Examining the utility of alternative video monitoring metrics for indexing reef fish abundance

Zeb H. Schobernd, Nathan M. Bacheler, and Paul B. Conn

SEDAR41-RD19

23 May 2014

NRC Research Press

ARTICLE

Examining the utility of alternative video monitoring metrics for indexing reef fish abundance

Zeb H. Schobernd, Nathan M. Bacheler, and Paul B. Conn

Abstract: Underwater video has become an important tool for monitoring reef fish populations worldwide because it is nonextractive and not strongly selective. A variety of approaches have been developed to enumerate fish on videos, but to our knowledge these metrics have not been tested to determine if they are proportional to true abundance. We compared the most commonly used metric, MaxN (i.e., the maximum number of fish in a single frame during the viewing interval), to a newly developed metric, MeanCount (i.e., the mean number of fish observed in a series of snapshots over a viewing interval), using simulations, a laboratory experiment, and an empirical study. MaxN was nonlinearly related to true abundance using all three approaches, providing increasingly dampened estimates of abundance with increasing true abundance (i.e., hyperstability). Therefore, MaxN may result in positively biased indices of abundance for declining fish stocks or negatively biased abundance indices when fish stocks are increasing. Alternatively, MeanCount was generally linearly related to true abundance and its variability was similar to MaxN, suggesting that MeanCount can be useful for indexing abundance of fish in underwater video surveys.

Résumé : La vidéo sous-marine est devenue un important outil pour la surveillance des populations de poissons récifaux à l'échelle planétaire en raison de son caractère non extractif et peu sélectif. Si différentes approches ont été mises au point pour dénombrer les poissons dans les vidéos, à notre connaissance, ces méthodes de mesure n'ont pas été testées pour déterminer si les résultats qu'elles donnent sont proportionnels à l'abondance réelle. Nous avons comparé la mesure la plus couramment utilisée, le MaxN (c.-à-d. nombre maximum de poissons dans une image donnée durant l'intervalle de visionnement), à une mesure nouvellement mise au point, MeanCount (c.-à-d. nombre moyen de poissons observés dans une série d'instantanés pris durant un intervalle de visionnement), à la lumière de simulations, d'une expérience en laboratoire et d'une étude empirique. MaxN était relié de manière non linéaire à l'abondance réelle pour les trois approches, l'atténuation des estimations de l'abondance augmentant parallèlement à l'abondance réelle (c.-à-d. hypersensibilité). Ainsi, MaxN peut donner des indices d'abondance biaisés positivement pour les stocks de poissons en déclin, ou négativement pour les stocks de poissons en expansion. Pour sa part, MeanCount présente généralement une relation linéaire avec l'abondance réelle, et sa variabilité est semblable à celle de MaxN, ce qui porte à croire que MeanCount pourrait être utile pour déterminer l'indice d'abondance de poissons dans les évaluations reposant sur la vidéo sous-marine. [Traduit par la Rédaction]

Introduction

Visual census techniques have become an increasingly important tool in monitoring reef fish populations worldwide. Being inherently nonextractive, visual census techniques are often used to sample no-take areas such as marine protected areas and sanctuaries (Bohnsack and Bannerot 1986; Willis et al. 2000; Denny and Babcock 2004; Cappo et al. 2007; Schobernd and Sedberry 2009). In addition, visual gears minimize many of the size (Cappo et al. 2004; Morrison and Carbines 2006; Wells et al. 2008) and species (Ellis and DeMartini 1995; Bacheler et al. 2013) selectivity issues that arise with typical fishing gears such as trawls, traps, and hooks. Video has the secondary benefit of being able to provide additional information on a species' habitat use and behavioral patterns (e.g., He 2003; Silveira et al. 2003).

Visual census techniques encompass a broad range of underwater sampling gears, each having benefits and drawbacks (Wells et al. 2008). Diver surveys, either using a stationary or transect approach, typically identify more individual fishes than video-based methods (Pelletier et al. 2011) and are generally less expensive to employ than remote video (Murphy and Jenkins 2010). Divers are limited by depth, however, and these surveys require better weather conditions than surveys using remote gear (Willis et al. 2000). In addition, divers can cause behavioral changes in fish being observed, potentially leading to bias in abundance estimates (Watson and Harvey 2007; Dearden et al. 2010). Powered underwater sampling gears such as autonomously operated vehicles and submersibles are often used to sample in deep water, but the cost of these vehicles can be prohibitive and the movement of the gear has been shown to alter the behavior of some fish taxa (Barans 1986; Parker and Ross 1986; Harvey et al. 2007). Stationary cameras represent a middle ground, combining the utility of the underwater vehicle sampling gears with a much lower cost (Cappo et al. 2004; Murphy and Jenkins 2010).

A variety of approaches have been developed to enumerate fish on videos, with the primary goal of producing an index of abundance that is proportional to true abundance (Kimura and Somerton 2006). Typically, a single value for relative abundance is obtained for each time and location sampled to ensure statistical independence. Simple approaches, such as determining the time

Received 12 February 2013. Accepted 21 November 2013.

Paper handled by Associate Editor Verena Trenkel.

Z.H. Schobernd and N.M. Bacheler. Southeast Fisheries Science Center, National Marine Fisheries Service, 101 Pivers Island Road, Beaufort, NC 28516, USA.
P.B. Conn. National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, 7600 Sand Point Way N.E., Seattle, WA 98115, USA.

Corresponding author: Z.H. Schobernd (e-mail: zeb.schobernd@noaa.gov).

Can. J. Fish. Aquat. Sci. **71**: 464–471 (2014) dx.doi.org/10.1139/cjfas-2013-0086

at which a species first appears (i.e., time at first occurrence), have been used successfully to index demersal fish abundance (Priede and Merrett 1996), although in some instances this method has not correlated well with more direct sampling methods (Stoner et al. 2008) and can be highly variable (Farnsworth et al. 2007; Merritt et al. 2011). Similarly, methods that simply count all the individuals that appear in a given amount of time (i.e., MaxCount) have been applied in some areas, but this approach is prone to counting the same individuals multiple times and has been shown to be highly variable, making it undesirable for use in a fishery-independent survey (Conn 2011). A more comprehensive review of the use of baited underwater video gear can be found in Cappo et al. (2006).

The most commonly used metric in reef habitats when employing baited stationary video gear is the MaxN (or MinCount) approach, which is defined as the maximum number of fish observed in a single frame during the viewing interval (Ellis and DeMartini 1995; Willis and Babcock 2000; Watson et al. 2005; Merritt et al. 2011). The MaxN method is popular because it is a conservative estimate of the number of fish on a given reef; while there may be more individuals present in the area than are recorded, MaxN provides a minimum number of individuals known with certainty to occur in the sampling area, and no fish is counted more than once. It also ensures that individual fish may only contribute to the indexing metric once during a viewing interval (unlike MaxCount, for instance). A potential downside of the MaxN approach is that it may increasingly underestimate abundance at higher levels of true abundance (Conn 2011), resulting in a nonlinear relationship to true abundance.

Here, we compare the MaxN approach to a newly developed approach, MeanCount, to determine which technique is more likely to be linearly related to true abundance. MeanCount is calculated as the mean number of fish observed in a series of snapshots over a viewing interval, and has been shown to track true abundance linearly in preliminary, theoretical simulations of Conn (2011). We compare MaxN and MeanCount using three approaches: (*i*) simulations that build upon Conn's (2011) approach, (*ii*) a laboratory experiment, and (*iii*) comparisons based on empirically derived field data on reef fish. Given the ubiquity of video methods to index the abundance of fish and wildlife populations around the world, determining the most appropriate metric to enumerate organisms on video is critically important.

Materials and methods

Simulation study

We first conducted several computer simulations to investigate possible biases in using MaxN or MeanCount to index the abundance of fish populations. Our simulation modeling approach was similar to Conn (2011). Simulations used a stochastic implementation of the following set of differential equations:

$$\frac{dP}{dt} = \lambda_2 O - \lambda_1 P \tag{1}$$

$$\frac{dO}{dt} = \lambda_1 P - \lambda_2 O \tag{2}$$

where P is the number of fish that are in the camera's field of view, O is the number of fish out of the camera's field of view, λ_1 is the per capita rate of leaving the camera's field of view, and λ_2 is the per capita rate of entering the camera's field of view. Thus, our simulation model is a special case of Renshaw's (1991) two colony model, without the birth or death component. Our model thus assumes that individual fish move independently, that the rates of movement remain constant during the viewing interval, and that the number of fish present at the particular reef being sampled remains constant while sampling is being conducted. We also assume that observers (video readers) have perfect knowledge of the number of fish present on video at all times (i.e., there is no measurement error). Although these assumptions are unlikely to hold in practice (e.g., fish behavior is likely to change throughout the viewing interval both with regard to other fish and with regard to the sampling platform), this simplified simulation setup provides a test of various indexing methods under idealized sampling conditions.

We obtained stochastic realizations from eqs. 1 and 2 by making use of the fact that inter-event times are exponentially distributed (Renshaw 1991). Starting with initial conditions $O = O_0$ and $P = P_0$, and setting $T = 0$, each simulation proceeded as follows:

(*i*) Simulate the time for a single fish movement to occur as $t \sim \text{Exponential}(\lambda_2 O + \lambda_1 P)$

(*ii*) Set $T = T + t$. If $T > T_{\max}$, stop simulation. Otherwise proceed to (*iii*).

(*iii*) Simulate $u \sim$ uniform (0,1). If $u < \lambda_1 P / (\lambda_2 O + \lambda_1 P)$, set $O = O + 1$ and $P = P - 1$. Otherwise, set $O = O - 1$ and $P = P + 1$. Proceed to (*i*).

We ran each simulation for $T = 20$ min, recording all incremental changes in O throughout each time series.

We considered two different simulation scenarios to assess the performance of MaxN and MeanCount for indexing reef fish abundance. In the first, we set $\lambda_1 = 0.20$ and $\lambda_2 = 0.05$, which indicates relatively low rates of movement (e.g., *Epinephelus* groupers). Also under this scenario, fish are more likely to move off screen than they are to move on screen, and the probability (p) that fish present on the reef will be in the camera's field of view at any instant is $p = 0.20$ (this value can be obtained by setting $dP/dt = 0$ and solving for P in terms of $N = P + O$). The value of p represents a variety of interacting processes, including the camera's relative field of view and the relative attraction-repulsion of fish to video gear. In the second scenario, we set $\lambda_1 = \lambda_2 = 0.5$, corresponding to a higher rate of movement (e.g., *Lutjanus* snappers). In this scenario, the probability that a fish is on camera at any given time is $p = 0.5$. The value for P_0 at the beginning of each simulation was obtained by sampling P_0 from a Binomial(N,p) distribution. An example of the simulated stochastic processes for fish counts is shown in Fig. 1.

For each scenario, we conducted 10 000 simulations at each of a number of levels of abundance (N = 1, 5, 10, 15,…, 100). For each such realization, we calculated MaxN and MeanCount; in this case, MeanCount was determined by integrating over the continuous time series as opposed to a finite number of snapshots. For instance, if we denote the count of fish $O(t)$ that are captured by video at any instantaneous time t, we calculated MeanCount as

$$\text{MeanCount} = 0.05 \int_{t=0}^{20} O(t)dt \tag{3}$$

In doing so, we are integrating the area under the stochastic process (see e.g., Fig. 1), and dividing by the total time surveyed. This is easily accomplished as follows:

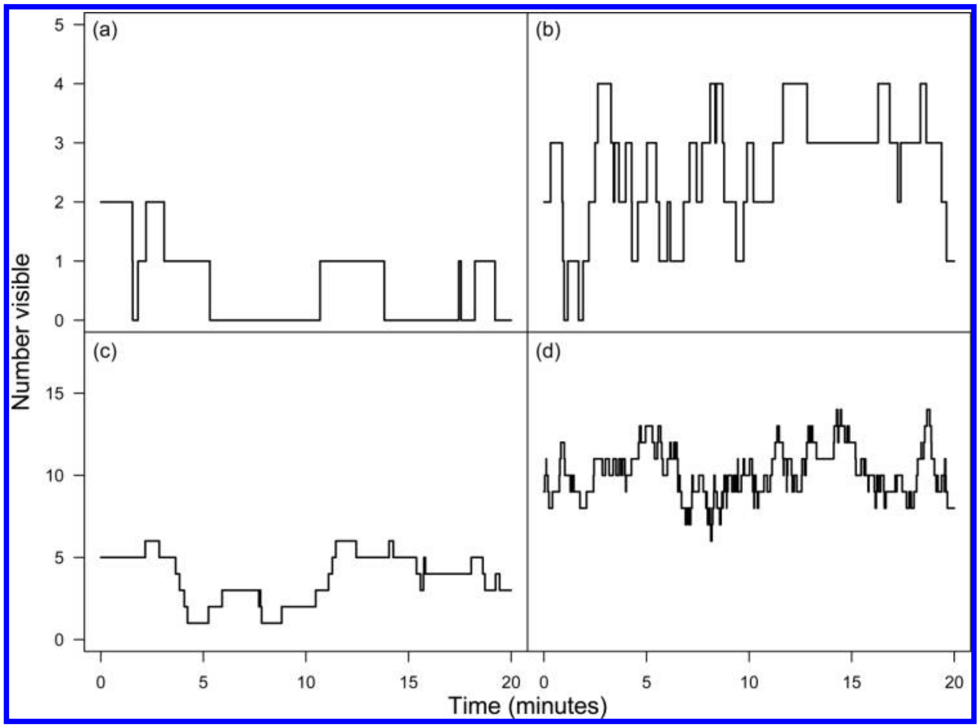
Define the start of the viewing interval as time $t_0 = 0$, and enumerate the times that fish abundance changes during the viewing interval as $t_1, t_2, \ldots, t_K$. The end of the viewing interval is set to $t_{K+1} = T$,

Describe abundance in the interval (t_k, t_{k+1}) as O_k,

Compute MeanCount as $\sum_{k=0}^{K} (t_{k+1} - t_k) O_k / T$.

For each value of N, we calculated the mean value of each index metric over all 10 000 simulations (i.e., $\bar{I}(N)$), as well as the coefficient of variation (CV = standard deviation divided by the mean)

Fig. 1. The number of fish visible by a theoretical camera based on simulations of two levels of true abundance (*a* and *b*: $N = 5$; *c* and *d*: $N = 20$) and two levels of movement rates (*a* and *c*: $\lambda_1 = 0.20$ and $\lambda_2 = 0.05$; *b* and *d*: $\lambda_1 = \lambda_2 = 0.5$).



for estimates. Each metric was then standardized by dividing by its value at $N = 1$ (that is, $\bar{I}_{\text{rel}}(N) = \bar{I}(N)/\bar{I}(1)$) to investigate possible departures from linearity. Under a strictly linear relationship, $\bar{I}_{\text{rel}}(N) = N$. All analyses were conducted using R, version 2.11.1 (R Development Core Team 2007).

Laboratory study

We next conducted a laboratory experiment using varying, known numbers of fish in a tank to directly determine the relationship between MaxN or MeanCount and true abundance. For this experiment, four replicates of eight treatments of pinfish (*Lagodon rhomboides*) abundance (i.e., 1, 5, 10, 15, 20, 30, 45, and 60 individuals) were placed in a circular 1.8 m diameter tank filled with 3.1 m^3 of water and allowed to acclimate for approximately 10 min. Pinfish were used for their hardiness, their relative ease of collection, and their behavioral characteristics that are similar to many reef fish species; individual fish used ranged in size from 15 to 25 cm. A Canon Vixia HF S200 video camera in an underwater housing was placed in the center of the tank and set to record for 20 min. The videos were then read using both MaxN and MeanCount approaches. For the MeanCount approach, we read snapshots every 30 s over the 20 min interval, for a total of 41 snapshots. We then fit regression models that expressed the response variable (MaxN or MeanCount) as (1) a linear function of abundance, and (2) a quadratic function of abundance (this model included both linear and quadratic terms). We used a likelihood ratio test to test whether the quadratic term was significantly different from zero, where a significant result provided evidence for a nonlinear relationship between MaxN or MeanCount and true pinfish abundance.

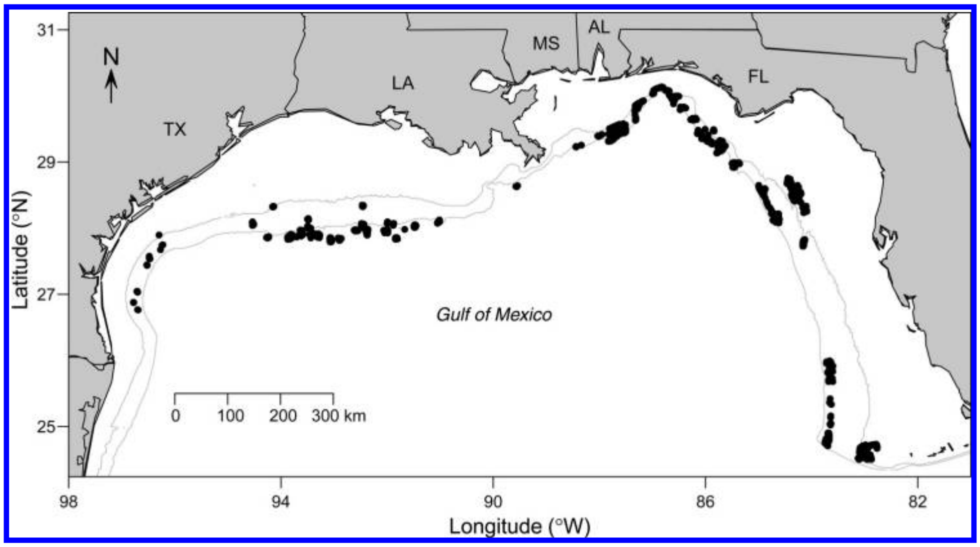
Empirical study

Finally, we examined the relationship between MaxN and MeanCount using empirically derived, field-collected video data from a long-term reef fish survey in the Gulf of Mexico. These data were collected by the National Marine Fisheries Service in 2001–2007 using a 4-camera array deployed on hard bottom habitat throughout the Gulf of Mexico. This survey, conducted using a baited stationary camera array with natural light in depths ranging from approximately 30–100 m, was developed to index reef fish populations in the region, and was typically conducted in the spring and summer on shelf-edge reefs from south Texas to the Dry Tortugas (Fig. 2). These particular years were selected because they were the most recent years before a significant methodology change in the video reading procedures. A detailed description of sampling and video processing protocols is available in Gledhill et al. (2006).

One video (of four) was randomly selected from each 60 min deployment, and 20 min of that video was analyzed, with reading starting once the gear touched the bottom and visibility was sufficient to identify taxa. The time each individual fish swam into and out of view (i.e., time in – time out) was recorded, allowing the abundance of fish in view to be calculated every second during the read time for each species. These abundances were used to calculate MaxN and MeanCount for each survey station, with MeanCount snapshots taken every second and integrated over the 20 min interval. For the purposes of this study we focused on 10 common and commercially important reef fish species in the region: almaco jack (*Seriola rivoliana*), gag (*Mycteroperca microlepis*), gray snapper (*Lutjanus griseus*), gray triggerfish (*Balistes capriscus*), greater amberjack (*Seriola dumerili*), red grouper (*Epinephelus morio*), red snapper (*Lutjanus campechanus*), red porgy (*Pagrus pagrus*), scamp (*Mycteroperca phenax*), and vermilion snapper (*Rhomboplites aurorubens*).

We used an errors-in-variable (EIV) modeling approach to investigate the relationship between MeanCount and MaxN from this survey. Ideally, one would compare these metrics to true abundance levels, but true abundance was unknown at each sampled station. Since simulations indicated that the relationship between true abundance and MeanCount was linear, we determined the functional relationship between MaxN and MeanCount directly to determine if it was linear or nonlinear. However, this approach violates the traditional regression analysis assumption that the independent variable is known without error (Zar 1999). Therefore, we used a hierarchical EIV model to account for errors in the independent variable (i.e., MeanCount). Our approach is similar

Fig. 2. Sampling locations for the National Marine Fisheries Service's reef fish video survey in the Gulf of Mexico, 2001–2007. Bathymetric contour lines indicate 50 and 100 m depth, and points indicate the location of each video sample included in the analysis. Note that symbols overlap in many cases.



to that of Jiao et al. (2006), who used an EIV model to analyze the relationship between relative abundance indices and biomass when both were subject to measurement error.

Adopting a hierarchical modeling perspective, we assumed that the expected MeanCount at each site i could be modeled through a Gamma(α,β) distribution, where α and β are shape and scale parameters to be estimated. The relationship between the expected MaxN (μ_{yi}) and expected MeanCount (μ_{xi}) at site i was then specified as

$$\mu_{yi} = a\mu_{xi} + b\mu_{xi}^2 \tag{4}$$

where a and b are parameters to be estimated. These parameters have interpretations similar to linear and quadratic regression parameters, respectively. For instance, if b differs substantially from zero, there is evidence that the relationship between the two metrics is nonlinear. A significantly negative b value would indicate an asymptotic MaxN with increasing MeanCount values, and a significantly positive b value would indicate an asymptotic MeanCount with increasing values of MaxN.

The relationship between these abstract parameters and data are provided by specifying a sampling model. We assumed that each MeanCount replicate j from site i, x_{ij}, followed a Normal(μ_{xi}, σ_{xi}^2) distribution; similarly, we also assumed that each MaxN replicate j from site i, y_{ij}, followed a Normal(μ_{yi}, σ_{yi}^2) distribution. In each case, we used the observed coefficient of variation (CV) from actual data to provide a value for σ_{xi} and σ_{yi}, with $\sigma_{yi}^2 = \mu_{yi}^2\text{CV}_y^2$ and $\sigma_{xi}^2 = \mu_{xi}^2\text{CV}_x^2$.

Model fitting was conducted using WinBUGS 14 software (Lunn et al. 2000), assuming the following relatively vague priors for α, β, a, and b:

$$\alpha \sim \text{Gamma}(0.1, 0.1),$$

$$\beta \sim \text{Gamma}(0.1, 0.1),$$

$$a \sim \text{Uniform}(0, 40),$$

$$b \sim \text{Normal}(0.0, 0.1).$$

The final prior distribution for b was chosen to constrain it to a reasonable range and prevent numerical overruns. We examined standard trace plots and Gelman–Rubin diagnostics in WinBUGS to confirm model convergence (Brooks and Gelman 1998).

Results

Simulation study

MeanCount was linearly related to true abundance in each of the two simulation scenarios (i.e., high and low movement rates; Fig. 3). In contrast, MaxN scaled nonlinearly with true abundance in each scenario, progressively underrepresenting true abundance at higher levels of true abundance (Fig. 3). For instance, with a doubling of true abundance from 40 to 80 in the high movement rate simulation, MeanCount doubled (i.e., 100% increase) but MaxN only increased 85%. As true abundance increased from 1 to 100 fish, MeanCount increased 100-fold but MaxN only increased about 40-fold when movement rates were low and 60-fold when movement rates were high. Average CVs for MeanCount and MaxN in the high movement rate simulation were 0.06 and 0.05, respectively, suggesting similar overall CVs between the two approaches (Fig. 4). However, MaxN had a 50% smaller CV at low abundance levels (i.e., ≤20 fish) and MeanCount had a 27% lower CV at higher abundance levels (i.e., ≥80 fish).

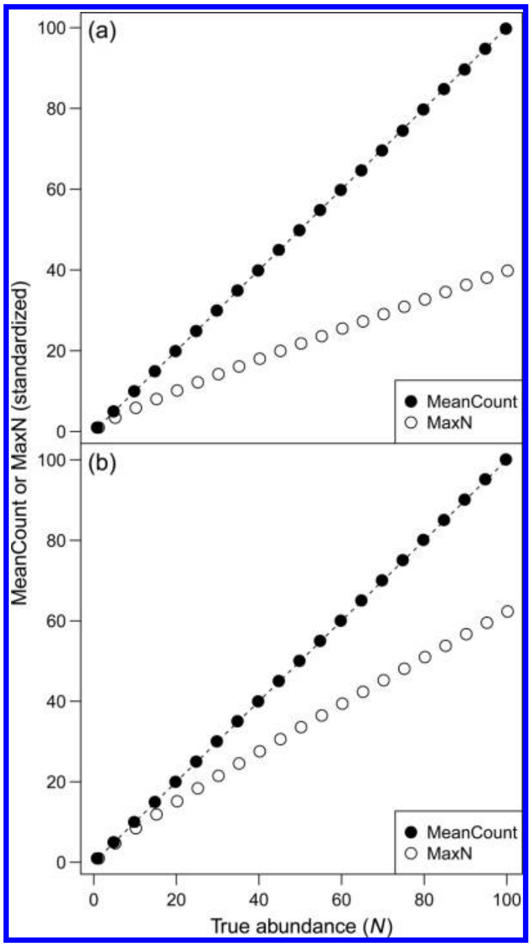
Laboratory experiment

MeanCount values for pinfish ranged from 0 in a 1-fish treatment to 12.2 in a 45-fish treatment, whereas MaxN ranged from 0 in a 1-fish density treatment to 17.0 in a 60-fish treatment (Fig. 5). The relationship between MeanCount and true pinfish abundance was slightly nonlinear ($\chi^2 = 4.4$, df = 1, $P = 0.04$), while the relationship between MaxN and true abundance was strongly nonlinear ($\chi^2 = 36.4$, df = 1, $P < 0.001$). MeanCount and MaxN were similar in terms of their standard errors, but MeanCount had higher coefficients of variation, likely because the denominator was smaller for MeanCount than MaxN (Table 1).

Empirical study

An initial examination of the data revealed a total of 2974 videos collected in the Gulf of Mexico available for inclusion in the analysis. Because not every species was observed on every video, there was a range of numbers of videos on which focal species were observed. Among the 10 species included in the empirical analysis, gray snapper were observed on the fewest videos ($n = 147$) and scamp were observed on the most videos ($n = 773$). For all

Fig. 3. Relationship between MeanCount (filled circles) or MaxN (open circles) and true abundance as estimated by the simulation model for (*a*) low and (*b*) high movement rates. Dashed lines represent linear relationships. Each series was standardized by dividing by its mean value at $N = 1$, so the y axis represents $\bar{I}_{\mathrm{rel}}(N)$ (see text for more details). As true abundance increases from 1 to 100, the standardized MeanCount index tracks true abundance, while MaxN underestimates the increase by 62% (*a*) and 30% (*b*).

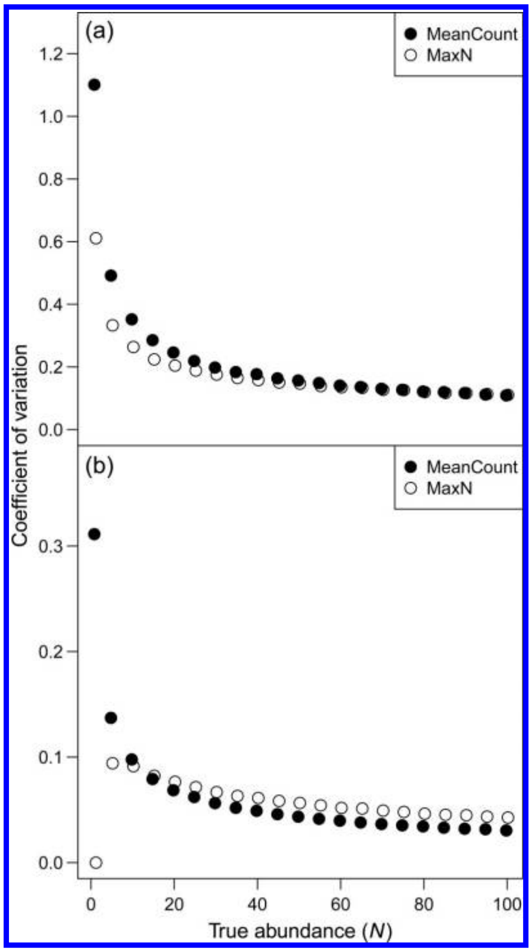


10 species examined, MaxN was nonlinearly related to MeanCount, with proportionally lower values of MaxN as MeanCount increased, even after accounting for potential errors in the independent variable (i.e., MeanCount) using an EIV model (Fig. 6). The parameter estimates for *b* varied between −0.61 (scamp) and −20.60 (almaco jack; Table 2), and in all cases *b* values were significantly different from 0, suggesting that MaxN was hyperstable with regard to true abundance (i.e., the index underrepresents the magnitude of changes in true abundance). While the patterns within each taxon as seen in Fig. 6 varied slightly, these results were consistent across all species despite a wide range of abundance (solitary fish to large schools) and behavioral (lie-and-wait predators to active schooling fish) characteristics.

Discussion

Through a variety of approaches (simulation, laboratory experiments, and empirical data), we have demonstrated a propensity for MaxN to be nonlinearly related to true abundance. Counts made using the MaxN approach underrepresented changes in true abundance, resulting in a nonlinear, hyperstable relationship between MaxN and true abundance (Hilborn and Walters 1992). The mechanism for the bias in MaxN is likely rooted in its reliance on a maximum order statistic to characterize population abundance at a given site. For instance, all simulations with $N = 1$ fish resulted in a MaxN value of 1; simulations with $N = 5$ fish would often result in a MaxN value of 4 or 5, but simulations with much higher N (e.g., $N = 80$) would be much more likely to take on a lower proportional value (e.g., $I(80) = 50$). In contrast, MeanCount appeared to be linearly related to true abundance.

Fig. 4. Coefficients of variation for MeanCount (filled circles) or MaxN (open circles) as estimated by simulation experiments when fish movement rates are (*a*) low or (*b*) high.



Another consideration when comparing MaxN to MeanCount is the variability surrounding relative abundance estimates. While there has been very little focus on directly comparing the variances associated with different video metrics, minimizing the variability of fishery-independent indices has long been a goal of stock assessment scientists (Maunder and Punt 2004). Our simulation results suggest that the variability around MaxN and MeanCount is generally similar, especially at higher levels of true abundance. In the laboratory experiment, the variability around MeanCount was somewhat higher than the variability around MaxN, particularly at low levels of true abundance. However, this increased variation for MeanCount in the laboratory experiment compared with the simulation is likely due to a slight difference in the viewing protocols between the laboratory experiment and the approach modeled in the simulation. The simulations assumed continuous monitoring of the fish visible on the screen, while the laboratory experiment utilized a count taken every thirty seconds. This alteration is because of the burdensome effort required to carry out a "time in – time out" protocol, where the times of entry and exit are recorded for each fish. While slightly higher variability is not ideal, it may be more useful in a stock

Fig. 5. The relationships between (*a*) MeanCount and true abundance or (*b*) MaxN and true abundance of pinfish (*Lagodon rhomboides*) in a laboratory experiment. The dashed line represents a linear relationship, and the solid line is the actual modeled relationship between the two metric observations and true abundance.

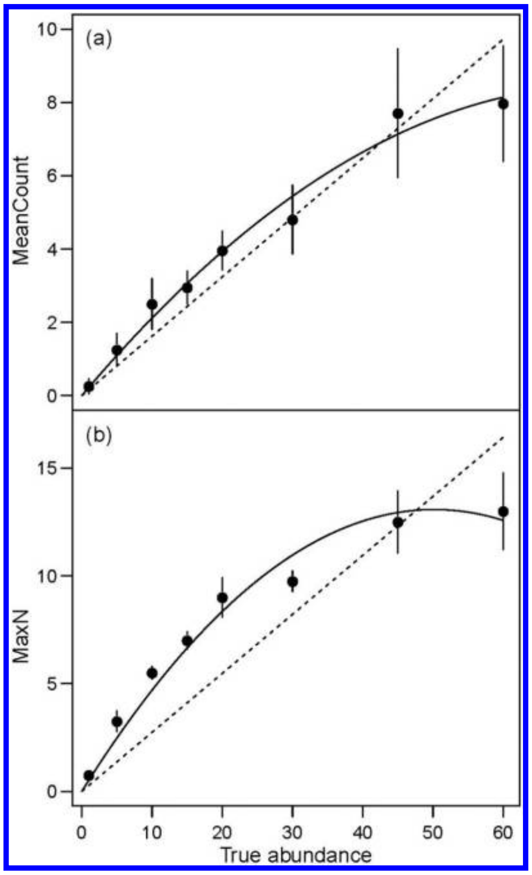


Table 1. Measurements of variance from the laboratory experiment comparing both MaxN and MeanCount to true abundance at eight densities of pinfish (*Lagodon rhomboides*).

True abundance	MeanCount			MaxN		
	Range	SE	CV	Range	SE	CV
1	0.00–0.88	0.21	163	0–1	0.25	67
5	0.34–2.44	0.45	71	2–4	0.48	30
10	0.71–3.63	0.69	56	5–6	0.29	11
15	1.83–3.73	0.45	30	6–8	0.41	12
20	2.83–5.10	0.53	27	7–11	0.91	20
30	3.05–7.24	0.95	39	9–11	0.48	10
45	4.27–12.20	1.76	46	9–16	1.44	23
60	5.32–12.07	1.58	40	10–17	1.78	27

Note: Range, minimum and maximum MeanCount and MaxN values observed for each treatment; SE, standard error; CV, coefficient of variation (%).

Table 2. Parameter estimates from the error-in-variables model relating MaxN to MeanCount for 10 reef fish observed on videos in the Gulf of Mexico, 2001–2007.

Species	*N*	$\hat{a}$ (95% CI)	$\hat{b}$ (95% CI)	$\hat{b}$ SD
Scamp	773	5.3 (5.1–5.5)	−0.61 (−0.62 – −0.52)	0.04
Red porgy	766	9.6 (9.1–10.0)	−2.11 (−2.41 – −1.81)	0.22
Almaco jack	677	25.7 (24.3–27.1)	−20.60 (−24.90 – −15.69)	2.33
Greater amberjack	596	29.8 (28.0–31.6)	−16.01 (−20.02 – −11.36)	2.18
Red snapper	578	9.4 (8.9–9.9)	−1.87 (−2.31 – −1.45)	0.22
Red grouper	557	8.0 (7.4–8.6	−9.53 (−10.97 – −8.07)	0.75
Gray triggerfish	350	11.2 (10.4–12.1)	−6.49 (−8.04 – −4.92)	0.79
Gag	330	9.8 (9.0–10.6)	−8.26 (−10.01 – −6.55)	0.88
Vermilion snapper	287	17.1 (15.7–18.7)	−3.49 (−4.57 – −2.46)	0.54
Gray snapper	147	12.5 (11.0–14.1)	−3.95 (−5.46 – −2.53)	0.75

Note: *N* is the number of video samples included for each species, $\hat{a}$ and $\hat{b}$ are the posterior means for linear and quadratic parameters used to model the relationship between MaxN and MeanCount, and $\hat{b}$ SD is the standard deviation of $\hat{b}$. The 95% confidence intervals are shown in parentheses after the parameter values for $\hat{a}$ and $\hat{b}$.

assessment context than a hyperstable index with a bias whose magnitude varies with abundance.

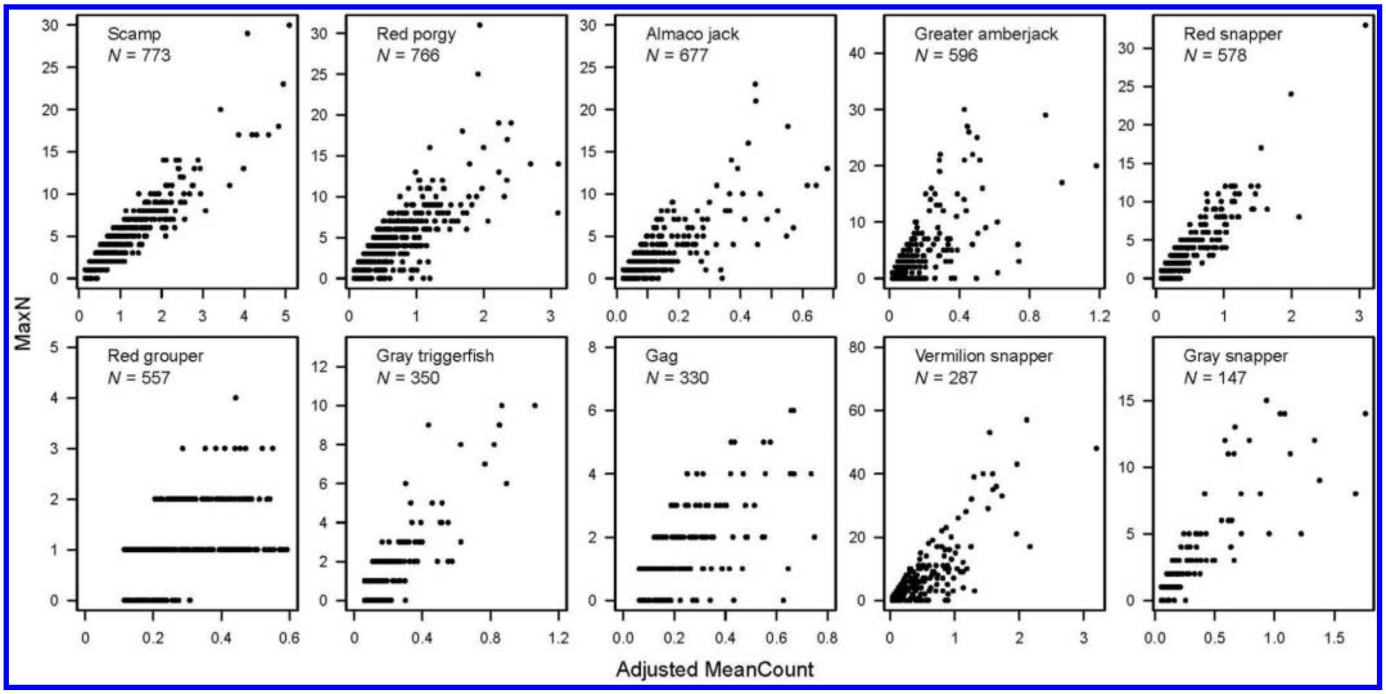
The implications of hyperstability in MaxN are potentially substantial. As underwater video becomes more prevalent in fishery-independent surveys, analysts will increasingly use video abundance metrics within fisheries stock assessments to help gauge relative abundance. Assuming that the properties of the relative abundance index are preserved in the final estimates of the stock assessment model, reliance on a hyperstable metric for increasing populations could contribute to overly conservative fishery regulations, unnecessarily limiting allowable catch limits. For populations that are decreasing in abundance, reliance on a hyperstable metric could contribute to overly permissive fishery regulations.

While our three-pronged approach clearly shows that there are potential benefits to using MeanCount over MaxN, it is not entirely clear how those advantages would translate to field data or subsequent index construction. All of our analyses indicate that MaxN provides a biased estimate of true abundance, but the amount of bias will ultimately depend strongly on the abundance, schooling behavior, and movement patterns of the organism under study. For example, many of the species seen on the videos that are of interest to fisheries managers, including large solitary predators, are seen in low abundances where the impact of hyperstability is the least. Further, relative abundance indices are often constructed with zero-inflated statistical models (e.g., delta-GLMs; Pennington 1983; Stefánsson 1996), so that the ultimate index constructed depends on both zeroes (i.e., sampling events for which no fish are detected) as well as the positive component (i.e., sampling events that detect at least one fish). We expect that relative abundance indices constructed in such a manner will tend to have bias in the second component if MaxN is employed, but it is unclear what the ultimate effect would be. Such an index would likely still be hyperstable, although less so than is indicated by examining only the positive ($I > 0$) component as we have done here. In addition, MeanCount should serve as an adequate approach when taking fish length measurements using the increasingly prevalent stereo–video systems (Harvey et al. 2002). Similar to MaxN, measurements could be taken at the highest MeanCount frame and provide comparable numbers of fish lengths for use in stock assessments.

MaxN has been a preferred metric in multiple studies in United States Pacific waters (Ellis and DeMartini 1995; Stoner et al. 2008; Merritt et al. 2011) and is common in monitoring programs for marine reserves in Australia and New Zealand (Willis and Babcock 2000; Cappo et al. 2004; Denny et al. 2004). There is potential that the efficacy of these reserves could be underestimated because of hyperstability — as fish stocks inside the reserves increase, the level of recovery that is reported could be negatively biased. Because of these concerns, a fishery-independent program monitoring reef fish populations in the southeastern United States using video chose to use MeanCount as their metric for video analysis (i.e., SouthEast Fishery Independent Survey; Bacheler et al. 2013).

Video surveys are certainly capable of providing information above and beyond relative abundance. For instance, they can be used to assess species richness or length distributions (e.g., using stereo-image photogrammetry) (Shortis et al. 2009). Because MaxN utilizes a continuous stretch of video read time, any fish that enters the field of view, no matter how briefly, gets counted

Fig. 6. Results of error-in-variable model comparing adjusted MeanCount (μ_{xi}) values to MaxN (μ_{yi}) values for 10 species of reef fish, using data from National Marine Fisheries Service Gulf of Mexico video reef fish survey, 2001–2007. Each point represents pairwise adjusted MeanCount and MaxN index values for each survey location that a species was detected in a video.



and included in species richness summaries. MeanCount can be constructed either using a continuous viewing protocol (as in the simulation study) or with counts that are taken at regular intervals (e.g., every thirty seconds) to reduce viewing times. When counts are taken at intervals, some rarer species will undoubtedly be missed. There is a broad consensus in the literature that ties an increase in viewing time to an increase in species richness (Bortone et al. 1989; St John et al. 1990; Gledhill 2001). However, to our mind, this issue is somewhat of a "red herring"; species that are present at a site may still be missed even with a continuous viewing protocol (e.g., if they never appear on frame). Approaches that account for imperfect species detection (e.g., Boulinier et al. 1998; Lekve et al. 2002) appear necessary for robust examination of species richness regardless of the viewing protocol chosen.

Sample sizes for fish length data will also be reduced using an interval-based viewing procedure. To eliminate the possibility of double counting fish it would be logical to base the length distribution for a given video sample on the video frame with the highest abundance to ensure an independent sample. However, this choice relies on the particular viewing procedure adopted, and is not necessarily reliant on the specific abundance metric (MaxN, MeanCount) selected.

In summary, underwater video surveys are an increasingly important tool for monitoring and regulating fish populations (Cappo et al. 2004). Utilizing the most statistically appropriate method to enumerate fish from these videos has, therefore, become of paramount importance. While there are a variety of different ways to attempt to survey fish underwater using video techniques (e.g., baited, unidirectional, stereo cameras (Watson et al. 2005); unbaited, rotating, camera (Pelletier et al. 2012)), we focused our efforts on a single fixed, unidirectional camera approach popular in the southeastern United States, and our study suggests that a new metric, MeanCount, can index abundance accurately and precisely and may have advantages over the commonly used MaxN approach. Further analysis, including recalculating past indices of abundance for key species using MeanCount, is needed before conclusively determining the impact, if any, of using MaxN (versus MeanCount) as a metric in the development of indices of abundance. However, given the weight of evidence from our study, we suggest MeanCount as a viable alternative when indexing fish populations using underwater video.

Acknowledgements

We thank C. Gledhill and the NMFS-Pascagoula laboratory for providing access to the Gulf of Mexico video data; A. Pollack for preliminary data summarizations; D. Berrane for helping to collect pinfish for the laboratory experiment; and J. Morris, B. Harrison, and A. Ostrowski for use of fish tanks. We also thank M. Campbell, L. Coggins, D. DeVries, C. Gledhill, T. Kellison, W. Mitchell, C. Schobernd, K. Shertzer, and E. Williams for helpful discussions; and L. Coggins, P. Marraro, K Shertzer, T. Kellison, A. Hohn, A. Chester, and D. Pelletier for reviews. Mention of trade names or commercial companies is for identification purposes only and does not imply endorsement by the National Marine Fisheries Service, NOAA.

References

Bacheler, N.M., Schobernd, C.M., Schobernd, Z.H., Mitchell, W.A., Berrane, D.J., Kellison, G.T., and Reichert, M.J.M. 2013. Comparison of trap and underwater video gears for indexing reef fish presence and abundance in the southeast United States. Fish. Res. **143**:81–88. doi:10.1016/j.fishres.2013.01.013.

Barans, C.A. 1986. Submersible avoidance by yellowfin bass, *Anthias nicholsi*. Northeast Gulf Sci. **8**(1): 91–95.

Bohnsack, J.A., and Bannerot, S.P. 1986. A stationary visual census technique for quantitatively assessing community structure of coral reef fishes. NOAA Tech. Rep. **41**, 15 pp.

Bortone, S.A., Kimmel, J.J., and Bundrick, C.M. 1989. A comparison of three methods for visually assessing reef fish communities: time and area compensated. Northeast Gulf Sci. **10**(2): 85–96.

Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E., and Pollock, K.H. 1998. Estimating species richness: the importance of heterogeneity in species detectability. Ecology, **79**: 1018–1028. doi:10.1890/0012-9658(1998)079[1018:ESRTIO]2.0.CO;2.

Brooks, S.P., and Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. **7**(4): 434–455. doi:10.2307/1390675.

Cappo, M., Speare, P., and De'ath, G. 2004. Comparison of baited remote underwater video stations (BRUVS) and prawn (shrimp) trawls for assessments of fish biodiversity in inter-reefal areas of the Great Barrier Reef Marine Park. J. Exp. Mar. Biol. Ecol. **302**(2): 123–152. doi:10.1016/j.jembe.2003.10.006.

Cappo, M., Harvey, E., and Shortis, M. 2006. Counting and measuring fish with baited video techniques—an overview. *In* Cutting-edge technologies in fish and fisheries science. *Edited by* J.M. Lyle, D.M. Furlani, and C.D. Buxton. Australian Society for Fish Biology Workshop Proceedings, Hobart, Tasmania, August 2006. pp. 101–114.

Cappo, M., De'ath, G., and Speare, P. 2007. Inter-reef vertebrate communities of the Great Barrier Reef Marine Park determined by baited remote underwater video stations. Mar. Ecol. Prog. Ser. **350**: 209–221. doi:10.3354/meps07189.

Conn, P.B. 2011. An evaluation and power analysis of fishery independent reef fish sampling in the Gulf of Mexico and US south Atlantic. NOAA Tech. Memo. NMFS-SEFSC-610. Available from http://docs.lib.noaa.gov/noaa_documents/NMFS/SEFSC/TM_NMFS_SEFSC/NMFS_SEFSC_TM_610.pdf [accessed 15 January 2013].

Dearden, P., Theberge, M., and Yasué, M. 2010. Using underwater cameras to assess the effects of snorkeler and SCUBA diver presence on coral reef fish abundance, family richness, and species composition. Environ. Monit. Assess. **163**:531–538. doi:10.1007/s10661-009-0855-3. PMID:19353295.

Denny, C.M., and Babcock, R.C. 2004. Do partial marine reserves protect reef fish assemblages? Biol. Conserv. **116**(1): 119–129. doi:10.1016/S0006-3207(03)00183-6.

Denny, C.M., Willis, T.J., and Babcock, R.C. 2004. Rapid recolonisation of snapper *Pagrus auratus*: Sparidae within an offshore island marine reserve after implementation of no-take status. Mar. Ecol. Prog. Ser. **272**: 183–190. doi:10.3354/meps272183.

Ellis, D.M., and DeMartini, E.E. 1995. Evaluation of a video camera technique for indexing abundances of juvenile pink snapper, *Pristipomoides filamentosus*, and other Hawaiian insular shelf fishes. Fish. Bull. **93**(1): 67–77. Available from http://fishbull.noaa.gov/931/ellis.pdf [accessed 15 January 2013].

Farnsworth, K.D., Thygesen, U.H., Ditlevsen, S., and King, N.J. 2007. How to estimate scavenger fish abundance using baited camera data. Mar. Ecol. Prog. Ser. **350**: 223–234. doi:10.3354/meps07190.

Gledhill, C.T. 2001. Reef fish assemblages on Gulf of Mexico shelf-edge reefs. Ph.D. dissertation, Department of Marine Sciences, University of South Alabama, Mobile, Alabama.

Gledhill, C.T., Ingram, G.W., Jr., Rademacher, K.R., Felts, P., and Trigg, B. 2006. NOAA Fisheries reef fish video surveys: yearly indices of abundance for gag (*Mycteroperca microlepis*). SEDAR10-DW-12. SEDAR, North Charleston, S.C.

Harvey, E.S., Fletcher, D., and Shortis, M.R. 2002. Estimation of reef fish length by divers and by stereo-video: a first comparison of the accuracy and precision in the field on living fish under operational conditions. Fish. Res. **57**(3): 257–267.

Harvey, E.S., Cappo, M., Butler, J.J., Hall, N., and Kendrick, G.A. 2007. Bait attraction affects the performance of remote underwater video stations in assessment of demersal fish community structure. Mar. Ecol. Prog. Ser. **350**: 245–254. doi:10.3354/meps07192.

He, P.G. 2003. Swimming behaviour of winter flounder (*Pleuronectes americanus*) on natural fishing grounds as observed by an underwater video camera. Fish. Res. **60**(2–3): 507–514. doi:10.1016/S0165-7836(02)00086-3.

Hilborn, R., and Walters, C.J. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York. 570 pp.

Jiao, Y., Reid, K., and Nudds, T. 2006. Variation in the catchability of yellow perch (*Perca flavescens*) in the fisheries of Lake Erie using a Bayesian error-in-variable approach. ICES J. Mar. Sci. **63**(9): 1695–1704. doi:10.1016/j.icesjms.2006.07.002.

Kimura, D.K., and Somerton, D.A. 2006. Review of statistical aspects of survey sampling for marine fisheries. Rev. Fish. Sci. **14**(3): 245–283. doi:10.1080/10641260600621761.

Lekve, K., Boulinier, T., Stenseth, N.C., Gjøsaeter, J., Fromentin, J.-M., Hines, J.E., and Nichols, J.D. 2002. Spatio-temporal dynamics of species richness in coastal fish communities. Proc. R. Soc. Ser. B Biol. Sci. **269**: 1781–1789. doi:10.1098/rspb.2002.2058.

Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. 2000. WinBUGS — a Bayesian modeling framework: concepts, structure, and extensibility. Stat. Comput. **10**(4): 325–337. doi:10.1023/A:1008929526011.

Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. **70**(2–3): 141–159. doi:10.1016/j.fishres.2004.08.002.

Merritt, D., Donovan, M.K., Kelley, C., Waterhouse, L., Parke, M., Wong, K., and Drazen, J.C. 2011. BotCam: a baited camera system for nonextractive monitoring of bottomfish species. Fish. Bull. **109**(1): 56–67. Available from http://fishbull.noaa.gov/1091/merritt.pdf [accessed 15 January 2013].

Morrison, M., and Carbines, G. 2006. Estimating the abundance and size structure of an estuarine population of the sparid *Pagrus auratus*, using a towed camera during nocturnal periods of inactivity, and comparisons with conventional sampling techniques. Fish. Res. **82**(1–3): 150–161. doi:10.1016/j.fishres.2006.06.024.

Murphy, H.M., and Jenkins, G.P. 2010. Observational methods used in marine spatial monitoring of fishes and associated habitats: a review. Mar. Freshwater Res. **61**(2): 236–252. doi:10.1071/MF09068.

Parker, R.O., Jr., and Ross, S.W. 1986. Observing reef fishes from submersibles off North Carolina. Northeast Gulf Sci. **8**: 31–49.

Pelletier, D., Leleu, K., Mou-Tham, G., Guillemot, N., and Chabanet, P. 2011. Comparison of visual census and high definition video transects for monitoring coral reef fish assemblages. Fish. Res. **107**(1–3): 84–93. doi:10.1016/j.fishres.2010.10.011.

Pelletier, D., Leleu, K., Mallet, D., Mou-Tham, G., Hervé, G., Boureau, M., and Guilpart, N. 2012. Remote high-definition rotating video enables fast spatial survey of marine underwater macrofauna and habitats. PLoS ONE, **7**(2): e30536. doi:10.1371/journal.pone.0030536.

Pennington, M. 1983. Efficient estimators of abundance for fish and plankton surveys. Biometrics, **38**: 281–286. doi:10.2307/2530830.

Priede, I.G., and Merrett, N.R. 1996. Estimation of abundance of abyssal demersal fishes: a comparison of data from trawls and baited cameras. J. Fish Biol. **49**(Suppl. A): 207–216. doi:10.1111/j.1095-8649.1996.tb06077.x.

R Development Core Team. 2007. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available from http://www.R-project.org [accessed 15 January 2013].

Renshaw, E. 1991. Modelling biological populations in space and time. Cambridge University Press, Cambridge, UK.

Schobernd, C.M., and Sedberry, G.R. 2009. Shelf-edge and upper-slope reef fish assemblages in the South Atlantic Bight: habitat characteristics, spatial variation, and reproductive behavior. Bull. Mar. Sci. **84**(1): 67–92.

Shortis, M., Harvey, E., and Abdo, D. 2009. A review of underwater stereo-image measurement for marine biology and ecological applications. *In* Oceanography and marine biology: an annual review. Vol. 47. *Edited by* R.N. Gibson, R.J.A. Atkinson, and J.D.M. Gordon. CRC Press-Taylor and Francis Group, Boca Raton, Fla. pp. 257–292.

Silveira, L., Jácomo, A.T.A., and Diniz-Filho, J.A.F. 2003. Camera trap, line transect census and track surveys: a comparative evaluation. Biol. Conserv. **114**(3): 351–355. doi:10.1016/S0006-3207(03)00063-6.

Stefánsson, G. 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES J. Mar. Sci. **53**(3): 577–588. doi:10.1006/jmsc.1996.0079.

St. John, J., Russ, G.R., and Gladstone, W. 1990. Accuracy and bias of visual estimates of numbers, size structure and biomass of a coral reef fish. Mar. Ecol. Prog. Ser. **64**: 253–262. doi:10.3354/meps064253.

Stoner, A.W., Laurel, B.J., and Hurst, T.P. 2008. Using a baited camera to assess relative abundance of juvenile Pacific cod: field and laboratory trials. J. Exp. Mar. Biol. Ecol. **354**(2): 202–211. doi:10.1016/j.jembe.2007.11.008.

Watson, D.L., and Harvey, E.S. 2007. Behaviour of temperate and sub-tropical reef fishes towards a stationary SCUBA diver. Mar. Freshwater Behav. Physiol. **40**(2): 85–103. doi:10.1080/10236240701393263.

Watson, D.L., Harvey, E.S., Anderson, M.J., and Kendrick, G.A. 2005. A comparison of temperate reef fish assemblages recorded by three underwater stereo-video techniques. Mar. Biol. **148**(2): 415–425. doi:10.1007/s00227-005-0090-6.

Wells, R.J.D., Boswell, K.M., Cowan, J.H., Jr., and Patterson, W.F., III. 2008. Size selectivity of sampling gears targeting red snapper in the northern Gulf of Mexico. Fish. Res. **89**(3): 294–299. doi:10.1016/j.fishres.2007.10.010.

Willis, T.J., and Babcock, R.C. 2000. A baited underwater video system for the determination of relative density of carnivorous reef fish. Mar. Freshwater Res. **51**(8): 755–763. doi:10.1071/MF00010.

Willis, T.J., Millar, R.B., and Babcock, R.C. 2000. Detection of spatial variability in relative density of fishes: comparison of visual census, angling, and baited underwater video. Mar. Ecol. Prog. Ser. **198**: 249–260. doi:10.3354/meps198249.

Zar, J.H. 1999. Biostatistical analysis. 4th ed. Prentice Hall, Upper Saddle River, N.J.