Optimising sampling of fish assemblages on intertidal reefs using remote underwater video

Katherine R Erickson [Corresp., 1, 2] , **Ana B Bugnot** [2, 3] , **Will F. Figueira** [2, 4]

[1] Centre for Marine Science and Innovation, University of New South Wales, Sydney, Kensington 2052, Australia

[2] School of Life and Environmental Sciences, University of Sydney, Sydney, Camperdown, 2006, Australia

[3] Oceans and Atmosphere, Commonwealth Scientific and Industrial Research Organisation, Brisbane, St. Lucia, 4072, Australia

[4] Sydney Institute of Marine Science, Sydney, Mosman, 2088, Australia

Corresponding Author: Katherine R Erickson
Email address: k.erickson@unsw.edu.au

Background. Assessing fish assemblages in subtidal and intertidal habitats is challenging due to the structural complexity of many of these systems. Trapping and collecting are regarded as optimal ways to sample these assemblages, but this method is costly and destructive, so researchers also use video techniques. Underwater visual census and baited remote underwater video stations are commonly used to characterise fish communities in these systems. More passive techniques such as Remote Underwater Video (RUV) may be more appropriate for behavioural studies, or for comparing proximal habitats where the broad attraction caused by bait plumes could be an issue. However, data processing for RUVs can be time consuming and create processing bottlenecks.

Methods. Here, we identified the optimal subsampling method to assess fish assemblages on intertidal oyster reefs using RUV footage and bootstrapping techniques. We quantified how video subsampling effort and method (systematic vs random) affect the accuracy and precision of three different fish assemblage metrics; species richness and two proxies for the total abundance of fish, $MaxN_T$ and $MeanCount_T$, which have not been evaluated previously for complex intertidal habitats.

Results. Results suggest that $MaxN_T$ and species richness should be recorded in real time, whereas optimal sampling for $MeanCount_T$ is every 60 seconds. Systematic sampling proved to be more accurate and precise than random sampling. This study provides valuable methodology recommendations which are relevant for the use of RUV to assess fish assemblages in a variety of shallow intertidal habitats.

1 **Optimising sampling of fish assemblages on intertidal reefs using remote**

2 **underwater video**

3 Katherine R. Erickson*[1, 2], Ana B. Bugnot[1,3], Will F. Figueira[1,4]

5 [1] School of Life and Environmental Sciences, University of Sydney, Camperdown 2006,

6 Australia.

7 [2] Centre for Marine Science and Innovation, School of Biological, Earth and Environmental

8 Sciences, University of New South Wales, Kensington 2052, Australia.

9 [3] Oceans and Atmosphere, Commonwealth Scientific and Industrial Research Organisation, St.

10 Lucia, Queensland 4072, Australia.

11 [4] Sydney Institute of Marine Science, Mosman, New South Wales 2088, Australia.

13 *Corresponding Author:

14 Erickson, Katherine[1, 2]

15 77 Margaret St, Petersham, Sydney, 2049, Australia

16 Email address: k.erickson@unsw.edu.au

Abstract

24   **Abstract**

25   **Background.** Assessing fish assemblages in subtidal and intertidal habitats is challenging due to

26   the structural complexity of many of these systems. Trapping and collecting are regarded as

27   optimal ways to sample these assemblages, but these methods are costly and destructive, so

28   researchers also use video techniques. Underwater visual census and baited remote underwater

29   video stations are commonly used to characterise fish communities in these systems. More

30   passive techniques such as Remote Underwater Video (RUV) may be more appropriate for

31   behavioural studies, or for comparing proximal habitats where the broad attraction caused by bait

32   plumes could be an issue. However, data processing for RUVs can be time consuming and create

33   processing bottlenecks.

34   **Methods**. Here, we identified the optimal subsampling method to assess fish assemblages on

35   intertidal oyster reefs using RUV footage and bootstrapping techniques. We quantified how

36   video subsampling effort and method (systematic vs random) affect the accuracy and precision

37   of three different fish assemblage metrics; species richness and two proxies for the total

38   abundance of fish, $MaxN_T$ and $MeanCount_T$, which have not been evaluated previously for

39   complex intertidal habitats.

40   **Results.** Results suggest that $MaxN_T$ and species richness should be recorded in real time,

41   whereas optimal sampling for $MeanCount_T$ is every 60 seconds. Systematic sampling proved to

42   be more accurate and precise than random sampling. This study provides valuable methodology

43   recommendations which are relevant for the use of RUV to assess fish assemblages in a variety

44   of shallow intertidal habitats.

45

46

47

Introduction

48

49   Accurate, rapid and robust quantitative techniques to assess fish communities are of

50   importance to marine ecologists as they provide the basis for ongoing monitoring and assessment

51   of community dynamics. This is especially important for systems in decline, such as shallow,

52   complex estuarine habitats like oyster reefs, mangroves, seagrass beds, and rocky reefs (Gaylard

53   et al. 2020). They provide nursery habitat within the estuary (Nagelkerken et al. 2015) and support

54   diverse fish assemblages that need to be monitored to maintain fishery and estuary health

55   (Whitfield and Elliott 2002). These habitats are also a focus for restoration, and optimising

56   methodology for fish assessments is vital to maximise the uptake of monitoring and ensure

57   restoration success (Bosire et al. 2008, Baggett et al. 2015).

58   In places such as marine reserves, non-invasive estimation methods are required over

59   destructive catch estimations, which can bias future samples and compromise management

60   objectives (Andrew & Mapstone 1987, Willis & Babcock 2000). Researchers thus rely on various

61   non-invasive methods such as diver underwater visual census, netting, mark/recapture, and video

62   (Cappo et al. 2006) to estimate population numbers.

63   Amongst these, baited (BRUV) and unbaited (RUV) remote underwater video systems are

64   especially beneficial due to their suitability for use in areas exposed to strong currents, or waters

65   too shallow or too deep for divers, and for the permanent record they supply (Cappo et al. 2006,

66   Langlois et al. 2010). To date, most studies assessing the applicability of video techniques have

67   focussed on baited videos (Cappo et al. 2006, Langlois et al. 2010, Whitmarsh et al. 2017, Langlois

68   et al. 2020). However, unbaited RUVs are rising in popularity as they can be more appropriate for

69   studies looking at fish behaviour, as they have little to no effect on the community they are

70   recording (Cappo et al. 2006, Mallet & Pelletier 2014). They are also better at capturing non-

71   carnivorous species (Goetze et al. 2015) and for comparing community assemblages between

72   proximal habitats, as it is impossible to know how far the attractant effect of the bait plume extends

73   in the water column (Asher 2017). However, it is recommended to employ more sampling effort

74   for unbaited RUVs than baited, due to lower abundances of fish in the frame — longer soak times,

75   or more replicates, to achieve the same level of statistical power (Bernard & Gotz 2012).

76   In addition, extracting data from the full videos can develop a 'bottleneck' in the analysis

77   workflow due to observer fatigue, with one hour of video taking up to 13 hours to process (Cappo

78   et al. 2006, Campbell et al. 2015). Recently, researchers have been recognising the value of time-

79   saving applications of deep learning to automatically count and identify fish in videos (Christin et

80   al. 2019, Ditria et al. 2021, Lopez-Marcano et al. 2021). However, models do not perform as well

81   in videos with complex backgrounds, poor visibility, differing light conditions, or cryptic

82   camouflaged fish (Salman et al. 2019). Methods are being developed to cope with these conditions

83   (Salman et al. 2019), but automated identification with greater than 70% accuracy is not yet

84   possible for shallow habitats, for example oyster reefs, with all four of these issues. Additionally,

85   each model must be trained with a dataset of thousands of annotated images of each species from

86   the particular habitat (Ditria et al. 2020), which would be extremely time consuming and difficult,

87   especially for less common species. Hence, it is key to identify the most efficient methods to

88   manually count fish in videos like these to ensure accuracy.

89   One way to reduce processing time is by subsampling videos to assess fish communities.

90   Studies have looked at how the number of RUV samples affects the accuracy of data (Garcia et al.

91   2021), and examined optimal soak times (Asher 2017, Wong et al. 2019, Garcia et al. 2021), but

92   no studies have assessed precision and accuracy associated with subsampling fish metrics for

93   RUVs.

94   The most common metrics used to assess fish assemblages from video are number of species

95   captured across the duration of the video (hereafter, species richness) and MaxN, defined as the

96   highest number of individual fish that appear in one frame (Ellis 1995). It does not double count

97   individual fish and thus does not result in overestimations (Cappo et al. 2006). According to the

98   most recent review, it is the primary metric employed to assess fish communities through BRUV

99   footage (81% of studies, Whitmarsh et al. 2017), and has a processing time:video length ratio of

100   0.5:1 (Gladstone et al. 2012). A new alternative metric, MeanCount, has only been used in 2% of

101   BRUV studies but is rising in popularity (Whitmarsh et al. 2017). MeanCount uses either

102   systematically or randomly selected individual frames from across the video to calculate a mean

103   of the number of fish in the frame (Conn 2011). Essentially, it is an occupancy-weighted MaxN

104   and can be a more informative and useful metric to obtain. There is debate over which metric is

105   more accurate, with one laboratory study showing that MeanCount is linearly related to true

106   abundance, whereas MaxN was found to be hyperstable and underestimated true abundance

107   (Schobernd et al. 2014). Other field and simulation studies have shown that it can be less precise

108   than MaxN, and potentially over-inflate zero counts (Stobart et al. 2015, Campbell et al. 2015).

109   Their relative value can change based on useage, as for RUVs in complex habitats there is high

110   correlation between the two for structure-oriented species, but less so for mobile species (Baker et

111   al. 2022).

112   Generally, previous studies have sampled species richness and MaxN across the duration

113   of the video, while MeanCount is calculated from a subsample of frames ranging from 10-150

114   seconds (Cappo et al. 2006, Schobernd et al. 2014, Follana-Berná et al. 2019, Cullen & Stevens

115   2020, Kilfoil et al. 2021). One study, however, has analysed the precision and accuracy associated

116   with subsampling species richness. Bacheler and Shertzer (2015) found that subsampling every 30

117 seconds for 20-minute BRUV footage results in 14% of species being missed. This study counted

118 210 species from 1543 videos. Two studies have compared the precision and accuracy of BRUV

119 data associated with MaxN and subsampling MeanCount. Campbell et al. (2015) found that

120 MeanCount had lower precision than MaxN for 8 economically important focal species, and

121 recommended sampling MeanCount every 15 seconds, despite slight differences between species.

122 Bacheler and Shertzer (2015) looked at three common focal species and advised sampling

123 MeanCount every 50 seconds for each. The Gulf of Mexico where these studies were performed

124 is very diverse, and therefore could have limited relevance to low abundance estuarine habitats,

125 like oyster reefs, that have a smaller species richness. Additionally, these results might not be

126 transferrable to unbaited RUVs, as BRUVs draw in more fish over time, whereas unbaited RUVs

127 have a steadier abundance of fish. One simulation study indicated that for three hours of RUV

128 footage, sampling every two minutes was sufficient for accurate and precise MeanCount data, but

129 this was done for just a single species (Follana-Berná et al. 2019, Follana-Berná et al. 2020).

130 Methods for estimating MaxN and MeanCount should be directly compared to decide on an

131 optimal sampling strategy that maximises efficiency and robustness.

132   Moreover, frames for subsampling can be chosen systematically or randomly across the

133 duration of the video. This has repercussions for the efficiency of data processing, as it could be

134 logistically easier to sample MeanCount systematically (in real time) than randomly if the whole

135 video is going to be watched regardless of other methods. No studies have experimentally

136 compared the two sampling methods, but some papers have suggested using either (Conn 2011,

137 Bacheler & Shertzer 2015). All metrics and both sampling methods should be explored when

138 optimising efficiency as, contrary to previous recommendations, there could be a difference.

139   In this study we used RUV samples of fish assemblages from an intertidal habitat, oyster

140   reefs, to 1) find the optimal soak time for RUVs to obtain accurate values of species richness in

141   these systems, 2) compare the precision and accuracy associated with subsampling MaxN,

142   MeanCount, and species richness using systematic and random methods, and 3) assess the trade-

143   off between effort and information gained to make recommendations about the number of frames

144   to sample.

145

146   **Methods**

147   Data collection for this study was carried out over two days in March 2019 at the Port

148   Hacking oyster reefs, located in Sydney NSW (34°4'25" S, 151°7'7" E). Local species richness and

149   true abundance of the fish community at this study site are not known. Video was collected with

150   permit P03/0029-5.1 issued by the Department of Primary Industries, New South Wales, and under

151   University of Sydney Ethics approval 2019/1571. Unbaited RUV footage was collected by setting

152   out GoPro cameras (wide FOV, 1080HD, 60fps), attached to 50x50x30cm metal frames, to film

153   for 70 minutes over high tide around the reefs (See Supplemental Figure 1). Placement of cameras

154   was split into three zones – Centre (interior of reef, 30cm from edge), Edge (30cm of reef edge

155   and 30cm of surrounding sand flat) and Off (1m off the reef, facing towards empty sand). Visibility

156   was consistent across all videos, and all fish visible were counted. Cameras were placed such that

157   roughly 40% of the field of view was benthic habitat and 60% was water column. Ten of these

158   videos, evenly spread over the three zones (Centre=4, Edge=3, Off=3), were selected for the

159   analysis presented here.

160   The first five minutes of each video was discarded to allow time for snorkelers to exit the

161   area. All frames of the following 60 minutes of footage were fully annotated using EventMeasure

162 software (SeaGis v. 5.12). Videos were watched in real time within the program and paused to

163 record the species, and time, each individual fish entered and left the field of view (time in-time

164 out, or TITO, from Schobernd et al. 2014). Hyper abundant fish (shoals of 200-800 fish) were

165 handled differently as it was impossible to track every individual fish. The total time the shoal was

166 in the frame was calculated, and then the number of fish in the frame was counted at five equal

167 time intervals over this period. Species richness was derived from this dataset, as well as MaxN

168 and MeanCount for each species. The average processing-time:video length ratio for all videos

169 was 8:1.

170   Optimal 'soak time', defined as length of time the camera is deployed and filming, was

171 investigated through species richness. The cumulative number of species seen was taken from

172 TITO data at 1-minute intervals across the 60 minutes of footage for all 10 videos (following

173 Cappo et al. 2006, Schram et al. 2020). A GAM curve and 95% CI were fitted using the *ggplot2*

174 package (v3.3.2; Wickham 2016) to aid visual interpretation.

175   The TITO data was used to create a dataset of the number of fish present for every second of

176 the video using a custom Microsoft Excel script (3600 second-points in total per video). For each

177 of the 10 videos, MaxN, MeanCount, and species richness were estimated by bootstrap sampling

178 this dataset with two subsampling methods, systematic and random, over a series of levels of effort.

179 Both methods used sampling intensities of 360, 180, 120, 60, 30, 20, 15 and 12 sampling points

180 over the 3600-second length of the video. These estimated values were then compared to the 'true'

181 values, where 'true' refers to the accurate values for each metric obtained from the TITO dataset,

182 and not the true abundance or richness of the study site (which is unknown).

183   For systematic sampling, points were distributed at equal intervals through the video (i.e.

184 for 12 sampling points, a frame was taken every 300 seconds over the course of the video). The

185   need for equal sample intervals restricted the number of unique subsets available for each video.

186   The highest sample size of 360 points required a sampling interval of 10 seconds, where there are

187   only ten different possible options within a 60 minute (3600 second) video. Where there were more

188   than ten possible datasets for a given sample size, the 'rand' function in Microsoft Excel (2016)

189   was used to pick a random selection of ten of these. For random sampling, points were drawn at

190   random from all possible points, with replacement, with the process repeated 10 times using the

191   *boot* package (v.1.3-27; Davison & Hinkley 1997, Canty & Ripley 2021) in R version 1.4.1717 (R

192   Core Team 2019).

193   For these two methods, accuracy and precision curves were calculated for each video and the

194   point of diminishing returns was found by averaging over all ten videos. Precision was defined as

195   the coefficient of variation (CV) within each sample size. Accuracy was defined as the percentage

196   absolute deviation of the observed estimate from the 'true' value (as calculated from TITO data).

197
$$Deviation = \frac{|Truth - Estimated|}{Truth} * 100$$

198   Our unbaited videos had low abundance and many zeros, which is to be expected for unbaited

199   videos on reefs with rare, cryptobenthic species. While this is not an issue for community analyses, it

200   does make species-level accuracy and precision estimates highly variable, and hence difficult to

201   evaluate (See Figure 1 for examples of three species, the common mobile species yellowfin bream

202   *Acanthopagrus australis*, the common juvenile tarwhine *Rhabdosargus sarba,* and the cryptobenthic

203   goby *Redigobius macrostoma*). For this reason, we have conducted the accuracy and precision analyses of

204   our abundance proxies by pooling across species to give the metrics total MaxN (MaxN$_T$) and total

205   MeanCount (MeanCount$_T$). Previous studies on fish abundance have combined metrics like MaxN

206   into 'total abundance' to facilitate analysis (Scott et al. 2013, Bond et al. 2018, Piggott et al.

207   2020, Smith et al. 2021).

208   We compared the accuracy and precision of each method and metric by testing for effects of

209   metric, method and sampling effort on CV and deviation, using generalised linear models in the package

210   *glmmTMB* (Brooks et al. 2017) in R 4.0.2. Two schooling species *Ambassis jacksoniensis* and

211   *Atherinomorus vaigiensis* obscured other patterns present in the data, therefore analysis of CV and

212   deviation (not species richness, as this metric was not affected) was repeated with the schooling species

213   removed. Interactions between sampling effort, metric and method were tested, and interpreted using

214   post-hoc comparisons in the package *emmeans* (Lenth et al. 2018).

215   To identify the optimal sampling point, beyond which increasing effort resulted in minimal

216   improvements to accuracy and precision, 95% confidence intervals were calculated for each sampling

217   effort to identify where the errors differed significantly from zero – aka the accurate values obtained from

218   the TITO dataset. However, the lower bounds were not negative when comparing to zero, even at small

219   sample sizes. Therefore, optimal values were assessed visually, based on the rate of curve deceleration.

220

Results

222 **Soak time**

223   A total of 16 different species from 10 different families were observed over the course of the

224   analysis (Table 1). Species richness increased over the duration of the video, with the greatest

225   difference between the first frame and 15 minutes (0-5.5 species), after which the curve began to

226   plateau and gradually increase up to the 60 minute mark (5.5 – 7.5 species) (Figure 2).

227

228 **Deviation from Truth (Accuracy)**

229   All three metrics generally had lower deviation, with smaller standard error, as sampling effort

230   increased, but the effect of the random and systematic methods differed between the metrics.

231   For the all-species analysis (Figure 3A), the systematic method had significantly lower

232   deviation than the random (Table 2). There was a significant interaction between sampling effort

233   and method (Table 2). Post-hoc analysis showed that MeanCount$_T$ had lower deviation than

234   MaxN$_T$ and Species Richness, while there was no difference between MaxN$_T$ and Species Richness

235   (Supplemental Table 1). The differences were more pronounced at low sampling efforts

236   (Supplemental Table 1).

237   Species Richness and MaxN$_T$ showed similar patterns across sampling effort — sampling from

238   12 frames to 120 frames resulted in a sharp 30% drop in deviation and, beyond this, the slope of

239   the curve plateaus yielding a 10% drop in deviation from 60 to 360 frames. For systematic

240   MeanCount$_T$, the metric and method with lowest deviation, increasing sampling from 12 frames

241   to 60 frames resulted in a steep 23% drop in deviation. Further increases in sampling effort to 360

242   frames decreased deviation more gradually by 7%.

243   When schooling species were removed from the analysis, the accuracy for MeanCount$_T$

244   improved between 5-10% across all sampling efforts, whereas MaxN$_T$ did not improve (Figure

245   3B). The systematic method had significantly lower deviation than the random (Table 2). There

246   was a significant interaction between sampling effort and metric (Table 2). Post-hoc analysis

247   showed that MeanCount$_T$ was always more accurate than MaxN$_T$, but this difference was more

248   pronounced at low sampling efforts (Supplemental Table 2).

249   For systematic MeanCount$_T$, increasing sampling from 12 frames to 60 frames returned a 14%

250   drop in deviation and, beyond this, the curve flattens yielding a 5% drop in deviation from 60 to

251   360 frames.

252

253   **Coefficient of Variation (Precision)**

254   All three metrics generally had a lower coefficient of variation (CV), with smaller standard

255   error, as sampling effort increased, but the effect of the random and systematic methods differed

256   between the metrics.

257   For the all-species analysis (Figure 3C) there was a significant interaction between sampling

258   effort and metric, and between sampling effort and method (Table 2). Post-hoc analysis shows that

259   there was no difference between the random and systematic methods for any metric, however there

260   was a strong trend where the systematic method had lower CV than the random for MeanCount$_T$

261   at high sampling efforts (360, 180, 120) (Supplemental Table 3). Systematic MeanCount$_T$ had

262   lower CV than the other approaches at high sampling efforts (360, 180) (Supplemental Table 3).

263   Species richness behaved differently to the abundance metrics as effort decreased, the CV did not

264   rise as sharply (Figure 3C).

265   For systematic MeanCount$_T$, the metric and method with the lowest CV, increasing sampling

266   from 12 frames to 120 frames results in CV dropping steeply from 0.44 to 0.05 and, beyond this,

267   the curve flattens yielding a 0.04 drop in deviation from 120 to 360 frames (Figure 3C). When

268   schooling species were removed from the analysis, the CV improved ~50% across all sampling

269   efforts (Figure 3D). There was a significant interaction found between sampling effort, metric and

270   method (Table 2). Post-hoc analysis shows that at low sampling efforts, there was no difference

271   between the metrics or methods (Supplemental Table 4). At high sampling efforts, MeanCount$_T$

272   was more accurate than MaxN$_T$, and the systematic method was more accurate than random for

273   MeanCount$_T$ only (Supplemental Table 4). The curve of the systematic MeanCount$_T$ approach did

274   not have a clear inflexion point (Figure 3D).

275

276

Discussion

277

278   Robust methodologies for monitoring fish assemblages are of critical importance for

279 maintaining the health, and improving the restoration success, of declining estuarine habitats such

280 as oyster reefs (Whitfield and Elliott 2002, Baggett et al. 2015, Gaylard et al. 2020). Unbaited

281 videos are the most appropriate technique for these habitats as they overcome scientific and ethical

282 issues associated with destructive or invasive sampling (such as trapping, UVCs or BRUVs) that

283 can stress fish and influence their behaviour and home range (Cappo et al. 2006, Mallet & Pelletier

284 2014, Asher 2017). This study provides useful insights into the differences in trade-offs between

285 effort and accuracy/precision for subsampling three fish community metrics, extracted from

286 unbaited underwater video focussed on characterising fish assemblages on intertidal oyster reefs.

287 While all three metrics were more accurate and precise as sampling intensity increased, they each

288 behaved differently across the sampling intervals. Subsampling $MeanCount_T$ tended to be more

289 accurate and precise than $MaxN_T$ or species richness, but differences in precision faded at low

290 sampling efforts. Accuracy and precision tended to increase when using the systematic

291 subsampling method (over the random) and when removing schooling species. We did not

292 compare $MaxN_T$ and $MeanCount_T$ directly, or comment on the accuracy of any of the metrics, as

293 the true abundance and local richness of the study site is unknown.

294   The species accumulation curve began to plateau at 15 minutes, at the point of 75% species

295 detection, matching the findings of a previous study on RUVs (Asher 2017). Species saturation

296 was not achieved by 60 minutes, due to the difficulties of spotting the less common goby and

297 blenny species such as *Redigobius macrostoma* and *Omobranchus rotundiceps*. Other methods

298 such as UVC may be more appropriate for accurately recording the abundance of cryptobenthic

299 species (Watson et al. 2005), however these methods are not usually possible in the shallow water

300   of intertidal habitats such as oyster reefs. In the videos, these species camouflage into the

301   background, and are far less abundant than oyster blennies (*Omobranchus anolius)* or bream

302   (*Acanthopagrus australis*). This would suggest, noting the large range of values over the ten

303   videos, that the optimal soak time for a general question regarding abundant or fishery-dependant

304   species, excluding cryptobenthics, would be 15 minutes. 60 minutes or more would be appropriate

305   for studies looking at less abundant species. This is similar to recommendations of 60-90 minutes

306   for BRUVS in seagrass, where the complexity of habitat and cryptic behaviour of species requires

307   a longer soak time (Gladstone et al. 2012). Estimates for species richness steadily became more

308   accurate and precise until a sample size of 120, with more modest benefits after this. If deviation

309   is to be kept under 10% and CV under 0.1 (in order to find less common species, for instance

310   camouflaged blennies and gobies), then we recommend not subsampling videos to capture species

311   richness. In a study focussed on abundant species, 120 subsamples would minimise error to 20%.

312   The random method tended to be less accurate, but not significantly so, therefore we recommend

313   systematic (sampling every 30 seconds) for ease of workflow. Bachelor and Shertzer's (2015)

314   study on BRUVs found a similar result for the comparatively lower effort of 50 subsamples. This

315   is likely due to unbaited cameras having a lower probability of filming less common species than

316   baited, so RUVs need more effort to capture true species richness. Additionally, their study was

317   based in the Gulf of Mexico, with a much higher overall species richness (210 vs 16) which is

318   likely to affect the sampling effort needed.

319   Overall, subsampling $MaxN_T$ led to lower precision and accuracy than subsampling

320    $MeanCount_T$. For example, at 60 subsamples, there is 9% variation from the 'true' TITO value for

321    $MeanCount_T$ and 32% for $MaxN_T$. Even at the highest subsampling effort (360 frames), deviation

322   for $MaxN_T$ already reaches 16%, which is why this study recommends recording $MaxN_T$ in real

323   time. No previous studies have attempted to subsample MaxN, but our result is reasonable in the

324   context of MaxN_T being a single value for the whole video – the real value may only occur once

325   or twice. MeanCount_T instead is an average of multiple values across the video and so is more

326   robust to variation caused by the subsampling.

327   Systematic video subsampling was more accurate than random for the MeanCount_T metric

328   (Table 2). We suggest the reason random video subsampling was less accurate is because of the

329   temporally correlated nature of fish presence. Some of the bootstrap replicates from random

330   sampling were entirely in the first or second half of the video, thus missing large groups of fish

331   that occurred, and resulting in higher variation. Systematic sampling has a higher chance of

332   catching these groups of fish that are only in the frame for a few minutes. This pattern can be seen

333   in the MaxN_T accuracy curve, though not as strongly, as MaxN_T is one value rather than an average

334   of values over time, and thus is less affected by temporal correlation. Therefore this study

335   recommends systematically sampling videos for MeanCount_T, rather than randomly. It was not

336   previously known whether systematic or random sampling had lower error, and it was

337   recommended to use either (Bacheler & Shertzer 2015). This recommendation is convenient as

338   systematic sampling is logistically simpler, being more easily accommodated in existing

339   laboratory video analysis work flow, and corresponds to methods in previous papers (Conn 2011,

340   Bacheler et al. 2017). Considering the higher accuracy and precision of systematic sampling for

341   the metric of MeanCount_T, we can recommend subsampling videos to calculate this metric to

342   reduce processing time. The 'inflexion point' on the curve for accuracy of all species (Figure 3A)

343   is at 60 frames, where reducing sampling effort further results in a sharp decrease in accuracy, and

344   an increase in variability (larger standard error bars). Analysis that excluded schooling species

345   (Figure 3B) had higher precision for both MaxN_T and MeanCount_T, and higher accuracy for

346   MeanCount$_T$. Excluding these species increases accuracy overall, as the small chance of missing

347   the schooling species results in a huge error (i.e. getting a MeanCount$_T$ of 3, instead of 250). The

348   'inflexion point' on the curve for accuracy is between 30 and 60 frames, but the variability

349   (standard error bar) is higher for 30 frames. The pattern is similar for precision, where reducing

350   effort to less than 60 frames results in a sharp decrease in precision and an increase in variability

351   (larger standard error bars). To minimise effort, and keep error low at ~6% and CV at 0.1, we

352   recommend sampling every 60 frames (see Table 3 for a summary).

353   Due to the true values being obtained with a TITO approach, the cost in time per frame cannot

354   be calculated for each of the sampling intensities conducted here. However, the recommended

355   method for video analysis was tested in a later study on oyster reefs (C. Pine unpublished data)

356   which found that sampling every 60 seconds for MeanCount$_T$, simultaneously recording MaxN$_T$

357   and species richness, resulted in a reduction of processing time:video length ratio from 8:1 to 1:1,

358   with only a small percentage increase in error. The metrics of MaxN$_T$ and species richness had an

359   even lower ratio of 0.5:1, as the entire video can be watched at 2x speed for RUV data in these low

360   abundance habitats. A recent study on RUVs suggested using Frequency of Occurrence, a

361   presence/absence metric derived from species richness, in situations where a quick and robust

362   assessment of fish assemblage composition is required (Baker et al. 2022). Researchers could

363   consider this method over subsampling when fast data processing is required.

364   Accuracy and precision estimates are typically directly related to the relative frequency of

365   occurrence of the subject of the monitoring (Lechene et al. 2018). It would thus be ideal to estimate

366   sampling effort needed for individual species within this dataset. Unfortunately, this was not

367   possible at the species level in this study given the low abundance within our RUV footage, as it

368   was unbaited in a complex habitat with multiple rare, cryptobenthic species. The high frequency

369   of zero-values resulted in erratic accuracy and precision curves (see Figure 1) at the species level.

370   Therefore, our recommendations are based on proxies of total fish abundance for all species, and

371   researchers interested in accuracy and precision for individual species, especially the more

372   uncommon species, would expect to need higher effort. Subsampling these species would result in

373   larger errors, and we caution researchers to account for the uncertainty arising from this in any

374   estimations and expectations calculated from this methodology.

375   The low abundance and richness of these videos means that our recommendations are

376   applicable for researchers studying unbaited videos in similar low abundance habitats, such as

377   estuarine mangroves, rocky reefs, seagrass and bivalve reefs. Rapid and robust assessments of fish

378   assemblages in these important habitats will not only assist in developing management strategies

379   to monitor and maintain the health of our estuaries (Whitfield and Elliott 2002), but are also

380   necessary for successful restoration by helping to define goals and evaluate the progress of projects

381   (Bosire et al. 2008, Baggett et al. 2015).

Conclusions

383   Estimating the true abundance of fish species is of critical importance for the monitoring and

384   management of fish communities. However, data processing time constraints are a significant

385   limitation for using video surveys to estimate fish abundance. Therefore, knowing precise, accurate

386   and efficient ways to subsample this data is essential. This study has generated guidelines for

387   subsampling unbaited remote underwater video on oyster reefs, or other similarly low abundance

388   complex intertidal habitats, for three different community metrics. In summary, we recommend a

389   workflow where the video is watched in real time (or potentially at 2x speed) to estimate species

390   richness and total MaxN (MaxN$_T$), while total MeanCount (MeanCount$_T$) is estimated every 60

391   seconds. This reduces the data processing effort from eight hours per hour of footage down to one

392 hour, while maintaining low errors in accuracy and precision. These guidelines will assist

393 researchers to evaluate fish community dynamics quickly and efficiently, thus promoting

394 successful monitoring and restoration of threatened estuarine habitats like oyster reefs, mangroves

395 and seagrasses.

396 **Acknowledgements**

397 We specially thank the researchers, students and volunteers involved in the fieldwork.

398 **Authors Contribution Statement**
399
400 KE and WF conceived the ideas and participated in the design of methodology and

401 collection of data, KE, WF and AB analysed the data, KE led the writing of the manuscript. All

402 authors contributed critically to the drafts and gave final approval for publication.

403 **Conflict of Interest**

404 The authors have no conflicts of interest to declare.

405 **References**
406
407 Andrew, N., and B. Mapstone. 1987. Sampling and the description of spatial pattern in marine

408 ecology. Oceanography and Marine Biology 25:13–90.

409 Asher, J. 2017. A Deeper Look at Hawaiian Coral Reef Fish Assemblages: A Comparison of

410 Survey Approaches and Assessments of Shallow to Mesophotic Communities. PhD Thesis.

411 Curtin University.

412 Bacheler, N. M., and K. W. Shertzer. 2015. Estimating relative abundance and species richness

413 from video surveys of reef fishes. Fishery Bulletin 113:15–26. 10.7755/FB.113.1.2

414 Bacheler, N., N. Geraldi, M. Burton, R. Muñoz, and G. Kellison. 2017. Comparing relative

415 abundance, lengths, and habitat of temperate reef fishes using simultaneous underwater

416   visual census, video, and trap sampling. Marine Ecology Progress Series 574:141–155.

417   10.3354/meps12172

418   Baker, R., D. Bilbrey, A. Bland, F. D'Alonzo, H. Ehrmann, S. Havard, Z. Porter, S. Ramsden,

419   and A.R. Rodriguez. 2022. Underwater Video as a Tool to Quantify Fish Density in

420   Complex Coastal Habitats. Diversity 14:50. 10.3390/d14010050

421   Bernard, A. T. F., & A. Götz. 2012. Bait increases the precision in count data from remote

422   underwater video for most subtidal reef fish in the warm-temperate Agulhas bioregion.

423   Marine Ecology Progress Series, 471, 235–252. https://doi.org/10.3354/MEPS10039

424   Bosire, J.O., Dahdouh-Guebas, F., Walton, M., Crona, B.I., Lewis Iii, R.R., Field, C., Kairo, J.G.

425   and N. Koedam. 2008. Functionality of restored mangroves: a review. Aquatic Botany 89:2,

426   251-259.

427   Bond, T., Partridge, J. C., Taylor, M. D., Cooper, T. F., and D.L. McLean. 2018. The influence

428   of depth and a subsea pipeline on fish assemblages and commercially fished species. PLOS

429   ONE, 13(11), e0207703. https://doi.org/10.1371/JOURNAL.PONE.0207703

430   Brooks, M.E., K. Kristensen, K.J. van Benthem, A. Magnusson, C.W. Berg, A. Nielsen, H.J. Skaug, M.

431   Maechler, and B.M. Bolker. 2017. Modeling zero-inflated count data with glmmTMB. BioRxiv,

432   132753.

433   Campbell, M. D., A. G. Pollack, C. T. Gledhill, T. S. Switzer, and D. A. DeVries. 2015.

434   Comparison of relative abundance indices calculated from two methods of generating video

435   count data. Fisheries Research 170:125–133. 10.1016/j.fishres.2015.05.011

436   Canty, A., and B.D. Ripley BD. 2021. boot: Bootstrap R (S-Plus) Functions. R package version

437   1.3-28.

438   Cappo, M., E. S. Harvey, and M. Shortis. 2006. Counting and measuring fish with baited video

439   techniques - an overview. Australian Society for Fish Biology 1:101–114.

440 Christin, S., Hervet, É., and N. Lecomte. Applications for deep learning in ecology. 2019.

441   Methods Ecol Evol. 10: 1632– 1644. https://doi.org/10.1111/2041-210X.13256

442 Conn, P. B. 2011. An evaluation and power analysis of fishery independent reef fish sampling in

443   the Gulf of Mexico and U.S. South Atlantic. National Oceanic and Atmospheric

444   Administration.

445 Cullen, D.W. and B.G. Stevens. 2020. A brief examination of underwater video and hook-and-

446   line gears for sampling black sea bass (*Centropristis striata)* simultaneously at 2 Mid-

447   Atlantic sites off the Maryland coast. J. Northw. Atl. Fish. Sci. 51: 1–13.

448   10.2960/J.v51.m725

449 Davison, A. C., and D. V. Hinkley. 1997. Bootstrap Methods and Their Applications. Cambridge

450   University Press, Cambridge.

451 Ditria, E.M., Sievers, M., Lopez-Marcano, S., Jinks, E. and R. M Connolly. 2020. Deep learning

452   for automated analysis of fish abundance: the benefits of training across multiple habitats.

453   Environmental Monitoring and Assessment 192 (11): 1-8.

454 Ditria, E.M., Connolly, R.M., Jinks, E., and S. Lopez-Marcano. 2021. Annotated Video Footage

455   for Automated Identification and Counting of Fish in Unconstrained Seagrass Habitats.

456   Frontiers in Marine Science 8: 160.

457 Ellis, M. D. 1995. Evaluation of video camera technique for indexing abundances of juvenile

458   pink snapper *Pristipomoides filamentosus,* and other Hawaiian insular shelf fishes. Fishery

459   Bulletin 93:67–77.

460 Follana-Berná, G., M. Palmer, A. Campos-Candela, P. Arechavala-Lopez, C. Díaz Gil, J. Alós, I.

461   Catalán, S. Balle, J. Coll, G. Morey, F. Verger, and A. Grau. 2019. Estimating the density of

462   resident coastal fish using underwater cameras: Accounting for individual detectability.

463   Marine Ecology Progress Series 615. 10.3354/meps12926

464 Follana-Berná, G., M. Palmer, A. Lekanda-Guarrotxena, G. Amalia, and P. Arechavala-Lopez.

465   2020. Fish density estimation using unbaited cameras: Accounting for environmental-

466   dependent detectability. Journal of Experimental Marine Biology and Ecology 527:151376.

467   10.1016/j.jembe.2020.151376

468 French, B., S. Wilson, T. Holmes, A. Kendrick, M. Rule, and N. Ryan. 2021. Comparing five

469   methods for quantifying abundance and diversity of fish assemblages in seagrass

470   habitat. Ecological Indicators 124:107415. 10.1016/j.ecolind.2021.107415

471 Garcia, G. S., M.S. Dias, and G.O. Longo. 2021. Trade-off between number and length of remote

472   videos for rapid assessments of reef fish assemblages. Journal of Fish Biology 99(3): 896–

473   904. 10.1111/jfb.14776

474 Gaylard, S., Waycott, M. and P. Lavery, P. 2020. Review of coast and marine ecosystems in

475   temperate Australia demonstrates a wealth of ecosystem services. Frontiers in Marine

476   Science 7:453.

477 Gladstone, W., S. Lindfield, M. Coleman, and B. Kelaher. 2012. Optimisation of baited remote

478   underwater video sampling designs for estuarine fish assemblages. Journal of Experimental

479   Marine Biology and Ecology 429:28–35. 10.1016/j.jembe.2012.06.013

480 Goetze, J. S., Jupiter, S. D., Langlois, T. J., Wilson, S. K., Harvey, E. S., Bond, T., and W.

481   Naisilisili. 2015. Diver operated video most accurately detects the impacts of fishing within

482   periodically harvested closures. Journal of Experimental Marine Biology and Ecology, 462,

483   74–82. https://doi.org/10.1016/J.JEMBE.2014.10.004.

484 Kilfoil, J.P., M. Campbell, M. Heithaus, and Y.Zhang. 2021. The influence of shark behavior

485   and environmental conditions on baited remote underwater video survey results. Ecological

486   Modelling 447:109507. 10.1016/j.ecolmodel.2021.109507.

487   Langlois, T. J., E. S. Harvey, B. Fitzpatrick, J. J. Meeuwig, G. Shedrawi, and D. L. Watson.

488   2010. Cost-efficient sampling of fish assemblages: Comparison of baited video stations and

489   diver video transects. Aquatic Biology 9:155–168. 10.3354/ab00235

490   Langlois, T., Goetze, J., Bond, T., et al. 2020. A field and video annotation guide for baited

491   remote underwater stereo-video surveys of demersal fish assemblages. Methods in Ecology

492   and Evolution, 11: 1401– 1409. https://doi.org/10.1111/2041-210X.13470

493   Lenth, R., H. Singmann, J. Love, P. Buerkner, and M. Herve. 2018. Emmeans: Estimated

494   marginal means, aka least-squares means. R package version, 1, 3.

495   Lopez-Marcano, S., Jinks, E., Buelow, C., Brown, C. J., Wang, D., Kusy, B., Ditria, E.,

496   Connolly, R., Buelow, C. A., and Connolly, R. M. 2021. Automatic detection of fish and

497   tracking of movement for ecology. Authorea Preprints.

498   https://doi.org/10.22541/AU.161519961.19174880/V1

499   Mallet, D., and D. Pelletier. 2014. Underwater video techniques for observing coastal marine

500   biodiversity: A review of sixty years of publications (1952-2012). Fisheries Research

501   154:44–62. 10.1016/j.fishres.2014.01.019

502   Nagelkerken, I., Sheaves, M., Baker, R., and R.M. Connolly. 2015. The seascape nursery: a

503   novel spatial approach to identify and manage nurseries for coastal marine fauna. Fish and

504   Fisheries 16 (2):362–371.

505   Piggott, C. V. H., Depczynski, M., Gagliano, M., and T. Langlois. 2020. Remote video methods

506   for studying juvenile fish populations in challenging environments. Journal of Experimental

507   Marine Biology and Ecology, 532, 151454. https://doi.org/10.1016/J.JEMBE.2020.151454

508   R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna,

509   Austria. URL http://www.R-project.org/.

510   Salman, A., Maqbool, S., Khan, A. H., Jalal, A., and F. Shafait. 2019. Real-time fish detection in

511   complex backgrounds using probabilistic background modelling. Ecological Informatics,

512   51, 44–51.

513   Schobernd, Z. H., N. M. Bacheler, and P. B. Conn. 2014. Examining the utility of alternative

514   video monitoring metrics for indexing reef fish abundance. Canadian Journal of Fisheries

515   and Aquatic Sciences 71:464–471. 10.1139/cjfas-2013-0086

516   Scott, M. E., Smith, J. A., Lowry, M. B., Taylor, M. D., Suthers, I. M., Scott, M. E., Smith, J. A.,

517   Lowry, M. B., Taylor, M. D., and I. Suthers. 2015. The influence of an offshore artificial

518   reef on the abundance of fish in the surrounding pelagic environment. Marine and

519   Freshwater Research, 66(5), 429–437. https://doi.org/10.1071/MF14064

520   Smith, S. M., Malcolm, H. A., Marzinelli, E. M., Schultz, A. L., Steinberg, P. D., and A. Vergés.

521   2021. Tropicalization and kelp loss shift trophic composition and lead to more winners than

522   losers in fish communities. Global Change Biology, 27(11), 2537–2548.

523   https://doi.org/10.1111/GCB.15592

524   Stobart, B., J. A. García-Charton, C. Espejo, E. Rochel, R. Goñi, O. Reñones, A. Herrero, R.

525   Crec'hriou, S. Polti, C. Marcos, S. Planes, and A. Pérez-Ruzafa. 2007. A baited underwater

526   video technique to assess shallow-water Mediterranean fish assemblages: Methodological

527   evaluation. Journal of Experimental Marine Biology and Ecology 345:158–174.

528   10.1016/j.jembe.2007.02.009

529   Stobart, B., D. Díaz, F. Álvarez, C. Alonso, S. Mallol, and R. Goñi. 2015. Performance of Baited

530   Underwater Video: Does It Underestimate Abundance at High Population Densities? PLOS

531   ONE 10:1–18. 10.1371/journal.pone.0127559

532   Watson, D.L., Harvey, E.S., Anderson, M.J. and G. Kendrick. 2005. A comparison of temperate

533   reef fish assemblages recorded by three underwater stereo-video techniques. Marine

534   Biology 148:415–425. https://doi.org/10.1007/s00227-005-0090-6

535   Watson, D. L., and E. S. Harvey. 2007. Behaviour of temperate and sub-tropical reef fishes

536   towards a stationary SCUBA diver. Marine and Freshwater Behaviour and Physiology

537   40:85–103. 10.1080/10236240701393263

538   Whitfield, A.K. and M. Elliott. 2002. Fishes as indicators of environmental and ecological

539   changes within estuaries: a review of progress and some suggestions for the future. Journal

540   of Fish Biology 61: 229-250.

541   Whitmarsh, S. K., P. G. Fairweather, and C. Huveneers. 2017. What is Big BRUVver up to?

542   Methods and uses of baited underwater video. Reviews in Fish Biology and Fisheries

543   27:53–73. 10.1007/s11160-016-9450-1

544   Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

545   ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

546   Willis, T. J., and R. C. Babcock. 2000. A baited underwater video system for the determination

547   of relative density of carnivorous reef fish. Marine and Freshwater Research 51:755–763.

548   10.1071/MF00010

549   Wong, M. Y. L., P. Gordon, K. C. Paijmans, and M. J. Rees. 2019. Finding rockpool fishes: a

550   quantitative comparison of non-invasive and invasive methods for assessing abundance,

551   species richness and assemblage structure. Environmental Biology of Fishes 102:81–94.

552   10.1007/s10641-019-0846-3

553
554

List of species found in the videos.

Order	Family	Species
Atheriniformes	Atherinidae	*Atherinomorus vaigiensis*
Beloniformes	Belonidae	*Tylosurus gavialoides*
Mugiloformes	Mugilidae	*Mugil cephalus*
Perciformes	Ambassidae	*Ambassis jacksoniensis*
	Blenniidae	*Omobranchus anolius*
		Omobranchus rotundiceps
	Gerreidae	*Gerres subfasciatus*
	Girrellidae	*Girella tricuspidata*
	Gobiidae	*Arenigobius bifrenatus*
		Cryptocentroides gobioides
		Favonigobius exquisitus
		Redigobius macrostoma
	Sillaginidae	*Sillaginodes punctatus*
		Sillago ciliata
	Sparidae	*Acanthopagrus australis*
		Rhabdosargus sarba

Results of generalised linear models for the effects of sampling effort, metric and method on deviation (accuracy) and coefficient of variation (precision).

Variable	χ^2	Df	p.value
Deviation from truth – all species			
i	191.487	1	**<0.001**
Method	5.145	1	**0.023**
Metric	135.847	2	**<0.001**
i:Method	0.016	1	0.899
i:Metric	12.397	2	**0.002**
Method:Metric	0.090	2	0.956
i:Method:Metric	0.141	2	0.932
Deviation from truth – shoaling species removed			
i	321.399	1	**<0.001**
Method	5.017	1	**0.025**
Metric	867.704	1	**<0.001**
i:Method	0.337	1	0.562
i:Metric	78.062	1	**<0.001**
Method:Metric	0.154	1	0.695
i:Method:Metric	0.070	1	0.791
Coefficient of variation – all species			
i	147.451	1	**<0.001**
Method	8.501	1	**0.004**
Metric	55.475	2	**<0.001**
i:Method	6.133	1	**0.013**
i:Metric	13.953	2	**0.001**
Method:Metric	3.519	2	0.172
i:Method:Metric	4.264	2	0.119
Coefficient of variation – shoaling species removed			
i	206.075	1	**<0.001**
Method	5.760	1	**0.016**
Metric	37.941	1	**<0.001**
i:Method	1.668	1	0.197
i:Metric	11.779	1	**<0.001**
Method:Metric	5.845	1	**0.016**
i:Method:Metric	11.243	1	**<0.001**

Summary of recommendations for subsampling unbaited RUV in shallow intertidal habitats.

Metric	Recommendation
Soak time	15 minutes for abundant sp. >60 minutes for rare sp.
Species richness	Every 20 seconds for abundant sp. Real time (no subsampling) for rare sp.
$MaxN_T$	Real time (no subsampling)
$MeanCount_T$	Every 60 seconds
Method	Systematic better than random

Figure 1

Mean (±SE) accuracy and precision of MaxN and MeanCount for three individual species.

Acanthopagrus australis accuracy (A) and precision (D), *Rhabdosargus sarba* accuracy (B) and precision (E), and *Redigobius macrostoma* accuracy (C) and precision (F).
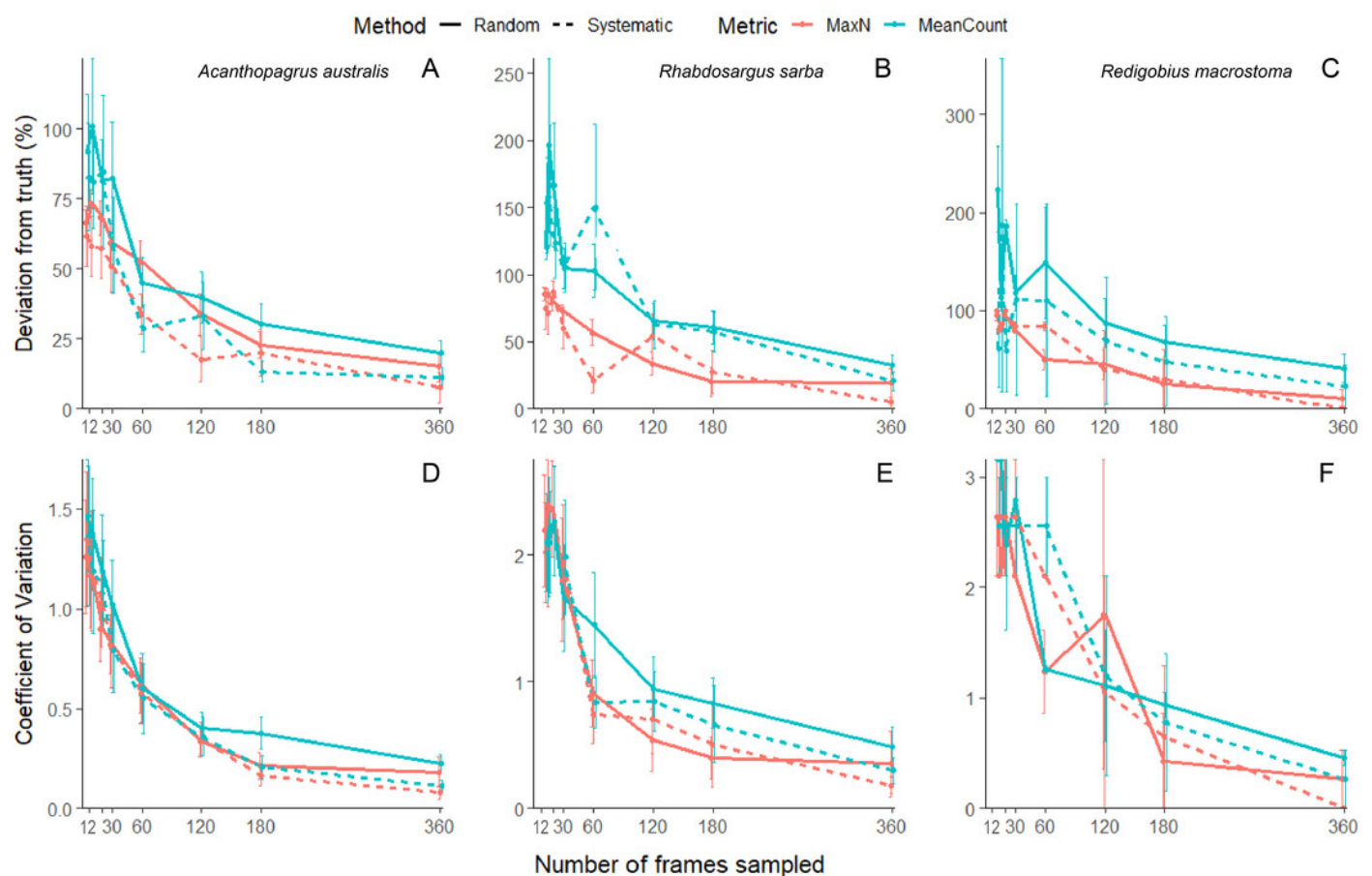
Figure 2

Mean species accumulation curve over the duration of all ten videos.

Colour represents the ten different videos. A GAM was fitted to the points to assist visual interpretation, grey band shows 95% confidence interval.
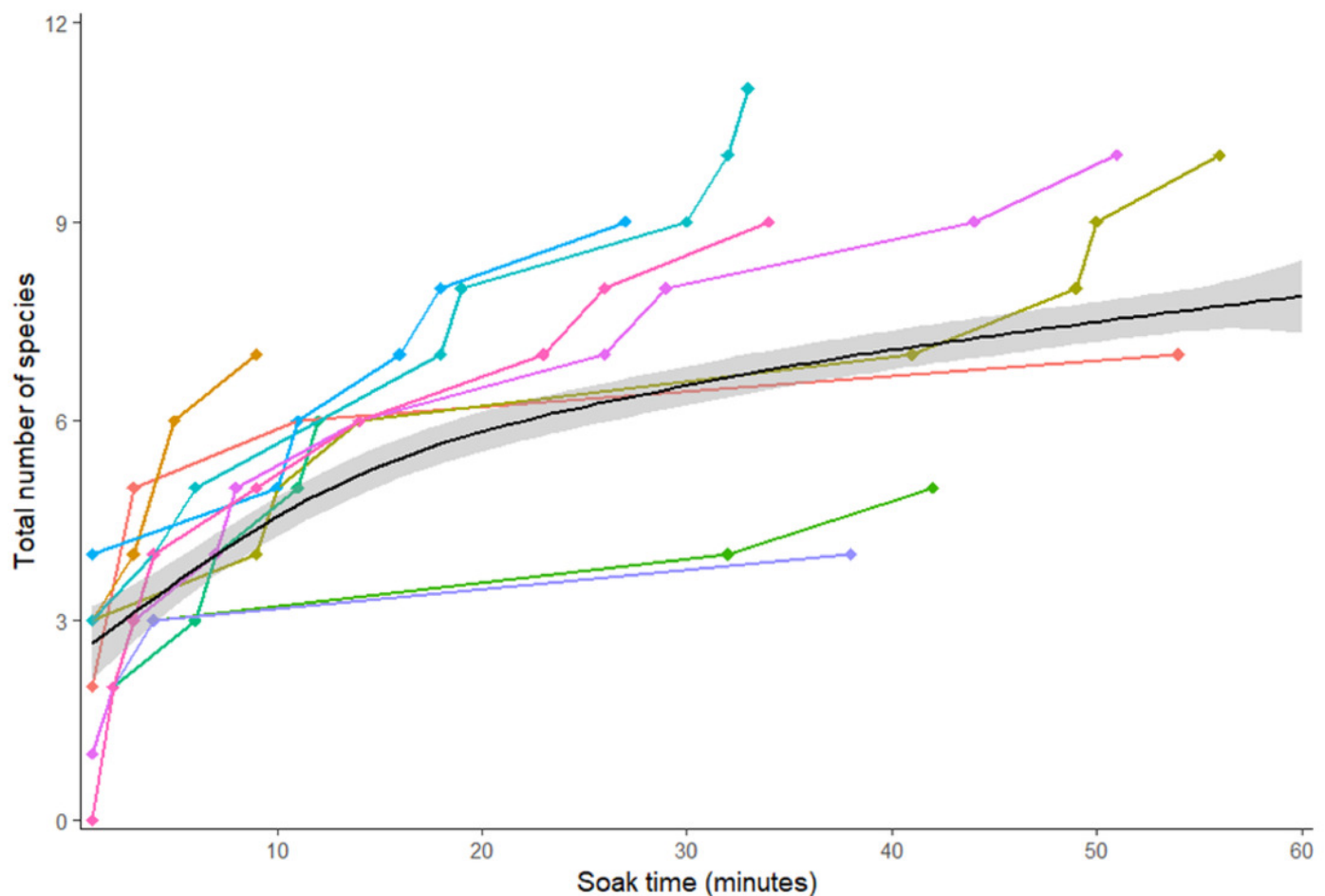
Figure 3

Mean (±SE) accuracy and precision of Species Richness, $MaxN_T$ and $MeanCount_T$.

All species accuracy (A) with shoaling species removed (B), all species precision (C) with shoaling species removed (D).