QjJ	2022	34.7	28.3	27.0	10.0
TgT	2021	18.9	53.3	24.7	3.1
TgT	2022	12.9	38.7	33.3	15.1
UfF	2021	20.9	50.7	25.2	3.2
UfF	2022	6.3	35.2	34.9	23.5
UfU	2021	20.0	50.5	25.9	3.6
UfU	2022	11.8	32.2	34.3	21.7
VvE	2021	23.0	44.3	28.0	4.7
VvE	2022	6.3	53.0	31.7	9.0
YbY	2021	28.8	46.0	23.1	2.2
YbY	2022	15.4	50.7	25.5	8.5
All vessels	2021	20.6	49.7	26.0	3.7
All vessels	2022	13.5	38.8	32.1	15.7

Table 6: Reasons designated by reviewers for reduced quality footage, in terms of percentage of footage recorded in a category.

	Footage quality (% hours, by category)		
Reason	Low	Medium	Unusable
Camera poor angle	1.6	1.0	4.3
Glare	56.0	31.1	71.2
Lighting poor	2.6	4.4	1.4
Out of focus	3.8	5.1	1.7
Video other	2.8	0.5	4.8
Water spots	33.1	57.9	14.7
Video camera failure			1.8

3.5.2 Multiple reviews

In the final dataset, there were 1214 reviews of 171 seabird capture events during the 2020–21 fishing year. The majority of capture events recorded were single captures; however, 4 capture events caught two seabirds and were confirmed by the expert reviewer. There were five reviewers (in addition to the expert) and every seabird capture event was reviewed by at least four of these reviewers.

Across the confirmed seabird capture events, 81.3% were found during the primary review. As video footage was only passed on for further review if a capture was detected during primary review, this indicates that there were additional captures found during subsequent reviews of the video footage. It is unknown how many captures were missed during primary review of footage that had no captures found.

The mean detection probability of the capture events during the secondary review was 75.3%. This is an estimate of the detection probability of the seabird captures by a single reviewer, given that the seabird

Table 7: Variation in capture detection between reviewers, for seabird captures during the 2020–21 fishing year. For each reviewer, the table gives the number of secondary reviews of confirmed seabird capture events, and the percentage of those reviews that detected a seabird capture.

Reviewer	Secondary reviews	Capture detection (%)
R9	163	91.40
R15	143	90.90
R14	101	83.20
R24	163	62.60
R38	134	48.50

capture is able to be identified. The detection probability of three birds that were retained on board the vessel was lower (66.7%) than that of birds that were released alive or returned to the sea (75.4%). The detection probability was similar for birds of different species (75.9% for black petrel and 74.9% for flesh-footed shearwater), and was similar for birds that were alive or dead (75.4% for live captures 75% for dead birds). There was a wide variation in the detection probability associated with individual reviewers, however (Table 7). Of the five reviewers, there were two who detected over 90% of the seabird capture events, and one who detected fewer than 50%.

3.5.3 Seabird captures

Based on expert review of the video footage, there were 176 seabird captures during 2021, and 21 during 2022 (Figure 16). Around a third of the seabird captures were dead (26.7% of captures during 2021, and 66.7% of captures during 2022). The highest number of captures recorded from a single vessel was 71 captures.

Flesh-footed shearwater was the most frequently caught species (139 captures during 2021, and 19 captures during 2022), followed by black petrel (35 captures during 2021, and 2 captures during 2022).

A fluttering shearwater (*Puffinus gavia*) and a sooty shearwater (*Ardenna grisea*) were also captured, both in 2021.

Flesh-footed shearwater were typically caught closer to shore, with black petrel caught further offshore (Figure 17). The small number of winter captures identified were all flesh-footed shearwaters, caught in October 2021 (Figure 18). There were no captures of any seabirds recorded during June, July, August, or September.

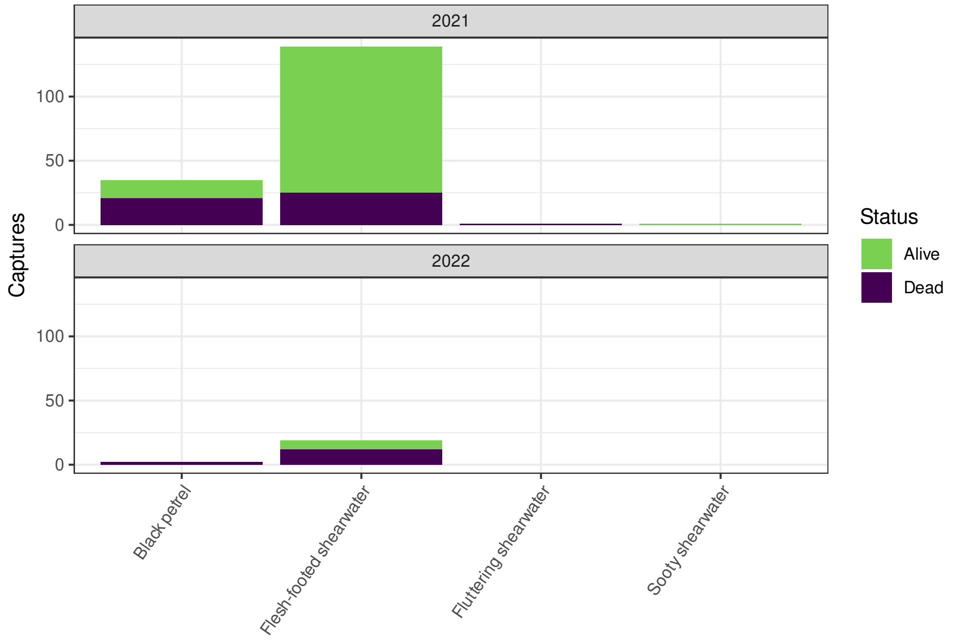


Figure 16: Seabird captures identified in the video observation programme.

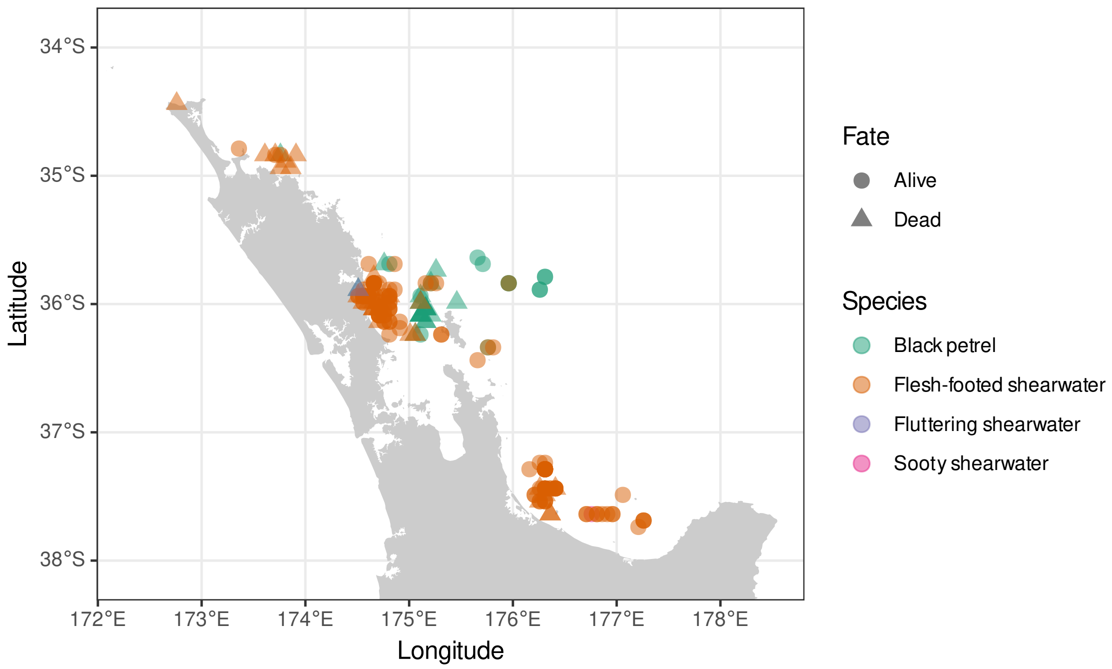


Figure 17: Locations of seabird captures identified in the video observation programme in the main season (November to May) of the 2021 and 2022 fishing years. Locations were jittered by first rounding the latitude and longitude to the nearest 0.05 degrees, then adding a random amount between plus or minus 0.025 degrees.

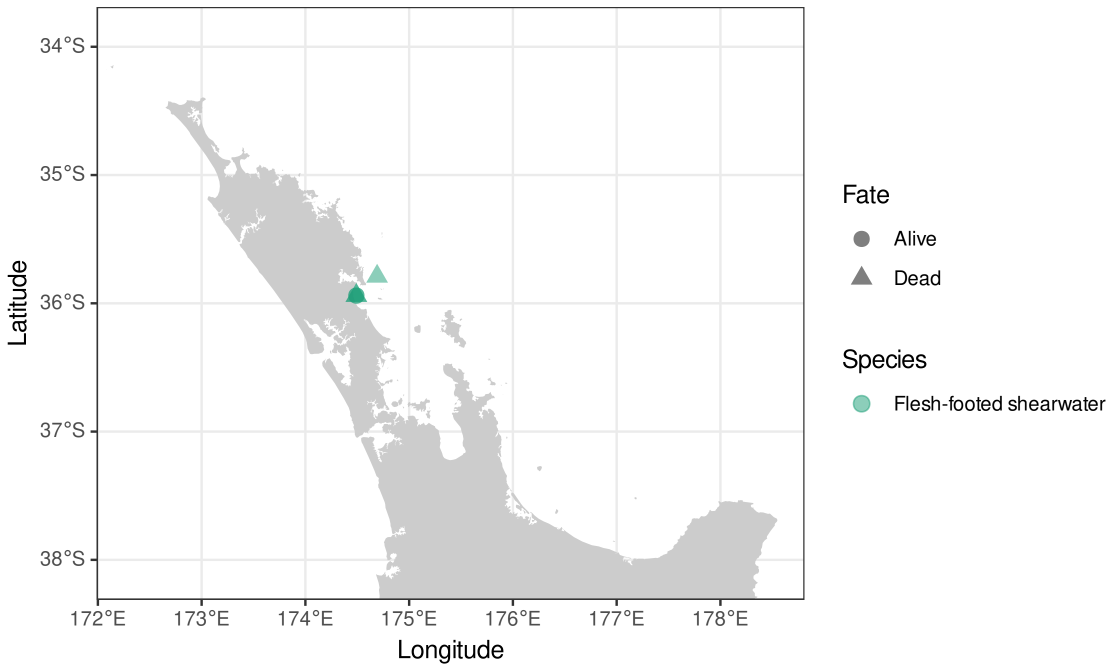


Figure 18: Locations of seabird captures identified in the video observation programme in winter (June to October) of the 2021 fishing year. Locations were jittered by first rounding the latitude and longitude to the nearest 0.05 degrees, then adding a random amount between plus or minus 0.025 degrees.

3.6 Estimated seabird captures

3.6.1 Linked dataset

The video observations were linked to the fisher-reported effort data at an event by event level. In addition to the data from 2020–21 and 2021–22, the linked dataset included video observed captures and fishing effort from 2016–17 to 2019–20. This linked dataset allowed for capture rates (seabird captures per 1000 hooks) to be determined, for comparison between video monitored and fisher-reported captures, and for estimation of total seabird captures.

The total FMA 1 bottom longline fishing effort data used for estimating seabird captures included 36 442 events (76.23M hooks). There were 10 554 events (29.76M hooks set) by vessels that participated in the video observation programme. Across all the years, 28.9% of the hooks set by vessels participating in the video programme were video reviewed; and 11.2% of the total fishing effort was video reviewed.

The linked data included a total of 367 video-observed seabird captures. The overall seabird capture rate from the video-monitoring programme was 0.042 seabird captures per 1000 hooks.

3.6.2 Estimated black petrel captures

The model of black petrel captures converged (the $\hat{R}$ of all parameters was 1.00, and there were no divergent transitions). When used to predict captures on the video-review data used to fit the model, the distributions of the numbers of black petrel caught per fishing event were similar to the observed captures (Figure 19). There was a single video-reviewed event with 9 black-petrel caught on a haul; and the model estimated that 99% of capture events were of 4 or fewer black petrel (with 99.9% being of 10 or fewer black petrel).

There were an estimated 40 (97.5% c.i.: 14 to 79) black petrel captures in all bottom longline fishing in FMA 1 between October 2021 and September 2022 (the 2022 fishing year; Figure 20). The 2022 fishing year had the lowest mean estimated captures of any of the fishing years included in the estimation. The capture rate during 2022 was 0.0039 (95% c.i.: 0.0013 to 0.0076) black petrel captures per 1000 hooks set over this period. The model and ratio estimates based on the video data were similar, with the ratio estimates typically lying within the 95% c.i. of the model estimates. The video-based estimates were lower than the estimates based on fisheries observer data in the years with estimates from both methodologies.

The highest captures were in the statistical area to the north and inshore of Great Barrier Island, while the estimated capture rate was highest immediately offshore from Great Barrier Island (Figure 21). Estimated captures were low in the inner Hauraki Gulf.

There was no evidence of a change in the black-petrel capture rate over the six-year period included in the analysis (Figure 22), with the credible interval of the year-effect including zero for each year. Similarly, there was no strong evidence of vessels having higher or lower capture rates than the mean across the fleet (although there were two vessels, EvV and ShS that had mean capture rates of more than 50% above the fleet average). The monthly effect showed an increase in capture rate in February and March, during chick-rearing. There was a decrease in catch rate during the winter season when black petrel are absent from New Zealand waters. The model also estimated a target effect (snapper target, relative to other targets) of -1.35 (95% c.i.: -2.09 to -0.53). This effect is on the natural log scale, with the model estimating that black petrel capture rates on snapper target fisheries are around one-quarter of the capture rates on bottom longline fishing targeting other species.

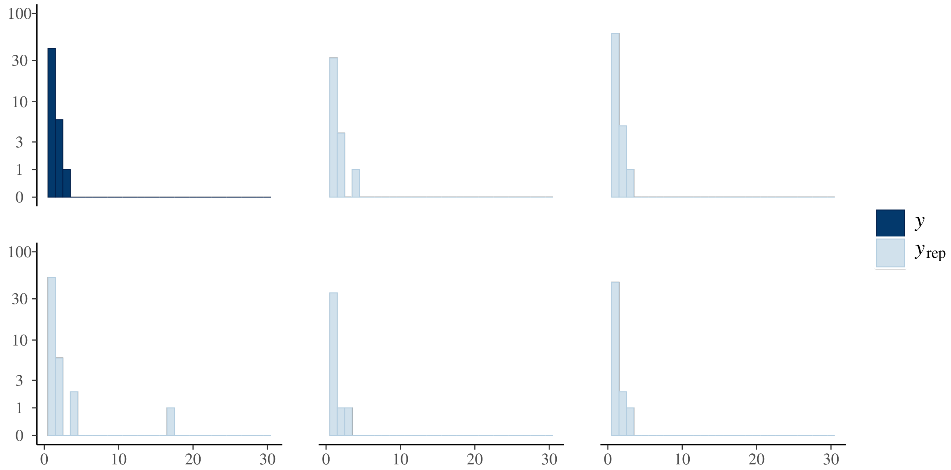


Figure 19: The distribution of black petrel caught per fishing event, taken from the video-review data used to fit the model (y), and from five samples of the posterior distribution of estimates taken from applying the model to the video-reviewed events (y_{rep}). The x-axis shows the number of black petrel captures caught (with captures), while the y-axis is the number of fishing events with that number of captures. Events with no captures were excluded from the figure.

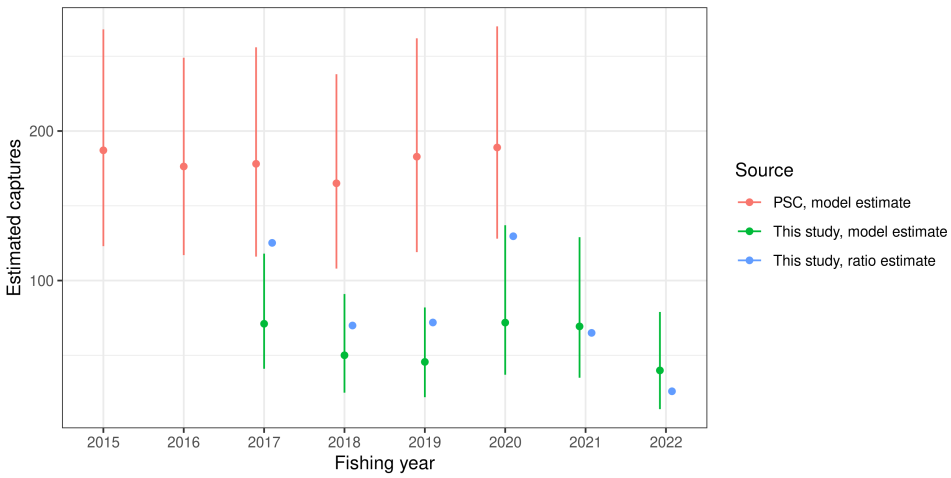


Figure 20: Annual estimated captures of black petrel in FMA1 bottom longline fishing by fishing year, from the model used in this study (green); from a ratio estimate that extrapolated the video-data derived capture rate to all the fishing effort (blue); and from estimation based on observer data (Protected Species Captures, PSC, red). For the two model estimates, the mean and 95% credible interval is shown.

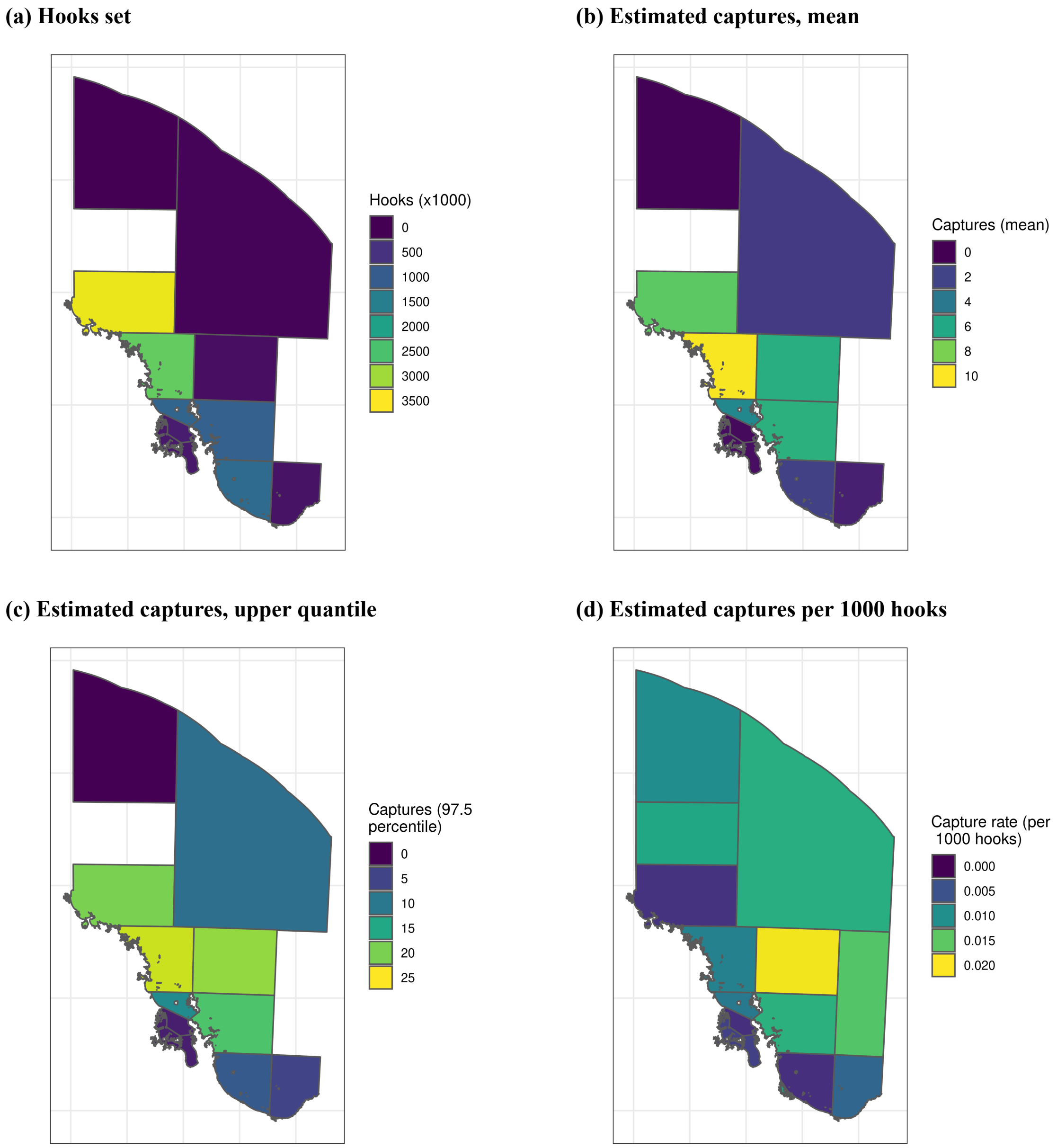
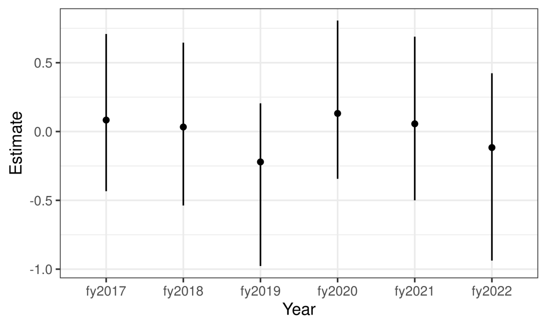
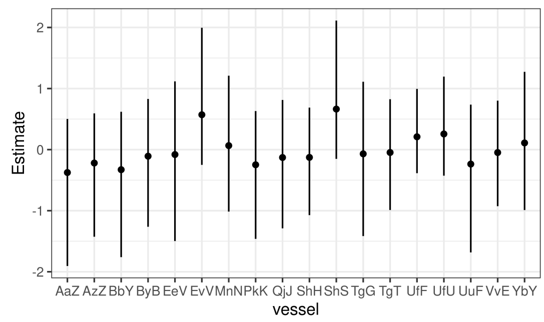


Figure 21: Estimated black petrel captures during the 2021–22 fishing year. The maps show the statistical areas in Fisheries Management Area 1. The maps are coloured by statistical area, with the colour indicating: (a) the number of hooks set (thousands); (b) the mean estimated number of black petrel captures (c) the upper quantile (97.5%) of the estimated black petrel captures; (d) the mean estimated black petrel capture rate (birds caught per 1000 hooks set). The rate was calculated for a single vessel (ShS), targeting snapper, and fishing during January 2022.

(a) Fishing year



(b) Vessel



(c) Month

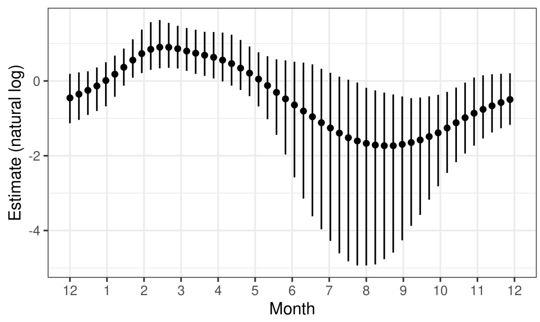


Figure 22: Estimated random effects of the black petrel model: (a) fishing-year (b) fishing vessel (c) month, estimated using a cyclic spline. For each value, the figure gives the mean and the 95% credible interval of the random effect. The estimated effect is shown on the natural logarithmic scale.

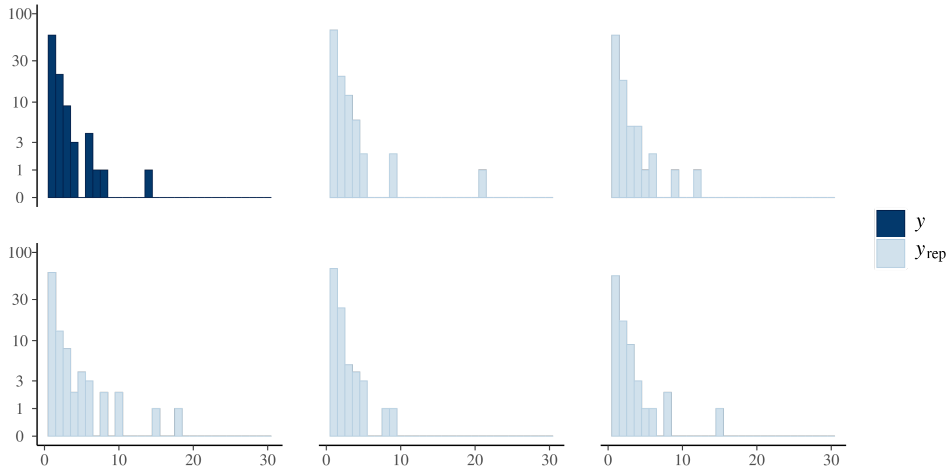


Figure 23: The distribution of flesh-footed shearwater caught per fishing event, taken from the video-review data used to fit the model (y), and from five samples of the posterior distribution of estimates taken from applying the model to the video-reviewed events (y_{rep}). The x-axis shows the number of flesh-footed shearwater captures caught (with captures), while the y-axis is the number of fishing events with that number of captures. Events with no captures were excluded from the figure.

3.6.3 Estimated flesh-footed shearwater captures

The model of flesh-footed shearwater captures converged (the $\hat{R}$ of all parameters was 1.01 or less, and there were no divergent transitions). When used to predict captures on the video-review data used to fit the model, the distributions of the numbers of flesh-footed shearwater caught per fishing event were similar to the observed captures (Figure 23). There was a single video-reviewed event with 14 flesh-footed shearwater caught on a haul; and the model estimated that 99% of capture events were of 14 or fewer flesh-footed shearwater (with 99.9% being of 36 or fewer flesh-footed shearwater).

There were an estimated 159 (97.5% c.i.: 56 to 392) flesh-footed shearwater captures in all bottom longline fishing in FMA 1 during the 2022 fishing year (Figure 24). There was a marked peak in the flesh-footed shearwater captures during the 2021 fishing year, with the mean estimated captures reaching 790 (95% c.i.: 349 to 1735). The ratio estimated captures were also high during this year. In contrast, the estimated captures of flesh-footed shearwater during 2017 were only 29 (95% c.i.: 6 to 83). The capture rate during 2022 was 0.0155 (95% c.i.: 0.0055 to 0.0382) flesh-footed shearwater captures per 1000 hooks set over this period. The model and ratio estimates based on the video data were similar, with the ratio estimates typically lying within the 95% c.i. of the model estimates. The video-based estimates were similar to the estimates based on fisheries observer data, although the estimates from the fisheries observer data did not have any inter-annual variation (other than changes due to changes in the fishing effort).

The highest estimated captures were in the statistical area to the north of the Bay of Islands, where the number of hooks set was highest (Figure 25). While the capture rate was low in the inner Hauraki Gulf, there was otherwise not strong variation in the capture rate through the region. The highest estimated capture rate was in the statistical area offshore from Tauranga, in the Bay of Plenty.

The flesh-footed shearwater catch rate was lower than the mean in the 2017 fishing year, and higher than the mean during the 2021 fishing year Figure 26, with the credible interval of the year-effect including zero for other years. There were two vessels (AzZ and YbY) that had an estimated vessel effect whose credible interval was higher than zero, with these vessels having higher capture rates than the rest of the fleet. There was a marked decrease in the catch rate during the winter season, from May to September. The model estimated a target effect (snapper target, relative to other targets) of -0.48 (95% c.i.: -1.31 to

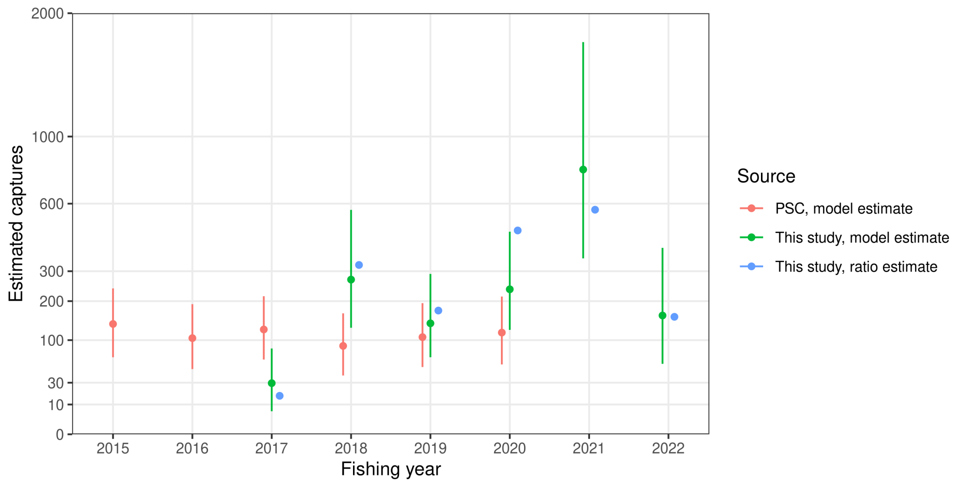


Figure 24: Annual estimated captures of flesh-footed shearwater in FMA1 bottom longline fishing by fishing year, from the model used in this study (green); from a ratio estimate that extrapolated the video-data derived capture rate to all the fishing effort (blue); and from estimation based on observer data (Protected Species Captures, PSC, red). For the two model estimates, the mean and 95% credible interval is shown.

0.34), indicating that there was no marked difference in flesh-footed shearwater capture rates between fishing targeting snapper or other species.

3.7 Fisher-reported captures

During the 2020–21 and 2021–22 fishing years, there were 257 seabird captures reported by fishers from bottom longline fishing within FMA 1. Of these captures, 218 were reported from vessels participating in the video monitoring trial (a capture rate of 0.0259 seabird captures per 1000 hooks). There were 39 captures reported from other vessels, with a capture rate of 0.0027 seabird captures per 1000 hooks. The rate of reporting by vessels participating in the trial was close to ten times as high as the rate of reporting by other vessels. During the 2018–19 and 2019–20 fishing years, the rate of fisher-reported seabird captures was around twice as high in the video-observation fleet as from other vessels (Middleton & Abraham 2023), and a more detailed analysis of data from the 2017 and 2018 fishing year found a doubling in the reporting rate when vessels began participating in the trial (Tremblay-Boyer & Abraham 2020).

The linked dataset allowed the captures reported from the video monitoring to be compared with the captures reported by fishers. During the 2020–21 and 2021–22 fishing years, there were 197 seabird captures from video review in the linked dataset. On the fishing events that were video-reviewed, there were 181 fisher-reported captures. While the overall numbers of reported captures were similar, only 84 (42.6%) of the video-review captures were on events that had the same number of fisher-reported captures, so there was considerable discrepancy at an event by event level, as was noted previously (Middleton & Abraham 2023).

The fisher-reported captures can be compared with the model-estimated captures (Figure 27). For the vessels participating in the video-monitoring, the fisher reported captures were within the credible interval of the model estimates and often close to the mean value (although less than the mean in all years but 2017). For the other vessels, the fisher reported captures were less than the lower limit of the credible interval in many of the years.

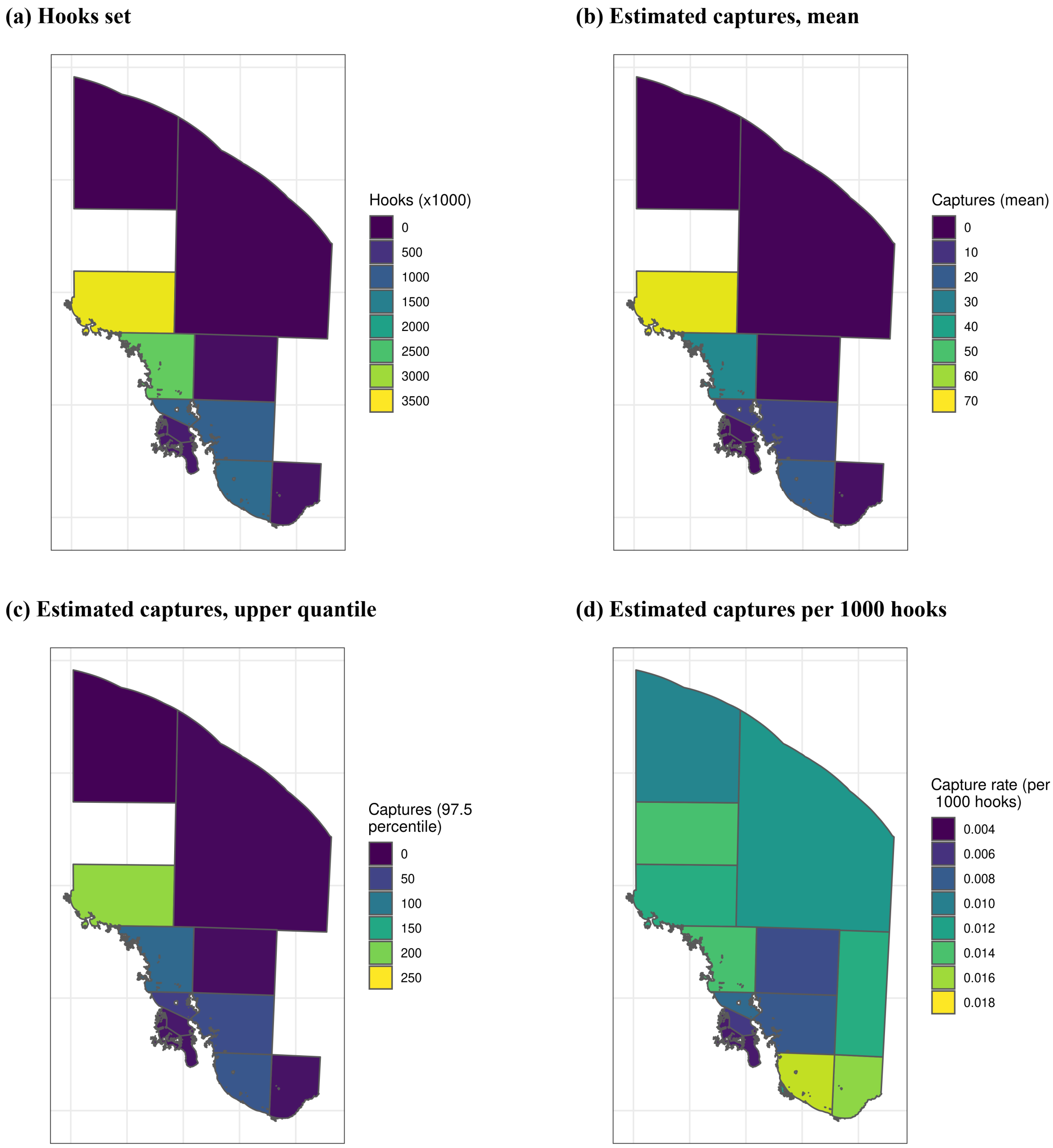


Figure 25: Estimated flesh-footed shearwater captures during the 2021–22 fishing year. The maps show the statistical areas in Fisheries Management Area 1. The maps are coloured by statistical area, with the colour indicating: (a) the number of hooks set (thousands); (b) the mean estimated number of flesh-footed shearwater captures (c) the upper quantile (97.5%) of the estimated flesh-footed shearwater captures; (d) the mean estimated flesh-footed shearwater capture rate (birds caught per 1000 hooks set). The rate was calculated for a single vessel (ShS), targeting snapper, and fishing during January 2022.

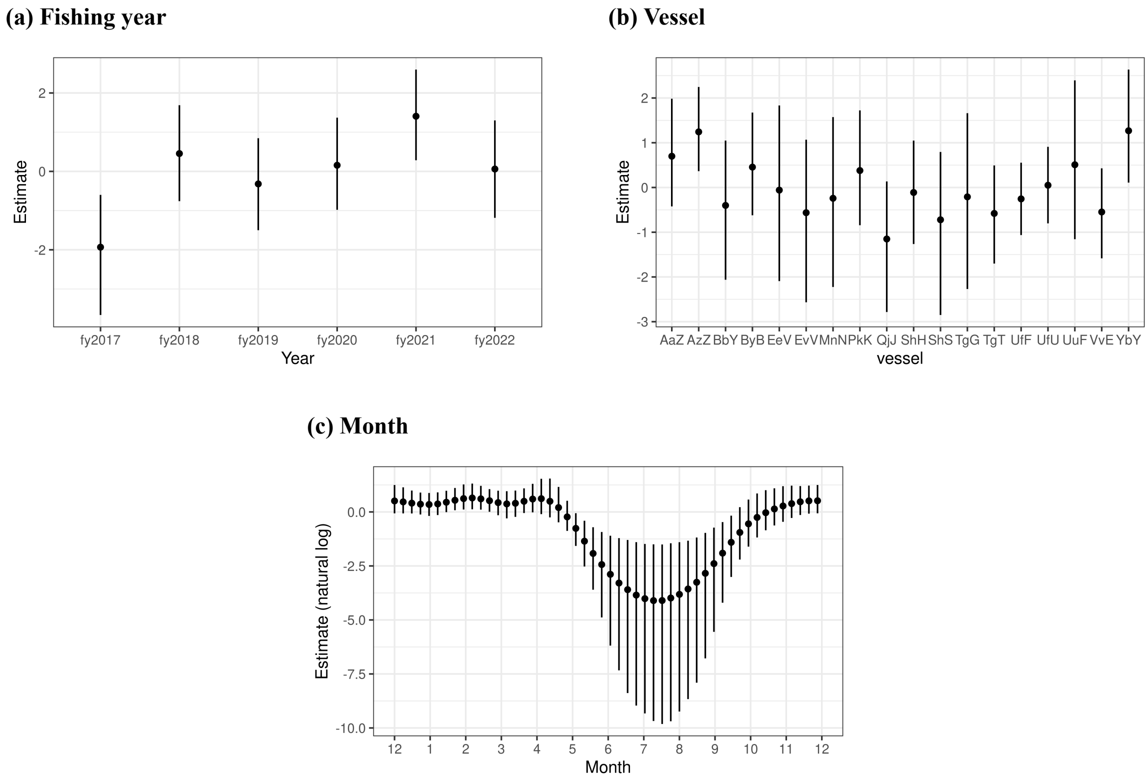


Figure 26: Estimated random effects of the flesh-footed shearwater model: (a) fishing-year (b) fishing vessel (c) month, estimated using a cyclic spline. For each value, the figure gives the mean and the 95% credible interval of the random effect. The estimated effect is shown on the natural logarithmic scale.

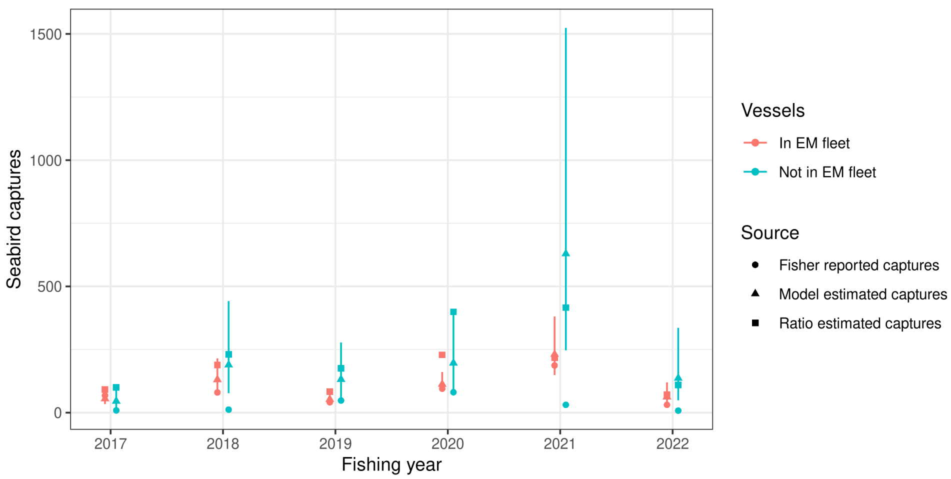


Figure 27: A comparison between fisher-reported captures, ratio-estimated captures, and model-estimated captures. The comparisons were made for fishing by vessels participating in the video-monitoring programme (In EM fleet) and other vessels (Not in EM fleet). In each year, and for each group of vessels, the figure shows the mean and 95% c.i. of the model-estimated captures (triangle and line), the ratio-estimated captures (square) and the fisher reported captures (circle).

4. DISCUSSION

Video observation provided coverage of 23.8% of bottom longline effort targeting snapper between November 2019 and May 2020 (Middleton & Abraham 2023). The video observation coverage in the 2021 fishing year (16.2% of effort) was somewhat lower and was further reduced (8.9% of effort) in 2022. Some reduction in coverage was anticipated due to reduced allowance for reviewing within the projects, although coverage in 2022 was also clearly impacted by low footage collection on some vessels in certain months (Table 4). These periods were due to technical issues with moisture affecting the wheelhouse units of the EM systems.

The higher level of observer coverage achieved by video observation allows for inter-annual variability to be included in the bycatch estimation. Because of the low coverage (around 2%) from human observers (Abraham & Richard 2020) estimation has previously assumed a constant capture rate over the whole period, with changes in the estimated captures been driven only by changes in the distribution and the amount of fishing effort. There was evidence from the video observation of strong variation in captures, with estimated captures of flesh-footed shearwater varying from a mean of 790 (95% c.i.: 349 to 1735) captures during the 2020–21 fishing year to a mean of 159 (97.5% c.i.: 56 to 392) captures during the 2021–22 fishing year. The estimated captures of black petrel were lower than the number of captures estimated from human observer data. It would be interesting to carry the lower estimate of black petrel captures through the risk assessment process (Richard & Abraham 2020) to understand how this affects the impact of fisheries bycatch on black petrel populations.

Video observation also provided data during the winter of 2021. These were the first observations of the bottom longline fishery during winter. There had previously been estimated captures of grey petrel in this fishery during the winter (Abraham & Richard 2020). These arose from a combination of a lack of winter observer coverage in this fishery, and observed captures in other bottom longline fisheries further south. No grey petrel captures were recorded during video review, and so the video observation will help to reduce uncertainty in the estimation of grey petrel captures. The winter video observation also helps to define the seasonal variation in the captures of black petrel and flesh-footed shearwater, with no seabird captures recorded from June to September, but some flesh-footed shearwater captures and a single black petrel capture recorded during October.

In this study, an extensive set of multiple-reviews was undertaken to investigate between-reviewer differences highlighted in Middleton & Abraham (2023). The multiple reviews allowed differences in reviewer-skill at locating seabird capture events in the footage to be reliably estimated. There was a large variation in reviewer skill, with one reviewer who detected fewer than 50% of the captures that had been found by other reviewers. We did not use this information to correct for the estimates of the total seabird bycatch. The estimation has assumed that all seabird captures were detected in the reviewed footage. Because of this, the total seabird capture estimates are likely to be an underestimate. It is likely that a similar under-counting occurs with human observations, and estimation models based on human observers also make the assumption that all captures on observed fishing events are recorded (Abraham & Richard 2020). A focus of future work could be on developing the statistical methods for estimating seabird bycatch that allow for the variation in reviewer skill to be accounted for.

Seabird capture events are rare, and finding them is a “needle in the haystack” problem. It is inevitable that reviewer fatigue and distraction will lead to some captures being missed. This motivates the development of machine learning (artificial intelligence) methods that are able to systematically process all the footage. An exploration of these approaches, applied to the petrel project footage, shows that they are able to detect the captures (Henry Zwart, Dragonfly Data Science, pers. comm.). There was a high false positive rate, however, with many events that were not seabird captures (e.g., shadows, fish captures) being scored as captures. Because of the false positives, the machine learning detected captures would need subsequent human review. A key advantage of the video observations is that the machine learning methods can be retrospectively applied to the archived footage. This allows for a continued improvement of the methodology.

The Fisheries (Electronic Monitoring on Vessels) Regulations 2017 introduced a requirement for fishing vessels, including those undertaking bottom longlining, to carry electronic monitoring equipment that records video footage of fishing activity. Under these regulations, Fisheries New Zealand is implementing an on-board cameras programme on commercial fishing vessels, with new electronic monitoring systems being installed between 2022 and 2024.[3] Monitoring carried out through this programme will replace the monitoring that has been carried out through the petrel project. It is hoped that the experience gained through the petrel project will help inform the broader development of video observation.

A limitation of the petrel project has been that the video observation has been restricted to participating vessels. It is possible that fishing practices differ between the rest of the fleet and the participating vessels. This is evidenced by the low fisher-reporting of seabird captures by the rest of the fleet, suggesting less engagement with seabird bycatch issues. The regulated electronic monitoring programme will extend the coverage across the whole fishery.

[3] https://www.mpi.govt.nz/fishing-aquaculture/commercial-fishing/fisheries-change-programme/on-board-cameras-for-commercial-fishing-vessels/

5. ACKNOWLEDGEMENTS

This work was carried out under Fisheries New Zealand projects PSB2020-10 and PRO2021-07. The cost recovered portion was attributed to SNA 1 and BNS 1 quota owners.

The vessels monitored in the project participated in the project on a voluntary basis and the participation of their operators, and the cooperation of their crews, is gratefully acknowledged.

TeemFish provided the video collection and review infrastructure. A team of reviewers contracted by Dragonfly Data Science and Guards Management Services reviewed the footage, and the determinative review of all seabird captures identified was carried out by Mike Bell of Toroa Consulting.

6. REFERENCES

Abraham, E.R.; Richard, Y. (2020). Estimated captures of seabirds in New Zealand trawl and longline fisheries, to 2017–18. *New Zealand Aquatic Environment and Biodiversity Report No. 249*. 86 p.

AI Forum of New Zealand (2019). *Towards Our Intelligent Future*. AI Forum of New Zealand.

Anonymous (2019). *Data trusts: legal and governance considerations* (tech. rep.). Open Data Institute.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software 80*: 1–28. https://doi.org/10.18637/jss.v080.i01

Gelman, A.; Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7 (4)*: 457–472.

Hardinges, J.; Wells, P.; Blandford, A.; Tennison, J.; Scott, A. (2019). *Data trusts: lessons from three pilots* (tech. rep.). Open Data Institute.

Knuth, D.E. (1992). *Literate Programming*. Cambridge University Press.

Mansell, J.; Laking, R.; Matheson, B.; Light, R. (2017). *Data Commons Blueprint, v1.3. A high-trust, lower-cost alternative to enabling data integration and reuse* (tech. rep.). Aotearoa New Zealand Data Commons Project.

Middleton, D.A.J.; Abraham, E.R. (2023). Video observation of the FMA 1 bottom longline fishery in 2018–19 and 2019–20. *New Zealand Environment and Biodiversity Report No. 302*. 42 p.

Middleton, D.A.J.; Guard, D. (2021). Summary and evaluation of the electronic monitoring programmes in the SNA 1 trawl and bottom longline fisheries. *New Zealand Fisheries Assessment Report 2021/37*. 69 p.

MPI (2011). *Research and Science Information Standard for New Zealand Fisheries*. Ministry for Primary Industries. https://www.mpi.govt.nz/dmsdocument/3692

Peatman, T.; Edwards, C.T.T.; Goad, D.; Webber, D.N. (2023). Review of biological inputs for the New Zealand Seabird Risk Assessment. *New Zealand Aquatic Environment and Biodiversity Report No. 312*. 193 p.

Privacy Commissioner (2009). *Privacy and CCTV. A guide to the Privacy Act for businesses, agencies and organisations*. Office of the Privacy Commissioner.

Richard, Y.; Abraham, E.R. (2020). Assessment of the risk of commercial fisheries to New Zealand seabirds. *New Zealand Aquatic Environment and Biodiversity Report No. 237*. 61 p.

Tremblay-Boyer, L.; Abraham, E.R. (2020). Increased fisher-reporting of seabird captures during an electronic-monitoring trial. *New Zealand Aquatic Environment and Biodiversity Report No. 238*. 32 p.

Whitcroft, A. (2019a). Data trusts — a model for Aotearoa. https://medium.com/proceeding-by-inquiry/data-trusts-a-model-for-aotearoa-pt-i-1cdbbb863c21

Whitcroft, A. (2019b). Data trusts and trusting openness. https://opendatacharter.medium.com/data-trusts-and-trusting-openness-e7b6ae285ad2

APPENDIX A: VIDEO OBSERVATION DATA WORKSHOP

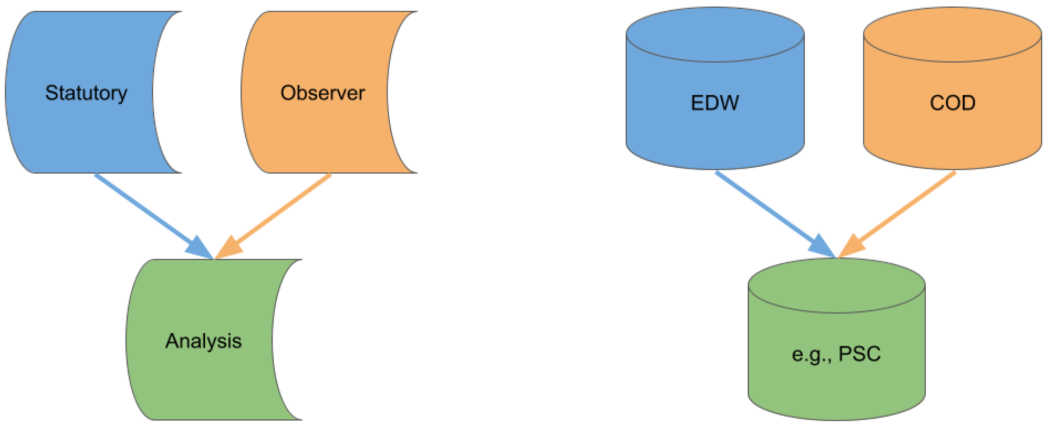
A workshop was convened on 8 February 2022 to discuss options for storage of video observation data. A specific question, that had arisen during contracting of the PRO2021-07 project, was whether video observation data should, or could, be stored in MPI's Centralised Observer Database (COD), alongside data collected through the traditional deployment of observers onto fishing vessels.

Here we summarise the background presentation given to workshop participants, and the resulting discussion.

A.1 Background

A.1.1 Introduction

Fisheries data come in many forms. The data used for assessing levels of protected species captures typically involves a combination of statutory reporting, that provides a comprehensive record of fishing activity, together with observer data that provides details of protected species captures from a subset of effort (Figure A-1a). The different data sources are typically stored in different databases (Figure A-1b); in New Zealand fisheries researchers access statutory data via MPI's Enterprise Data Warehouse (EDW) database, and observer data via the COD database. In some cases, a special-purpose analysis database is generated from these source data; for example, the Protected Species Captures (PSC) database.



(a) Some types of fisheries data. **(b) Examples of New Zealand fisheries databases.**

Figure A-1: Examples of types of fisheries data, and the corresponding databases used for storage/reporting.

We use the term **statutory** data to refer to data that Government *requires from fishers* (usually commercial fishers). In New Zealand, the legislative basis for such requirements is the Fisheries Act 1996, including Part 10 *Recordkeeping, reporting, disposal of fish, and provisions relating to taking and possession of fish for purpose of sale*, s297 *General regulations*, and s304 *Circulars*. More specific details of the data required from all commercial fishing activity are laid out in the Fisheries (Reporting) Regulations 2017 and a series of circulars, such as the Fisheries (E-logbook Users Instructions and Codes) Circular 2021.

Observer data collection also has a statutory basis (Part 12 of the Fisheries Act 1996), but the statutory powers are focussed primarily on the obligation for fishers to carry an observer, rather than the detailed nature of the data collected. Observer data are collected *independently*[4] by observers, and are usually only available from a subset of the effort in a fishery. Because observer data are collected by specific

[4] of the normal fishing activity and fishery participants

data-collection personnel, these can be more detailed than the statutory data, which must balance the detail sought against the onerous nature of providing data from each fishing event.

Many analyses conducted to inform fisheries management decision making can be made using just the statutory or just the observer data. However, other analyses require using the two datasets together; for example, when scaling estimates made from observer data up to the comprehensive 'fleet level'. In such cases, the task of providing a linked dataset (such as that assembled in the PSC database) may often require a degree of analysis and task-specific decision-making in order to reconcile data collected from different sources over some decades.

The storage of data collected using new technologies, such as 'Electronic Monitoring' (EM) systems installed on fishing vessels, does not necessarily fit neatly into any existing data system. In the case of the petrel project (Middleton & Guard 2021, Middleton & Abraham 2023) the specific intention is to collect observational data that are largely equivalent to data collected by observers physically present on a fishing vessel, at least from the perspective of analyses using observer data. However, supplementing observer data is only one potential role for EM data; an alternative use would be in the *verification* of statutory data.

To date, key outputs from the petrel project have been datasets that are assembled and formatted in a way that facilitates their incorporation into the PSC database, alongside traditional observer data from the COD database (Figure A-2). However, these 'prepared EM data' are a derived dataset, summarising a variety of other 'raw' data records. The question of whether these 'raw' EM data could usefully be stored alongside the raw data from on-board observers in the COD database requires consideration of the nature of EM data.

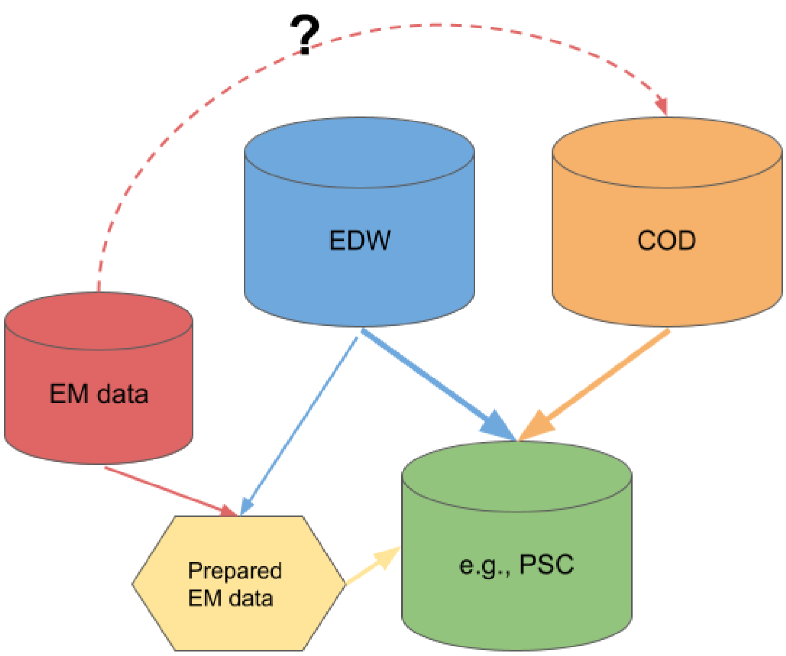


Figure A-2: Illustrating the current flow of EM data collected by the petrel project into the PSC database for estimation of fleet-level seabird captures, alongside data from the EDW and COD databases, and the potential need to store EM data alongside traditional observer data in future.

A.1.2 Video observation data

The terminology used for EM/video observation varies, both around the world and within New Zealand. Here we follow, and extend where necessary,[5] the terminology introduced by Middleton & Guard (2021). The same basic terminology should be appropriate whether EM is being used in video observation (i.e., generating observer-like data) or being used to validate statutory data.

We make a high-level distinction between **EM footage**, the raw electronic data from EM systems, including the video footage, GPS track data, and data from any other sensors that are part of the EM system, and **EM data**, the observational data resulting from review of the EM footage, usually by shore-based reviewers (or 'video observers') but potentially also from automated processing of the EM footage. Other key terminology is identified in Figure A-3; the data (or metadata) generated in the different stages of the video observation process is discussed further below.

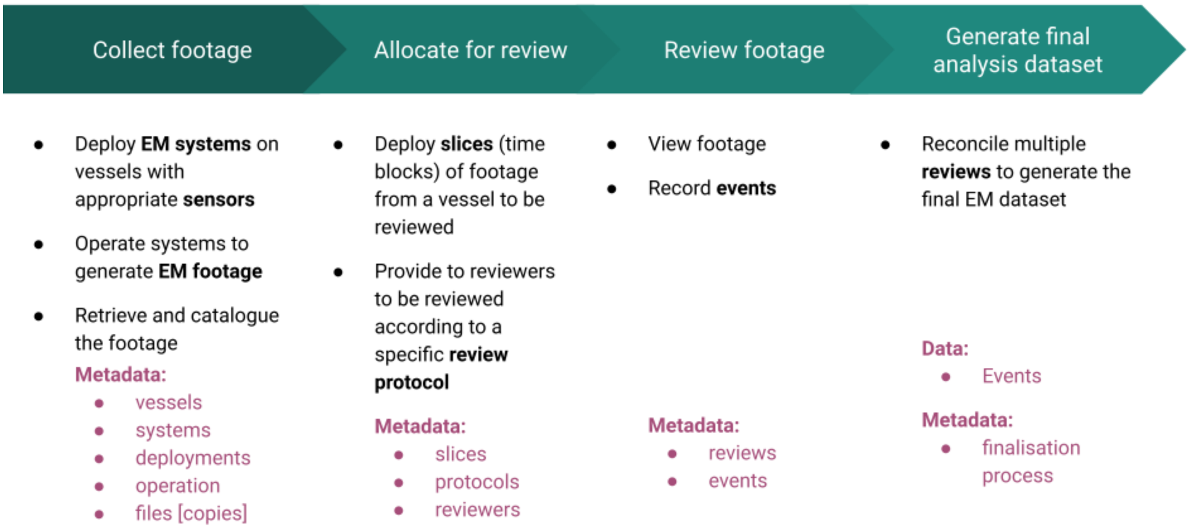


Figure A-3: Stages in the video observation process, with key terminology (in bold) and stage-specific metadata, and the ultimate EM data, noted in purple.

All stages of the video observation process generate a range of data and metadata. These are described below, and can be considered as describing:

- operational data relating to the collection of footage;
- a catalogue of EM footage;
- EM data, generated from the review of EM footage; and
- analysis datasets generated from EM data.

In the remainder of this report, we use JSON (JavaScript Object Notation[6]) to illustrate the scope of the data attributes associated with a particular class of data. JSON is a lightweight data-interchange format, used here because it is easy for humans to read and write. It is also easy for machines to parse and generate, but in a real database of EM data it would be normal to store data in a format that requires less overhead to query.

A.1.3 EM systems and deployments

Any use of Electronic Monitoring first requires the deployment of an **EM system** onto a vessel. An EM system will typically comprise:

[5] i.e., for machine learning applications
[6] https://www.json.org/

- a range of **sensors**, such as:
  - cameras;
  - a Global Positioning System (GPS) unit;
  - gear sensors;
- a **storage** device, such as a solid-state drive;
- a **power supply**, possibly with additional sensors to detect power interuptions etc.;
- **data transmission** components such as removable drives, cellular or satellite modems, etc.

In order to effectively operate a video observation programme, it is necessary to keep track of the **deployments** of EM systems onto vessels (Figure A-4). A record of EM system deployments functions in the same way as a record of observer trips: it creates the expectation that EM footage will be available. Like many of the data types illustrated here, key attributes of a deployment record are the start and end time that the system was deployed on a particular vessel.

```
{
  "vessel_id": "99887766",
  "em_system_id": "cc00xx99yy77",
  "deployment_start": "2015-10-29 12:45:17 NZDT",
  "deployment_end": "2018-11-29 15:18:17 NZST"
}
```

Figure A-4: An example of the data collected recording an EM system deployment on a vessel.

An EM system is unlikely to be fully operational throughout a deployment; electronic equipment on vessels, especially when some of the components are installed in exposed locations, is susceptible to malfunctions. Ideally these will be rare, but it will nevertheless be necessary to keep records of EM systems **operational data** to maintain a complete picture of when EM footage should be available. Operational data may include:

- power supply data;
- known hardware issues;
- scheduled maintenance;
- routine maintenance, e.g., crew cleaning lenses;
- communications with vessels/operators;
- snapshots from systems at sea.

A.1.4 EM footage

The raw footage (e.g., the video) collected by EM systems will be stored in files, created on the EM system and ultimately transferred to storage. Such files will usually be of a standard length (e.g., 15 minutes) designed to ensure that the files remain within the limits of the file systems used, and can be expected to be transferred in a single operation. Splitting a continuous stream of footage across short files also adds robustness; for example, in the event of a power failure the file currently being written is at greatest risk of corruption. Shorter files implies less data loss if a file becomes corrupted, although this must be balanced against any overhead created by creating new files.

Although a basic catalogue of files is available simply by listing the contents of the storage directory, this has limitations. Simple directory listings can be made more useful by encoding information in the filename and/or directory structure but, to be useful for dynamic queries, it will be necessary to maintain a catalogue of the source files in a database table. The contents of such a catalogue entry for an individual file are illustrated in Figure A-5. Key attributes are the origin of the file (e.g., the vessel or deployment, and the specific sensor, that generated the file) and the period to which the recording relates. Features of the file, such as the size and duration, may be usefully recorded during cataloguing. Information should also be included to verify the integrity of the files, such as a checksum. Ideally, these statistics will be associated with a file immediately after it has been generated by the EM system and subsequently used to ensure that the file remains intact as it is transmitted from the vessel to long-term storage.

The successful operation of an EM programme requires a 'parcel tracking' system for footage, ensuring that all recorded files are faithfully transmitted to storage. Files should generally be transmitted sequentially, with the oldest files transmitted first. If files are transmitted out of sequence, this risks delaying the review process while waiting for gaps in the footage to be filled. However, in some circumstances it may be possible to prioritise some footage for transmission first; for example, based on time of day, using fisher-reported data on when fishing operations occurred, or by making use of features of the sensor data themselves (i.e., gear sensors, or motion detection in video, etc.).


```
{
  "vessel_id": "99887766",
  "em_system_id": "cc00xx99yy77",
  "sensor_id": "aa27hh79yy45",
  "file_name": "aa27hh79yy45_1446289200000.mp4",
  "file_start": "2015-11-01 00:00:00 NZDT",
  "file_end": "2015-11-01 00:14:59 NZST",
  "file_size": "58.9 MB",
  "duration": "900",
  "checksum": "595f44fec1e92a71d3e9e77456ba80d1"
}
```


Figure A-5: An example of a data record cataloguing a video file from an EM system.

A.1.5 Identifying footage for review

Once the required footage is available, it can be reviewed to generate EM data. Conceptually, the EM system on a vessel delivers a continuous stream of footage, and it is necessary to define a **time slice** of this stream as the basis for a particular review. Many different units could form the basis of this slicing for review; for example, some reviewing may focus on a port-to-port trip, while other approaches may focus on a particular fishing event (e.g., the haul of a longline), or on other events evident in the fisher-reported data, such as a protected species capture. These considerations define the *type* of the slice; other data fields relate to the vessel involved, the time boundaries of the slice, and potentially the specific sensor data that must be reviewed (Figure A-6).

A.1.6 Review metadata

Once slices have been defined, these will be the subject of **reviews**. A slice may be reviewed a number of times, either for different purposes, or to test the reproducibility of the review process. Most reviewing that seeks to generate 'observer-like' EM data will require that reviewing is carried out according to a well-defined **review protocol**. Such protocols will usually be defined in standalone documentation, and good version control practices should be applied as it is likely that protocols will require to evolve over time, and interpretation of the resulting data may change as a result. There is likely to be a close relationship between the review protocol and the data entry screens available in the reviewing software. Consideration could usefully be given to the use of literate programming techniques (Knuth 1992) to generate both the review interface and protocol documentation from the same source file, to ensure a clear match between the data fields required and the instructions and interface provided to reviewers. A review protocol should define:

```
{
  "slice_id": "1500977",
  "vessel_id": "99887766",
  "type": "trip",
  "slice_start": "2017-01-13 08:45:17 NZDT",
  "slice_end": "2017-01-16 14:38:53 NZST",
  "sensors": [
    {
      "type": "GPS",
      "sensor_id": "gg57pp69ss77"
    },
    {
      "type": "camera",
      "sensor_id": "aa27hh79yy45"
    }
  ]
}
```

Figure A-6: An example of a data record defining a time slice of footage from an EM system that will be reviewed.

- the purpose of the review (e.g., the independent identification of seabird captures);
- the footage expected to be available;
- the review approach to be taken, including the permitted playback speeds;
- the data that the reviewer must generate;
- optional data that a review may generate, if they deem it appropriate.

Each review will give rise to a data record (Figure A-7) that records the slice reviewed, by whom, and according to what protocol. The time expended in undertaking the reviewing will also be recorded, at a minimum as a review start and end time, but additional data on the review effort (breaks in reviewing, playback speeds used, etc.) may also be captured. Where reviewing is undertaken by humans, rather than an automated process, then such metadata may assist in a variety of situations including managing reviewer Health and Safety (e.g., ensuring breaks from screen time), or detecting when reviewers may miss events due to lapses in attention.

```
{
  "review_id": "150097701",
  "slice_id": "1500977",
  "protocol_id": "psb2020-10-01",
  "reviewer_id": "177",
  "review_start": "2017-03-07 08:45:17 NZDT",
  "review_end": "2017-03-07 15:18:17 NZST",
  "events": [
    {},
    {}
  ]
}
```

Figure A-7: A data record describing a single review of slice of footage from an EM system, according to a given protocol. During each review, a series of individual data records, 'events', will be defined - conceptually these are part of the top level review record, illustrated here as a list of associated events.

A.1.7 EM data

Much of the data generated during a review of EM footage will be in the form of **events**, with an identified start and end time. For example, events may describe the vessel activity observed for a period (Figure A-8) or the capture of a seabird, including detail on the species and fate of the bird (Figure A-9).


```
{
  "event_id": "15009770101",
  "review_id": "150097701",
  "event_start": "2017-01-14 08:37:11 NZDT",
  "event_end": "2017-01-14 14:33:57 NZDT",
  "event_type": "vessel",
  "event_value": "hauling"
}
```


Figure A-8: A data record describing a vessel activity recorded during a review of EM footage.


```
{
  "event_id": "15009770101",
  "review_id": "150097701",
  "event_start": "2017-01-14 11:33:11 NZDT",
  "event_end": "2017-01-14 11:37:57 NZDT",
  "event_type": "seabird capture",
  "event_value": {
    "species_code": "XFS",
    "status": "alive",
    "fate": "released"
  }
}
```


Figure A-9: A data record describing a seabird capture observed during a review, with detail added on the seabird species, and the status ad fate of the bird.

A additional class of EM data, that is particularly relevant to the development of machine-learning approaches to the review of EM footage, is **annotations**. Traditional EM data are in the form of events, defined by the start and end time during which something was observed to occur. Annotation data require the reviewer to also define a region of the image in which the event occurred (e.g., a polygon drawn around a seabird visible in the footage; Figure A-10). Annotation data therefore require the ability to define a region of the image where an event is occurring. Furthermore, with video footage, the annotation region may move from frame to frame; this requires that an annotation record is able to refer to the frame that was viewed when it was created, and also indicates the need to link a sequence of annotations together to adequately describe the movement of the object of interest through a sequence of frames. Metadata should be recorded with each annotation describing who created it, when, and why.

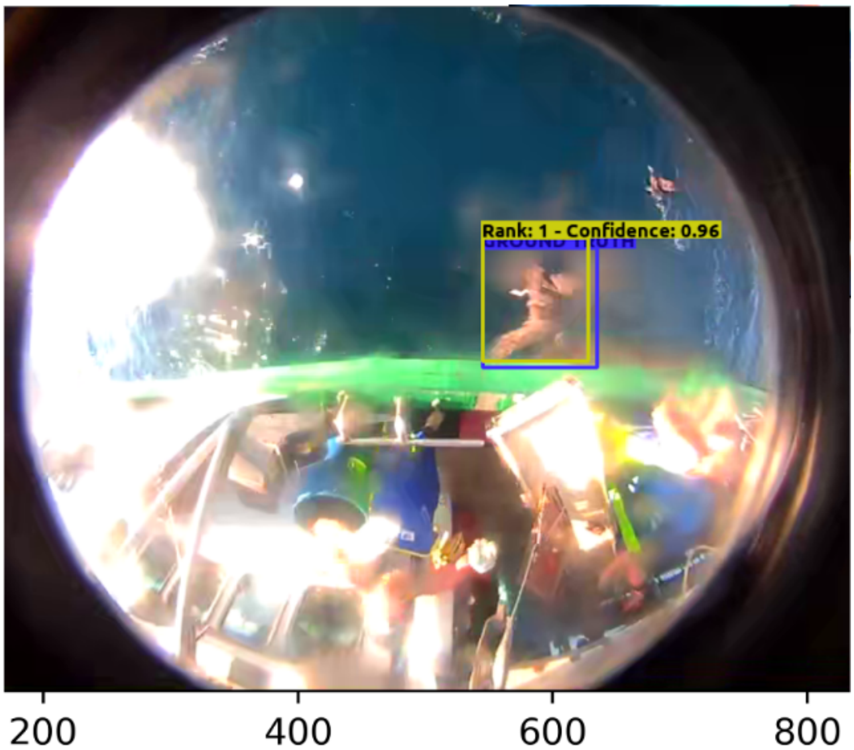


Figure A-10: Examples of annotations added to a single frame of footage captured by an EM system. In this case the original annotation, manually added by a reviewer, is shown as a blue box, overlaid with a further annotation (yellow box) generated by a model developed to identify seabird captures in the footage.

A.1.8 EM data for analyses

A further category of EM data is datasets prepared for analyses. This is the step that ultimately gives rise to observations that are equivalent to the observations that would be collected by a human observer onboard a vessel. The 'raw' EM data, described above, are often not immediately usable – for a variety of reasons. For example, video reviewers would typically stop recording hauling activity by a vessel when the vessel relocates to a different part of the line, or when operations pause due to crew breaks. This fine-scale activity recording would need to be 'rolled up' to the level of the hauling event typically recorded in statutory or observer data. Furthermore, there may be multiple reviews of a slice; the EM data generated from these individual reviews must be reconciled to generate a dataset with unique entries for each event. This is analogous to the process of reconciling data that has been double-entered into a database after recording on paper forms, or producing a final 'agreed age' for a fish from multiple readings of an otolith.

In the case of the petrel project (Figure A-11), multiple reviews of a footage from a haul may be undertaken in order to identify seabird capture events, and to allow for estimation of detection rates. The EM data from these initial reviews are reconciled to identify unique capture events which are then reviewed by a seabird expert who is responsible for assigning the final species identification, the status and fate of the bird, and also confirming whether the event meets the definition of a capture. The expert review data are then matched to fishing events recorded in the statutory ('ERS') data in order to generate an analysis dataset for incorporation in the protected species captures ('PSC') database.

The workflow associated with the generation of final EM datasets in the petrel project is illustrated in Figure A-12. Footage is stored in cloud storage (`AWS`), where it is accessed by the review system used by reviewers to generate EM data. The review data are used to build a reporting database (`ems3`) that sits alongside a reporting database (`kahawai`) built from extracts of statutory data provided by MPI. Code that is used for both review scheduling, and the generation of final datasets, is maintained in the

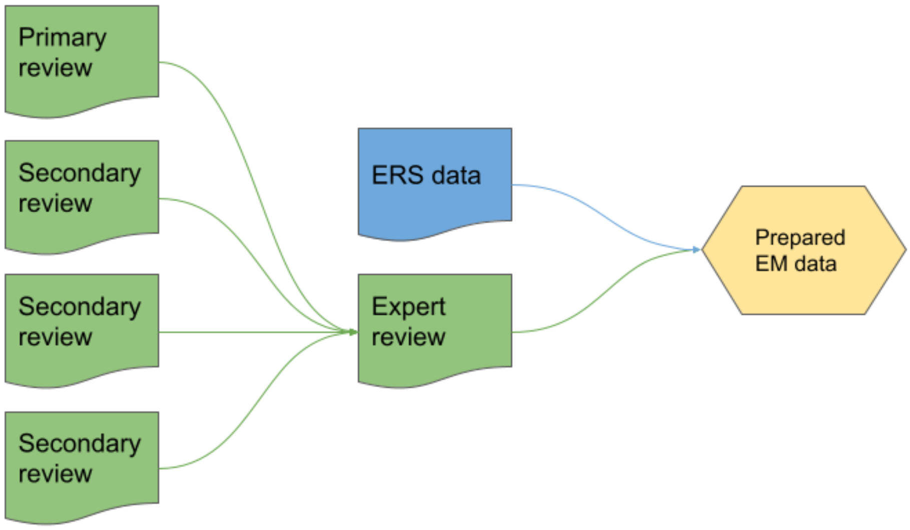


Figure A-11: The review and data reconciliation steps involved in the petrel project.

`github` version control system, and the code is run in the Kahawai Collective's reproducible reporting system (`kahawai.io`). The key feature relevant to the generation of datasets for analysis is that this step is defined in analysis code, maintained on a version control system, and this can be run repeatedly to ensure the reproducibility of the data preparation step.

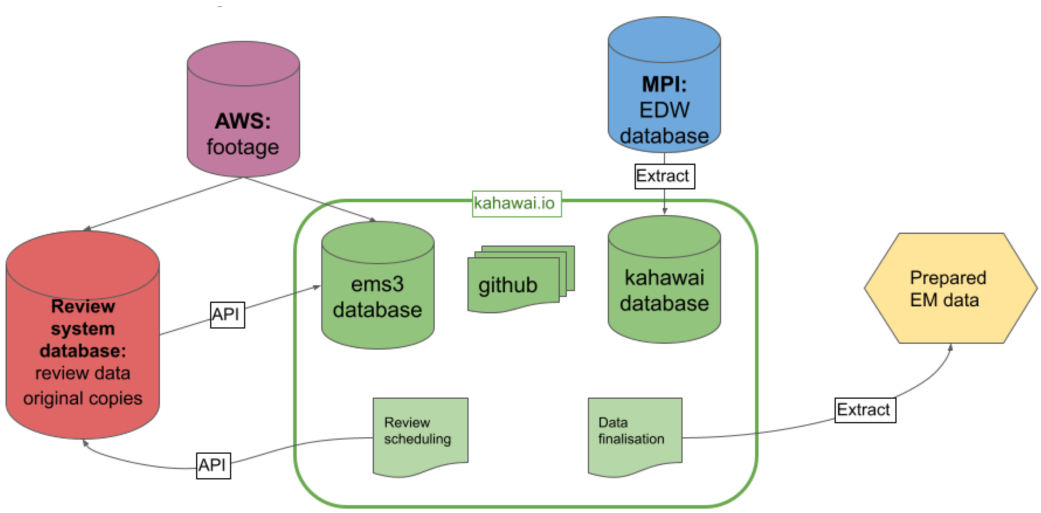


Figure A-12: The workflow involved in the delivery of prepared EM data from the petrel project, illustrating the different systems involved in the storage of generation and storage of data, and the code involved in review scheduling and dataset preparation.

A.2 Workshop discussion

A number of issues discussed in the workshop are briefly summarised here:

- The question of **data ownership and governance** was raised. This was acknowledged to be important and, in the case of the petrel project is discussed further in Appendix B. However, this was considered to be peripheral to the question of the nature and storage of EM footage and data, which are similar whatever the governance arrangements around the programme. The logistics of data storage can be separated from data governance. Data that fall under different governance arrangements can potentially be managed in a common database, so long as there are processes for implementing the requirements of different governance arrangements.

- Fisheries New Zealand noted that the focus of the developing statutory on-board cameras programme is likely to be more on **validation of fisher-reported data** than independent observation. The workshop organisers indicated that the nature of EM footage and much of the resulting EM data was likely to be similar in this situation, with the key differences arising in the selection of data for review, and when the EM data were prepared for analysis alongside other datasets.

- It was noted that **timezone recording** has caused confusion with observer data in the past (with Fisheries New Zealand Observers instructed to record all data in NZST). It was noted that modern databases support a 'timestamp with timezone' data type that can be used to remove ambiguity.

- The importance of **clear terminology** was noted on a number of occasions; for example, it was noted that while footage may provide information on a range of issues, a review protocol is required to provide usable data for a particular analysis. However, a benefit of the footage is that it can potentially be reviewed again to provide different data for other analyses.

- Monitoring of **reviewer performance** was discussed, noting that data-quality metrics are typically not a focus in other data collection programmes (e.g., traditional human observer programmes). It was accepted that monitoring of reviewer performance for data-quality has the potential to overlap with employment issues. Monitoring implemented to data has typically been around the time taken to undertake reviewing; other forms of monitoring are possible, but may be intrusive and need to be developed carefully. However, the barriers are also technological, with limited implementation or these concepts in current review software.

- It was agreed that the reason for selecting a particular slice of footage for review, and the protocol used for that review, are important parts of the **metadata for a review**. For example, the petrel project now distinguishes between primary, secondary and expert reviews, and between slices selected for review at-random, due to fisher-reported captures, or for other reasons.

- It was noted that the choice of the term **annotations** to indicate regions of an image defined in support of machine-learning applications was potentially unhelpful, as other projects used this term to refer to event-type data. It was agreed that terminology had been diverse, and could usefully be standardised. There was some discussion of the relative merits of producing annotations as part of the standard review process, which might add to the workload of reviewers, versus through a specific process in support of machine-learning.

- There was general support for the use of **open standards** for data formats, noting that different uses may need to convert data to specific formats.

- It was noted that different ownership and governance arrangements may apply to different parts of the footage and/or data arising from an EM programme. This influences the question of **what is stored, and where?** It was recognised that management of video observation data has to accommodate the different stages in the data, including raw footage, reviews, and finalised datasets. Participants agreed that storing video observation data alongside traditional observer data in the Centralised Observer Database (COD) would be difficult, and that development of a new database focussed on EM data was appropriate.

- Workshop participants had a range of views on the extent to which **generic or specific data storage solutions** were desirable. From the perspective of the petrel project, data were provided in a format designed to be linked to the PSC database, but these outputs were supported by a much richer underlying set of data and metadata. One option is to simply archive these data in the form they are used within the project, another is to generalise these so that they are useful for other projects using EM systems. Producing generalised and flexible databases comes at a cost, but the absence of general solutions implies that individual programmes are ‘reinventing wheels’ and value in the underlying data may be lost. Project specific data outputs tend to be documented in project specific reports; a more general approach would require the development of specific database documentation.

- Participants were aware of a range of projects where **footage was being analysed to produce data**. Apart from ‘fisheries EM projects’, including the petrel programme, the MPI pilot programme in the West Coast North Island fisheries, and the FINZ tarakihi fishery programme, other examples included citizen science projects and drone footage, such as that collected by the Maui63 programme.

A.3 Recommendations

The workshop conveners noted that recommendations from the workshop were not binding on any party, and that no process or funding had been identified to follow up on the recommendations. There was judged to be consensus around:

- the need for **well-defined and consistently applied terminology**;
- the use of **open standards** for data storage;
- the desirability of **capturing the different stages of EM data** rather than just the final outputs;
- the need to **develop new solutions for long-term storage of video observation data** rather than trying to fit these into existing databases, such as COD.

There was no clear consensus on the extent to which a generic solution for storing video observation data was desirable. It was noted that the petrel project data management experience may provide a good stating point for designing a more generic solution, but would need to consider the experience of similar projects.

A.4 Workshop participants

David Middleton Pisces Research

Edward Abraham Dragonfly Data Science

Finlay Thompson Dragonfly Data Science

Richard Mansfield Dragonfly Data Science

William Gibson Fisheries New Zealand

Christopher Dick Fisheries New Zealand

Dan Kerrigan Fisheries New Zealand

Dave Wetherall Fisheries New Zealand

Tosin Olateju Fisheries New Zealand

Kim George Fisheries New Zealand

Shannon Weaver Department of Conservation

Lyndsey Holland Department of Conservation

Jade Maggs NIWA

Janice Molloy Southern Seabirds Solutions Trust

Richard Wells Deepwater Group

Rosa Edwards Fisheries Inshore New Zealand

Steph Borelle Birdlife International

APPENDIX B: MANAGEMENT AND REUSE OF VIDEO OBSERVATION DATA

B.1 Background

The governance arrangements for the footage and data collected by the petrel project have been more complex than those of many fisheries research projects. There has been a long-standing recognition that, while fisheries data has clear and obvious value for fisheries management, and a range of other uses, it is also commercially sensitive. As a result, the Ministry for Primary Industries has well established processes around the management and use of fisheries data.[7] In particular, fine-scale location information on fishing activity has traditionally been particularly sensitive, and data have typically been both anonymised, and aggregated to coarser temporal and spatial scales, before they are made publicly available.

The collection of video footage on board fishing vessels involves issues of commercial sensitivity and of privacy (Privacy Commissioner 2009). As a result, in the petrel programme, and similar projects where vessel operators have participated on a voluntary basis (Middleton & Guard 2021), the participating vessel operators have been identified as the owners of the footage collected on their vessels and the process of reviewing the footage to create EM data has been conducted according to agreements where the footage owners allow the use of the footage for this particular purpose.

However, it is also recognised that data (in the broad sense, so here including EM footage) may have enduring value, including for purposes other than those for which they were originally collected. Data governance and stewardship processes must find a balance between maintaining data security (both of existing data, and the supply of future data) and facilitating the use of data. One initiative to develop new models for data sharing in New Zealand was the Data Commons Project, whose 'Data Commons Blueprint' (Mansell et al. 2017) observed:

> *Data integration and reuse at scale can create significant value for all parties – data contributors, and data reusers – but only if people can create and maintain a high-trust relationship in regard to the transactions they are participating in.*

More recently, the concept of a 'Data Trust' has emerged (Anonymous 2019, Hardinges et al. 2019). The Open Data Institute define a Data Trust as a legal structure that provides independent stewardship of data. The data trust, which is functionally independent of the institutions that collect and hold data, becomes a steward of the data, taking responsibility to make decisions about the data and ensure they support the data trust's purpose. The concept of such organisations has gained some traction in New Zealand (AI Forum of New Zealand 2019, Whitcroft 2019a, 2019b) but, to date, few concrete examples have emerged.

B.2 Data Trusts

Under the auspices of The Kahawai Collective, the concept of Data Trusts has been developed to provide longer-term stewardship of various datasets originally entrusted to members for particular projects. In many cases these data are provided voluntarily by a range of contributors, and over an extended period. Long term access to these data is required in order to ensure ongoing reproducibility and transparency for the associated data analyses. In addition, the data may be useful in subsequent projects.

Despite the name, trust law is not considered to be an appropriate legal structure for data trusts (Anonymous 2019); however, a legal structure is required. The Kahawai Collective is working towards formalising such an approach, building on agreements developed for use in the US by Brighthive[8]; key components are outlined below.

[7] https://www.mpi.govt.nz/dmsdocument/34803-Guidelines-for-Release-of-Fisheries-Information
[8] https://github.com/brighthive/data-trust-legal

B.2.1 Ethical principles

Brighthive's agreements propose that a first policy adopted by a Data Trust states the ethical principles by which the trust will operate, with the default example proposing the following principles:

Fairness: Understand, mitigate and communicate the presence of bias in both data practice and consumption

Benefit: Set people before data and be responsible for maximizing social benefit and without causing harm

Openness: Practise humility and openness. Transparent practices, community engagement, and responsible communications are an integral part of data ethics. No Member, third party affiliate, or approved data user shall act from a place of political motivation, ties, or secret affiliations

Reliability: Ensure that every effort is made to glean a complete understanding of data, where it came from, and how it was created. Extend this effort for future users of all data and derivative data

Principles of this nature are likely to be a useful foundation, emphasising both established principles of good scientific practice (MPI 2011) and the importance of acting responsibly towards data providers.

B.2.2 Participants and roles

Data Trusts, as currently envisaged, will consist of three types of participant: a **Governance Board** which sets the policy and procedures of the Trust, and supervises the **Trustee** who manages the Trust. The other key participants are the **Trust Members**, who contribute the data managed by the Trust.

Users of the data managed by the Trust and not part of the Trust per se, but nevertheless may include the Trustee and Members of the Data Trust, in addition to approved third parties.

B.2.3 Data Trust Agreement

The legal Agreement establishing the Data Trust is made between the Trustee and the Members of the Data Trust. A key role of the Agreement is establishing the Governance Board for the Data Trust which comprises representatives of the Trustee and of the Members.

The Board is responsible for (i) making decisions around new uses of the Trust's data and (ii) ensuring that the Trustee manages the Trust's data in accordance with the Data Trust Agreement, essentially enforcing the agreement on behalf of the Trust members.

B.2.4 Schedules

A range of schedules to the Data Trust Agreement detail the range of issues that are necessary to address in the establishment of a Data Trust. These include:

- defining the goals of the Trust, ideally with example use cases and projects;
- maintaining a Data Registry of Member-contributed data managed by the Trust, and the default level of access for different classes of these data;
- maintaining a data access register which lists the approved users and uses of the data;

- listing any processing routinely applied to the data managed by the Trust, including processes for data archival and disposal;
- providing technical information on the storage of the Trust data holdings, including procedures relating to the storage of Personal Information;
- listings the Governance procedures that define how the Data Trust Board operates;
- providing a change log recording changes to either the Data Trust Agreement or the Data Registry;
- recording all additional documents (e.g., MoUs, data access agreements) relating to use of the Trust's data;
- listing any fees and charges applying to use of the Trust's data.

B.2.5 Other key considerations

A key reason for establishing a Data Trust is to assist in building the high-trust relationships required to generate value from data (Mansell et al. 2017). However, in establishing a legal entity to underpin these relationships it is necessary to define the recourse (and, conversely, the liability) that parties to the agreement have in the event that the terms are breached, whether intentionally or otherwise. This must balance the need to achieve the goals of the Data Trust while accepting that some failures (e.g., reputational damage or IP leakage) result in losses that are extremely difficult to quantify.

A further consideration, relevant to setting the goals for a Data Trust, is to whom any value accrues from the use of the data. A starting point is the expectation that Members receive the benefit from their data contributions. However, in most cases where fisheries data are shared there is a recognition that the benefits are often indirect, through improved management, rather than immediate financial reward. A more direct issue is that maintaining a Data Trusts results in costs, that will ideally be covered by the operation of the Trust rather than requiring external funding.

B.3 Experience with the petrel project

At the time of writing, the intention to establish a Data Trust for the petrel project footage has been discussed with the footage owners, and received their support. However, the use of the petrel project footage for AI investigations has so far been handled on an *ad hoc* basis. There would need to be greater demand (or similar requirements for other datasets) to justify further legal expenses involved with formally establishing a Data Trust.

The small number of participants has, however, has allowed testing of the concept of a Governance Board. In particular, permission was sought and received to reuse the footage collected during the project to build a training dataset, and to build and test computer vision models, that assist in automating the process of detecting seabird captures in the footage.