Integrating underwater video into traditional fisheries indices using a hierarchical formulation of a state-space model

Daniel C. Gwinn[a,b*], Nathan M. Bacheler[c] and Kyle Shertzer[c]

[a] Biometric Research, 3 Hulbert Street, South Fremantle 6162, Western Australia, Australia.
[b] The University of Western Australia, School of Biological Sciences, Perth, Western Australia, Australia.
[c] National Marine Fisheries Service, Southeast Fisheries Science Center, Beaufort, North Carolina, United States of America.

[*]Corresponding author: Daniel C. Gwinn, dgwinnbr@gmail.com

Abstract

Indices of abundance are commonly used in fisheries stock assessment models to represent trends in population size over time; however, an index can misrepresent such trends when catchability varies, sampling gears change or spatial sampling frames shift. Here we develop a state-space model in a Bayesian framework that combines both chevron trap catches and video counts into a single integrated index. The modeling approach accounts for variation in sampling efficiency (catchability) of both sampling gears and adjusts for aspects of changes in the spatial sampling frame (sampling intensity and spatial coverage) through time due to monitoring program development. We validate the model using a simulation study and then demonstrate its utility using data on vermilion snapper *Rhomboplites aurorubens* from the period 1990-2016. The index suggests high variation in the abundance of vermilion snapper, particularly for years previous to 2000 and a systematic decline in abundance between the early 1990s and 2016. This pattern culminates (2016) with vermilion snapper at about 16% of their average early 1990s abundance which is a stronger decline than is indicated by the current index used for stock assessment of the species.

Keywords: Vermilion snapper, abundance index, Bayesian model, catchability

1. Introduction

Fisheries harvest policies are typically based on the results of fitting population dynamics models with a variety of data types (Hilborn and Walters, 1992; Walters and Martell, 2004). One essential piece of information used in many fisheries stock assessments is a metric that indexes changes in total stock size through time (Maunder and Punt, 2004). These indices are typically derived from some form of fishery-dependent or -independent catch or count per-unit-effort data and are assumed to change in proportion to abundance, and thus reflect a scaled version of the total stock size. The resultant indices are used to "tune" stock assessment models, affecting estimates of population dynamics quantities and management reference points, such as harvest targets. Due to their importance for effective fisheries management, much attention has been paid to fisheries index development (e.g. Maunder and Starr, 2003; Maunder and Punt, 2004; Maunder et al., 2006); however, obtaining efficacious indices reflecting true changes in total stock size can be quite difficult (Kimura and Somerton, 2006).

Analysts face many challenges when developing abundance indices for stock assessments, particularly regarding the assumption of proportionality. One of the simplest representations of an abundance index is $I_t = N_t q$, where the index I_t is the product of the true abundance N_t and the catchability q (i.e. the proportion of N_t sampled). As long as catchability is constant through time, the assumption of proportionality is met and the index will have the desirable property of reflecting true proportional changes in abundance. However, when catchability varies, changes in q (or q_t) are confounded with changes in N_t, such that I_t may not adequately represent true abundance trends. Complicating this matter is that q (almost) always varies when sampling fish for a variety of reasons (Monk, 2013; Gwinn et al., 2016). For example, the influence of vessel effects on catchability and fishery-dependent indices is well known with variation in q related to variables such as vessel size, crew size, GPS technology, power of motor, and specific gear characteristics (Maunder, 2001; Maunder and Punt, 2004; Thorson and Ward, 2014). Fisheries-independent data for use in developing abundance indices are generally recognized as superior to fishery-dependent data (e.g. Dennis et al., 2015), however, these data are similarly vulnerable to variable q. For example, the catchability of reef fish with common baited traps can be strongly related to environmental variables such as temperature, depth, soak time, and substrate characteristics (Coggins et al., 2014; Bacheler et al., 2014; Shertzer et al., 2016). At best, these influences on catchability add noise into catch data but can also result in spurious patterns in I_t that do not reflect N_t when influential variables change systematically across space and time (e.g. Walters and Maguire, 1996; Ward, 2008; Langseth et al., 2016).

Shifts in sampling design elements such as the spatial frame of sampling and sampling methods commonly occur in long-term monitoring programs. Typically intended to improve sampling, these idiosyncrasies can also create challenges when developing fisheries indices (Conn et al., 2017). Similar to changes in catchability, changes in the spatial sampling frame and associated environmental characteristics can influence the component of the stock targeted by sampling (e.g. Walters and Maguire, 1996; Langseth et al., 2016). This is also true for changes in sampling gears and is the reason why there is a continuous development of methods to create spatially-explicit indices (e.g. Walters, 2003; Cao et al., 2017; Ducharme-Barth et al., 2018) and indices that integrate multiple sampling methods (e.g. Conn, 2010; Gibson-Reinemer et al., 2017; Kotwicki et al., 2018; Ono et al., 2018). Thus, methods that are robust to variation in catchability due to environmental variables, as well as shifts in sampling frame and sampling methods, are important tools for stock assessment scientists (Maunder and Piner, 2015).

The management of many economically important reef fisheries along the southeast U.S. Atlantic coast rely on indices derived from surveys using fishery-independent chevron traps. These traps have been used in this region since 1990 but were fitted with video cameras beginning in 2011 to further understand the quality of chevron trap catch data for indexing reef fish abundance, including species that do not enter traps (Bacheler et al., 2013a; Shertzer et al., 2016). The use of underwater video to assess the properties of various sampling gears is becoming increasingly common in the literature (e.g. Ward, 2008; Parker et al., 2016; Streich et al., 2018) and can result in a form of replicated count data that may be used to index abundance (Bacheler et al., 2013b; Schonbernd et al., 2014). However, appropriate statistical methods that create indices from data collected with these two sampling gears have yet to be developed. Combining data from multiple gears presents the opportunity for improved inference, but in this case, introduces two prominent challenges. Firstly, the paired samples of the chevron trap and video count are not fully independent. Although each gear represents an independent sample of the vulnerable fish community at the survey location, they are non-independent at the spatial scale of inference (i.e. the region) because samples are collected from the same locations and thus do not represent two independent measures of stock size at the regional level. Secondly, early research comparing trap catches to video counts revealed substantial variation between the two (Bacheler et al., 2013b), likely due to differences in how environmental conditions influenced the catchability of traps and videos for various species of fish (Bacheler et al., 2014; Coggins et al., 2014).

Here we develop a novel fishery-independent index of abundance that integrates paired trap catches and video counts into a single index of stock size using a Bayesian hierarchical formulation of a state-space model (SSM). The SSM has three key features that make it potentially useful for this application: (i) The model incorporates the baited trap catches and video counts into a single index that accounts for dependence between the gears;

(ii) the model accommodates changes in catchability due to temporal and spatial variation in the environment through the use of covariates and random effects of the observation processes; and (iii) the model can account for aspects of variable catchability due to shifts in the sampling frame by modeling temporal variation at the meta-population scale separate from spatial variation at the sub-population scale. We apply this model to vermilion snapper (*Rhomboplites aurorubens*) data collected along the southeast U.S. Atlantic coast by the Southeast Reef Fish Survey as an example and compare it to an index developed with the current methods (Conn, 2010) taken from the most recent stock assessment for the species (SEDAR, 2018). The current method of index development (i.e. Conn, 2010) treats the chevron trap catches and camera counts as independent measures of the stock and does not explicitly account for shifts in the spatial frame of the surveys, thus offering a useful comparison for the SSM method.

2. Methods

2.1 Overview of methods

We organize our methods in four main parts. First, we describe the sampling design and data treatment in the context of the Southeast Reef Fish Survey sampling of *R. aurorubens* along the southeastern coast of the U.S., which is the motivation behind our model; second, we describe the general model structure, covariate structure, model fit evaluation and model optimization methods used in the example analyses; third we describe how we compare our index to the current index used for stock assessment of *R. aurorubens*; and last we describe validation methods of our model on a set of simplified simulated data sets.

2.2 Sampling design

R. aurorubens count data were collected along the southeast United States Atlantic Coast from Florida to North Carolina by the Southeast Reef Fish Survey (Fig. 1a). All baited traps were set on or near hard-bottom reef locations. There were 15,629 chevron trap samples available covering a period of 27 years (1990-2016). The number of locations sampled has varied substantially among years due to program development and funding. In the early years, the number of samples collected annually was typically in the range of several hundred; however, this number has expanded severalfold to over thirteen hundred in the most recent years. Along with increased sampling intensity, the sampling frame of the program has expanded in both the latitudinal and longitudinal directions, thus shifting the sub component of the stock vulnerable to sampling (for a detailed description of the sampling frame shift, see Appendix A). Traps were set no closer than 200 m from one another to maintain spatial independence relative to fish movement, and at depths between 13 and 115 m. All trap sites used for this analysis were selected randomly from a defined sampling frame of hard-bottom sampling points (Bacheler et al., 2014). Traps were baited with menhaden and set for approximately 90 min. For the time period of 2011-2016, the chevron traps were fitted with an outward-looking video camera (Fig. 1b) resulting in 7,644 41-frame video samples (Fig. 1c). The camera (Canon Vixia HFS200 in 2011 - 2014 and GoPro Hero 3 or 4 in 2015 and 2016) recorded at least 20 minutes of video from the bottom, and videos were read according to Schobernd et al. (2014). Specifically, a series of video frames spaced 30 seconds apart were read 10 to 30 minutes after the trap landed on the bottom. This resulted in 41 replicate camera samples and one baited trap sample per site.

2.3 Data and treatment

For trap data, we analyzed the un-transformed catch and for the video data, the sum of the counts across the 41 camera frames (*SumCount*). We chose to use the *SumCount* of the camera data because *SumCount* changes linearly with the *MeanCount* (Bacheler and Carmichael, 2014), which is often the preferred camera metric (Conn, 2011; Schobernd et al., 2014; Campbell et al., 2015), and using the *SumCount* preserves the discrete nature of the camera counts allowing for the use of derivations of the Poisson distribution to describe both the chevron trap and camera observation processes. We applied several data filters to either simplify predictor variables, remove records with missing predictor variables, or to remove unusual values. Detailed methods of the data cleaning process are reported in Appendix B.

2.4 Model development

The model was formulated with three distinct hierarchical layers such that the relative abundance at the meta-population level (representing our index of interest, denoted as I_t) was modeled separately from the relative abundance at the sub-population level (i.e. at sample sites, denoted as $n_{s,t}$) and separately from the observation processes. By modelling the meta-population level abundance separately from the sub-population abundance, we were able to isolate the fishery index of interest from components of spatial variation among sampling locations. This is the key component that separates shifts in spatial sampling frame relative to latitude, longitude, and depth from the changes in the average meta-population abundance. Furthermore, by modeling the abundance processes and observation processes with separate sub-models we were able to separate observation error from process error and account for systematic variation in catchability.

The most general version of our model describes the latent meta-population level relative abundance (hereafter referred to simply as abundance, I_t) for each year as an independent, freely estimated parameter represented as:

$$\log\ I_t\ \sim\text{Normal}\ \ 0, 100 \tag{1}$$

however more constrained formulations that assume that meta-population level abundance is a random effect among years (i.e. $\log\ I_t\ \sim\text{Normal}\ \bar{I}, \sigma$), a Markovian random walk (i.e. $\log\ I_t\ = \log\ I_{t-1}\ + r_t$, where $r_t \sim \text{Normal}\ \bar{r}, \sigma$), or any population dynamics model (e.g. logistic model, age-structured model) could be applied based on the intended use of the index. If the index will be used to fit a more complex population dynamics model for stock assessment, it may be desirable to impose as little constraint on the temporal pattern of the index as possible; thus, we present the model that assumes indexes are independence among years to represent this case.

Spatial variation in abundance across sample sites each year (sub-population level) was modeled on the log scale as:

$$\log(n_{s,t}) = \log\ I_t\ + co\ {}_{s,t}^{n} + \varepsilon_{s,t}^{abun} \tag{2}$$

where the term $\log\ I_t$ is the year specific intercept of the linear model, $co\ {}_{s,t}^{n}$ is a linear combination of spatial covariates, and $\varepsilon_{s,t}^{abun}$ describes random site-level variation in abundance that is not explained by the covariate structure.

We approximated the baited trap catches ($c_{s,t}^{trap}$) and the camera *SumCounts* ($c_{s,t}^{cam}$) as deviates drawn from Poisson log-Normal distributions, which are similar in character to negative binomial distributions (Ntzoufras, 2009, p. 315-317), but can demonstrate better mixing properties than negative binomial distributions when applied in Bayesian programs such as JAGS. We specified these models as:

$$c_{s,t}^{trap} \sim \text{Poisson}\left(e^{\log(n_{s,t}) + cov_{s,t}^{trap} + \varepsilon_{s,t}^{trap}}\right) \tag{3}$$

$$c_{s,t}^{cam} \sim \text{Poisson}\left(e^{\log(n_{s,t}) + cov_{s,t}^{cam} + \varepsilon_{s,t}^{cam}}\right) \tag{4}$$

where the mean on the log scale is the site-specific abundance $n_{s,t}$ plus a linear combination of environmental and/or sampling covariates (i.e. $co\ {}_{t,j}^{trap}$ and $co\ {}_{t,j}^{cam}$) to account for systematic variation in catchability. The parameters $\varepsilon_{s,t}^{trap}$ and $\varepsilon_{s,t}^{cam}$ are gear-specific log-Normal distributed random observation errors modeled as, $\varepsilon_{s,t} \sim \text{Normal}\ \ 0, \sigma$, with a mean of zero and an estimated standard deviation specific to each sampling method (i.e. σ^{trap} and σ^{cam}).

2.5 Model covariates

To account for systematic variation in our count data, we incorporated a suite of covariates into the abundance and observation sub-models. We selected covariates based on two key considerations. Our first consideration was to separate covariates that influenced the spatial distribution of fish from those that influenced temporal patterns in fish abundance. This was important because spatial and temporal patterns of abundance are modeled in two separate hierarchical layers (i.e. equation 1 and 2) to create a distinction between the fishery index, i.e. temporal patterns in abundance at the meta-population level (I_t), from spatial variation in the data due to patterns in the spatial distribution of fish ($n_{s,t}$) and shifts in the sampling frame through time. Thus, we included nonlinear (quadratic) effects of latitude (*lat* and lat^2), longitude (*lon* and lon^2) and depth (*dep* and dep^2), as well as the potential interaction between latitude and longitude as covariates of local-scale abundance. We included both main and quadratic effects of these variables to account for any optimal ranges in latitude, longitude and depth within our sampling frame that vermilion snapper may prefer. The interaction between latitude and longitude was included to allow any preferred range of one variable to be dependent on the other. For example, if vermilion snapper demonstrated a preferred distance from shore, a positive interaction between latitude and longitude could approximate this spatial distribution. Lastly, we included a measure of bottom relief (*rel*) and the percent of the substrate that was hard-bottom (*sub*) as these habitat features may affect the local density of fish. Spatial covariates of abundance were incorporated into the model as:

$$co\ {}^{n}_{s,t} = \beta_1 lat_{s,t} + \beta_2 lat_{s,t}^2 + \beta_3 dep_{s,t} + \beta_4 dep_{s,t}^2 + \beta_5 lon_{s,t} + \beta_6 lon_{s,t}^2 + \beta_7 lat_{s,t} lon_{s,t} + \beta_8 rel_{s,t} + \beta_9 sub_{s,t}. \tag{5}$$

Our second key consideration was to separate covariates of the abundance and observation processes. This was important because our model likely has limited ability to disentangle systematic patterns in abundance from systematic patterns in catchability when they are similar. Thus, we do not expect to be able to resolve the effects of covariates that have similar influences on patterns in abundance and catchability (Barker et al., 2017). Given this limitation, the most useful covariates for predictive purposes are those that either, (i) only influence abundance or catchability, or (ii) have very different influences on abundance and catchability. Thus, we included main and quadratic effects of trap soak time (*E* and E^2), and main and quadratic effects of temperature (*temp* and $temp^2$) as continuous variables; we included water turbidity (*turb*) as a categorical variable with two levels (low as *turb* = 0 and high as *turb* = 1); and we included current direction as a categorical variable with three levels (current away from the lens and trap opening indicated by $dir1 = 0$ and $dir2 = 0$; current towards the side of camera and trap indicated by $dir1 = 1$ and $dir2 = 0$; and current away from the lens and trap mouth indicated by $dir1 = 0$ and $dir2 = 1$). We incorporated these covariates into our chevron trap observation model as:

$$co_{s,t}^{trap} = \eta_1 E_{s,t} + \eta_2 E_{s,t}^2 + \eta_3 temp_{s,t} + \eta_4 temp_{s,t}^2 + \eta_5 turb_{s,t} + \eta_6 dir1_{s,t} + \eta_7 dir2_{s,t}. \tag{6}$$

In the camera catchability sub-model, we included turbidity, current direction, and main and quadratic effects of bottom temperature as:

$$co_{s,t}^{cam} = \varphi_1 + \varphi_2 turb_{s,t} + \varphi_3 dir1_{s,t} + \varphi_4 dir2_{s,t} + \varphi_5 temp_{s,t} + \varphi_6 temp_{s,t}^2 + \nu_t \tag{7}$$

where the intercept φ_1 allows for a systematic difference in the catchability of the camera relative to the chevron trap. The parameter ν_t is a fixed value (i.e. log(1.72), Bacheler and Ballenger, 2018) that accounts for the increased field of view of the video cameras used in 2015 and 2016. All continuous covariates were centered on zero and scaled to one standard deviation with the exception of the effort covariate. We scaled effort by subtracting 60 and dividing by 60 to ease interpretation (effects are relevant to one hour). The absolute value of the Pearson correlation coefficient between covariates were all < 0.6 with the exception of latitude and longitude ($\rho = 0.87$); however, we chose to retain both covariates as we expected that they would both be important for describing site-level variation in abundance and the correlation would not impact adversely on the abundance index after model regularization. All covariate definitions are provided in Table 1, the correlation matrix of all covariates is presented in Table B1 (of Appendix B) and JAGS model code and fitting methods are provided in Appendix C.

2.6 Model fitting and prior specification

The posterior distributions of all parameters were estimated using a Gibbs sampler implemented in JAGS (Plummer, 2003). We called JAGS from program R (R Core Team, 2015) using the library R2jags (Su and Yajima, 2015). All prior distributions of log-scale covariate effect parameters, including model intercepts and the fisheries index I_t were specified as diffuse normal distributions (N[0,100]). Standard deviation parameters including all random effects were specified as scaled half Student-t distributions with input parameter values chosen to stabilize fit while inducing negligible parameter shrinkage (i.e. $\mu = 0, \tau = 2.78, k = 2$). Inference was drawn from 10,000 posterior samples taken from two chains of 10^6 samples. We discarded the first 500,000 values of each chain to remove the effects of initial values and thinned the chain to every 100th value. Convergence of all models was diagnosed by visual inspection of trace plots and Gelman-Rubin statistic ($\hat{R} \leq 1.1$ indicate model convergence, Gelman et al. 2004).

2.6 Model fit evaluation and regularization

There are two common purposes of models in applied ecology, (i) causal explanation and (ii) empirical prediction, and the same model will often not perform well for both purposes (Shmueli, 2010; Authier et al., 2016). A model used for the purpose of explanation requires that the uncertainty in parameter estimates are appropriately accounted for such that the realized 95% credible interval coverage is equivalent to the *a priori* expectation (i.e. true parameter value contained within 95% CI 95% of the time). In practice, this requires that the model error structure adequately explains the residual error and, thus, can be determined with model fit tests. Alternatively, the optimal predictive model will often be a model where the covariate effect estimates are removed or shrunk towards zero through a process termed regularization (e.g. Reineking and Schroder, 2006; Hooten and Hobbs, 2015). Thus, some level of increased bias is accepted for the predictive advantage of decreased variance. Although optimal prediction of our index is our main purpose, we were also interested in the influence of our covariates on abundance and catchability. Thus, we first used a posterior-predictive check to determine an adequate error structure for our fully parameterized model for the purpose of evaluating covariate relationships (termed 'global model'). Covariates were considered statistically different than zero when the associated 95% Bayesian credible intervals (quantile based) did not include zero. Second, for our best error structure, we used a process termed Stochastic Search Variable Selection (SSVS) to induce shrinkage of covariate effects and generate a model with optimal predictive properties to produce the fisheries index (termed 'reduced model'). Using SSVS to produce models with desirable predictive properties was first introduced by George and McCulloch (1993) but has been thoroughly discussed in more recent ecological literature by O'Hara and Sillanpaa (2009), Tenan et al. (2014), and Hooten and Hobbs (2015).

We evaluated model fit of the global model for eight general model error structures with Bayesian p-values (Kéry, 2010). The Bayesian p-value is a posterior-predictive check that provides a measure of under- or over-dispersion of the data relative to the model (Kéry, 2010; Hooten and Hobbs, 2015). The eight error structures were models that either included or excluded the random variables ε^{abun}, ε^{trap}, and/or ε^{cam}. We performed our model fit evaluation by simulating our data directly from each model for each Markov Chain Monte Carlo (MCMC) iteration and calculating a Pearson residual between the simulated and expected values (i.e. predicted χ^2) and observed and expected values (i.e. observed χ^2). The simulated data are considered "perfect" because they are generated directly from the model and, thus, the resulting Pearson residual represents the fit of the model when all model assumptions are perfectly met (Kéry, 2010). We then created a fit metric that is equal to zero when the Pearson residual was greater for the observed data than the simulated data and is equal to one, otherwise. The Bayesian p-value was then calculated as the mean of the posterior sample of the fit metric for each data type, where a mean of 0.5 indicates perfect model fit to the data and a mean approaching 1 or 0 indicates under- or over-dispersion of the data relative to the model, respectively.

We chose the procedure of SSVS to produce the reduced model and optimize prediction because preliminary analysis indicated that processing times in excess of four days may be expected for the example data. Thus, many common approaches to variable selection that either employ iterative model runs such as information theoretic methods (e.g. AIC, WAIC, DIC, etc.) or k-fold cross validation are prohibitive. Therefore, we employed SSVS which took approximately five days to complete two million MCMC iterations. We applied the SSVS method for each covariate effect parameter in Eq. 5, 6, and 8 to invoke parameter shrinkage. Specifically, we applied a hierarchical structure for each of our covariate priors that is conditional on a random effect indicator variable as: $P(\beta_j|w_j) \sim \mathrm{Normal}(0, \sigma_j)$, where $\sigma_j = 100w_j + 0.01$. The variable w_j is a random effect for each covariate that has a prior distribution of $P(w_j) \sim \mathrm{Bernoulli}\ \ 0.5\ \ $, such that when $w_j = 1$, $\sigma_j = 100.01$, approximating a standard uninformative prior on the covariate effect parameter β_j. Alternatively, when $w_j = 0$, $\sigma_j = 0.01$ which approximates a highly informative prior for a $\beta_j \cong 0$. Thus, the conditional prior creates a region of high probability around zero similar to ridge regression or a "slab and spike" prior (Tibshirani, 1996; Ishwaran and Rao, 2005). Furthermore, the posterior mean of w_j can be interpreted as the relative support of a non-zero value of β_j

similar to the posterior probabilities for different model structures obtained via reversible jump methods (e.g. Hillary 2011). However, one advantage of the SSVS process is that model predictions are automatically model averaged, providing a more refined level of regularization. Thus, we produced the index from the regularized model that included all covariates as well as the indicator variables and conditional priors.

2.7 Comparison of indices

To increase our insight into the value of the SSM index, we compared it to an index developed for use in the most recent stock assessment of *R. aurorubens* (Conn, 2010; SEDAR, 2018; hereafter referred to as the "Conn index"). The Conn index utilized a hierarchical analysis to combine multiple indices into a single index for use in stock assessment (Conn, 2010). The method requires prior knowledge of sampling error and constraints on process error, which may be difficult to inform. A detailed description of methods is provided in Conn (2010). In brief, the approach treats multiple, independently developed indices of abundance as measurements of the same underlying quantity (the true relative abundance), with each index subject to sampling and process error. For this application toward *R. aurorubens*, two indices were combined, one developed from video gear (Cheshire et al., 2017) and one from chevron traps (Bubley and Smart, 2017). Thus, the data were the same as those used for the SSM index, and the primary difference in methodology is that the Conn (2010) approach operates on previously created indices, whereas the approach presented here operates at the level of the observed data. By doing so, our approach more naturally accounts for the lack of independence between the gears that might be expected when sampling co-occurs (i.e., cameras are mounted on traps) and the potential impact of non-independence between the sampling methods on the index uncertainty.

To simplify comparison of the indices we used a parametric bootstrap method to estimate the linear slope of population change through time for each index. For each year and index, we sampled 10,000 random values drawn from log-Normal distributions with the means specified as the annual index point estimates and the associated standard deviations. For each random sample, we use least square methods to estimate the intercept and slope of the index through time on the log scale. This results in a probability distribution of the log-scale linear trend for each index.

2.7 Model validation

To validate the efficacy of our model, we addressed two questions with a simulation experiment; (i) is our model identifiable and (ii) does it produce unbiased parameter estimates when applied to perfect data? Our methods were to define a data-generating model, simulate multiple datasets, and analyze the simulated datasets with the data-generating model. Our data-generating model is presented in Table 2 and was identical to the model described for our *R. aurorubens* analysis where the temporal abundance process is modeled as an independent variable for each year (Eq. 1, Table 2) drawn from a log Normal distribution with a mean and standard deviation set to represent observed variation in the *R. aurorubens* index. However, we excluded the random variables ε^{abun}, ε^{trap}, and ε^{cam} to reduce the limitations of computation time. We simulated nine covariate relationships influencing sub-population level abundances, trap catchability and camera catchability (θ_{1-9}, Table 2), where the simulated covariates, *x*1-*x*9 (Table 2) are nine separate vectors of random draws from normal distributions with mean of zero and standard deviation of one to simulate generic centered and scaled covariates. We chose covariate effect sizes arbitrarily to represents different levels of effects and the absence of effects. The input values for the simulated covariate effects were, $\theta_{1,4,7} = 1$, $\theta_{2,5,8} = -.5$, and $\theta_{3,6,9} = 0$. Simulated data sets were fit with the data generating model. We report the mean absolute error as a measure of bias and, to evaluate if the credible interval coverage was appropriate, we reported when the true parameter value was excluded from the 95% credible intervals for each iteration of the simulation. All simulation code is included in Appendix D.

3. Results

3.1 Simulation study

Our simulation study revealed that the SSM model does indeed return unbiased estimates of meta-population abundance (I_t) and covariate relationships with appropriate credible intervals. The mean absolute error of all covariate effect estimates centered on zero (Fig. 2a) and the true value was included in the 95% Bayesian credible intervals between 91.5 and 96.5% of the time. The results for the simulated relative abundance index were similar with little to no systematic bias (Fig. 2b) and 95% credible interval coverage of the true index value for 90.2 to 97.2% of simulation interactions. These results indicate that the model is identifiable and produced unbiased parameter estimates with appropriate levels of uncertainty.

3.2 Vermilion snapper analysis

All eight error structures of the global model fit to *R. aurorubens* count data converged after 10^6 iterations and each required up to 96 hours of computer processing of two MCMC chains run in parallel. Our posterior-predictive check indicated that three model error structures adequately fit the data (models 1, 2, & 3 in Table 3). All of these models included a site-specific random effect in the abundance sub-model (ε^{abun}) and either a site-specific random effect in the camera sub-model (ε^{cam}), the trap sub-model (ε^{trap}) or both. Global model 2 and 3 had the simplest structure of only two random effects, allowing us to exclude global model 1 as the most parsimonious error structure. Global model 3 produced a lower model deviance than global model 2 and the posterior estimates of the standard deviation of the random effect ε^{trap} were near zero when estimated with global model 1 offering additional support for global model 3 as the best error structure (See Appendix E). Thus, we used global model 3 for the remaining analyses in this paper. This model included the random effects ε^{abun} and ε^{cam} and excluded the random effect ε^{trap} (Table 3).

Most of the covariates evaluated with global model 3 (had a statistically significant influence on abundance or catchability (Table 4). We found that the strongest determinants of sub-population abundance ($n_{s,t}$, equation 3) were the latitude, longitude, depth, and percent hard-bottom substrate at the sample site (Fig. 3, Table 4). However, the interaction between the latitude and longitude of the location was also a strong influencer (Fig. 3, Table 4). The only covariates of abundance that were not statistically different than zero were the main effect of depth (*lon*^{2}, $\beta_6 \approx 0$, Table 4) and the bottom relief at the site (*rel*, $\beta_8 \approx 0$, Table 4). The catchability of the chevron trap was found to be strongly related to the amount of time the trap was set (*E*, $\eta_1 > 0$, Table 4) and its square (*E*^{2}, $\eta_2 < 0$, Table 4). This relationship suggests that the number of *R. aurorubens* captured increases with trap soak time to a maximum (at ~110 minutes of soak time), beyond which the catch declines (Fig. 4a). Temperature and its square (*temp* η_3 φ_5 and *temp*^{2} η_4 φ_6, Table 4) defined a pattern in catchability for both the chevron trap and camera that stayed fairly constant at lower temperatures and increased rapidly at temperatures greater than ~25° C (Fig. 4b); however, this pattern was less pronounced for the camera (Fig. 5a). Current direction influenced the catchability of both the chevron trap and camera (η_6, η_7, φ_2, and φ_3, Table 4) but was a stronger effect for the camera (Fig. 4c and 5c). For both gears, the lowest catchability was when the current direction was towards the mouth of the trap and camera lens, while it was the highest when the current direction was away from the trap mouth and camera lens (Fig. 4c and 5c). Finally, higher levels of turbidity increased the catchability of the camera (Table 4, Fig. 5b), but had no influence on chevron trap sampling efficiency (Table 4).

Our model regularization procedure resulted in a reduced model with the effective removal of eight covariates relative to the global model (w_j << 0.05, Table 4). For example, the three non-statistically significant covariates (i.e. β_6, β_8, and η_5, Table 4) had w_j values equal to zero. Additionally, six statistically significant covariates (i.e. β_5, η_6, η_7, φ_3, φ_4, and φ_5) were effectively removed from the model with w_j values ≤ 0.23. Although statistically different than zero in the global model, these covariates tended to have small effects sizes with relatively high levels of uncertainty (Table 4). We found that the value of several covariates with high inclusion probabilities (i.e. $w_j \approx 1.00$) differed between the global and reduced model (i.e. β_1, β_3, β_4, φ_2, Table 4). This is likely a result of some level of multicollinearity among covariates (particularly for latitude and longitude covariates). Our index generated from the reduced model tended to be equally precise as the index generated from the global model with an average coefficient of variation of 0.40 (range across years = 0.35, 0.52) and 0.42 (range across years= 0.36, 0.51), for the reduced and global model, respectively. The observed different of 0.02 is likely not large enough to be biologically relevant.

Our index of *R. aurorubens* suggests high annual variation in abundance (Fig. 6a). For example, our model predicted a nine-fold increase in abundance between 1990 and 1991. After 1991, annual variation in abundance ranges between a 168% increase in 1994 and an 87% decrease in 2003. This level of variation was fairly consistent across the time series (Fig. 6a). The index also suggests a linear decline in *R. aurorubens* since the 1990s. A bootstrapped slope of this decline on the log scale was statistically negative (μ = -0.60, 95% CI = -0.69, -0.51, Fig. 6b) and suggests that *R. aurorubens* are currently (2016) at about 16% of their average abundance in the early 1990s (i.e. 1990-1995).

The SSM index described a very similar pattern in abundance to the index generated from the methods of Conn et al. (2010); however, there were some differences (Fig. 6c). For example, the Conn index had a smaller average coefficient of variation than the SSM index (Conn = 0.35, SSM = 0.40) and demonstrated some differences in year-to-year variation in the index, however these differences were subtle (Fig. 6a, and b). Most notably, the SSM index described a stronger pattern of decline across the time frame of the data than the Conn index. The bootstrapped slope of the Conn index was statistically different than zero but nearly half the value of the SSM index (μ = -0.36, 95% CI = -0.49, -0.24, Fig. 6d). This decline suggests that *R. aurorubens* in 2016 are at approximately 33% of their mean abundance in 1990-1995, which is over twice the value predicted by the SSM index.

4. Discussion

We developed a state-space model that integrates data from multiple gears that are non-independent relative to the sampling process into a single fisheries index. We demonstrated its use for indexing *R. aurorubens* abundance from paired count data derived from underwater video cameras and catch data from traditional fisheries-independent baited traps. The method provides a means to account for random and systematic variation in the catchability of both sampling gears and adjusts for aspects of non-proportionality due to changes in the spatial frame of sampling expected when monitoring programs are developing. The model produced unbiased estimates of meta-population level relative abundance when the model is correctly specifies and demonstrated good fitting properties. We see this modelling approach as a flexible tool that has the potential to be useful for generating fisheries indices for stock assessment for a variety of fish species sampled with paired non-independent gears, particularly traditional gears paired with underwater video cameras.

One of the key strengths of the SSM model is its ability to account for variation in catchability for both sampling gears. The importance of the covariates of catchability was highlighted by our SSVS model regularization procedure that indicated the optimal predictive model included many of these covariates. Furthermore, it is important to note the advantage that multiple sampling gears provide in addition to covariates when estimating parameters of state-space models. The addition of multiple gears, and thus, multiple observation sub-models to the SSM provides contrast between the residual error of each gear and the covariates that describe it. This contrast between patterns in residual error provides greater information for the model to disentangle process error from observation error. For example, when relative catch rates of the gears deviate from the expected value differently, at least one of these deviations must be due to observation error. Alternatively, when only one observation sub-model is included in the SSM, the pattern in the residual and the *a priori* choice of covariates to describe it are the only sources of information that the model has to distinguish observation error from process error. Furthermore, it is the inclusion of multiple sampling methods that allows a model that assumes independence of the index among years to be identifiable, which is a desirable option when the index will be further used to fit a stock assessment population dynamics model. With only a single observation model, a more confining structure must be imposed on the index to obtain identifiability, such as a Markovian process commonly applied in state-space models (e.g. Clark and Bjornstad 2004; Jiao et al. 2008). Furthermore, greater contrast between the variation in catchability of the gears will provide the most informative data and likely result in greater index precision. Thus, the inclusion of multiple gears can be quite advantageous in this context.

In our case, the direction of the covariate effects on the observation sub-models tended to be consistent with previous research on the sampling efficiency of these gears (Bacheler et al., 2013b; Bacheler et al., 2013c; Coggins et al., 2014; Shertzer et al., 2015). This comes as no surprise because we based our choice of covariates, in part, on these studies. For example, we found a dome shaped relationship between trap soak time and catchability that resulted in a maximum catch at about 110 min of soak time. A similar relationship has been found for other reef fish species such as black sea bass (*Centropristis striata*) and is likely the result of entry and exit rates of fish into and out of the trap that change through time inversely proportional to each other (Bacheler et al., 2013c; Shertzer et al., 2015). Similarly, we found that the effect of temperature on both the trap and camera was positive with the appearance of a threshold-like response at ~25° C. Bacheler et al. (2014) found a comparable relationship between chevron trap catch of *R. aurorubens*, with a threshold at ~20° C, but did not detect this relationship for cameras. We observed a positive relationship between camera counts of *R. aurorubens* and turbidity which has also been observed for red snapper (*Lutjanus campechanus*, Coggins et al., 2014). Although this response may be counterintuitive, our data filtering process was similar to Coggins et al. (2014), which removed high turbidity data points that demonstrably impacted the counting of fish in video frames; thus, this effect may be a result of fish behavioral changes with variation in water clarity (e.g. McMahon and Holanov, 1995; De Robertis et al., 2003; Andersen et al., 2008).

Another important benefit of our SSM is that it can account for shifts in the sampling frame from year to year. For example, over the length of time of the Southeast Reef Fish Survey sampling program, the number of chevron traps set each year has systematically increased as the program expanded (particularly since 2011). The expanding of the program has led to changes in the distribution of traps relative to latitude, longitude, and depth (Fig. 7a, b, c), resulting in variability in the mean covariate values among years with apparent systematic increases in depth and decreases in latitude over the life of the program (Fig. 7). Our model accounts for this shift by modeling the index I_t at a fixed point in space (relative to latitude and longitude) and for a fixed depth. The limitation of this method is that it only accounts for the shift in the sampling frame relative to these covariate relationships. Thus, any unaccounted for systematic spatial patterns in abundance that coincide with the expansion of sampling may still result in a biased index. This provides high incentive to determine the important drivers and structure of the spatial distribution when using this method, which could include environmental covariates as well as modeling a spatially autocorrelated residual. In our case, inspection of the residual did not reveal any non-random patterns in the spatial distribution relative to the covariates and the expanding sampling design, nor did calculating the index from only the sample locations contained within a core area that was sampled every year produce an index substantially different from the one presented in Figure 6a. These two diagnostics suggest low risk of a biased index due to shifting sampling frame, in our case (see Appendix F for details about the diagnostics). However, unaccounted for changes in the average abundance due to shifts in the sampling frame or shifts in the species distribution should be carefully considered when applying this method. This is particularly the case if the count data are derived from a fishery-dependent source, where preferential sampling that is often related to fish density is common (Pennino et al. 2018). As accounting for preferential sampling in the analysis of count data can be analytically challenging (e.g. Conn et al. 2017, Pennino et al. 2018), we recommend, first, that appropriate spatial designs be used for sampling and, when this is not possible, that appropriate diagnostics be used to evaluate the risk of induced bias.

4.1 Model extensions

There are several possible extensions to the SSM that would allow it to accommodate various idiosyncrasies of different data sets worth discussing. One prominent extension is to accommodate various levels of zero inflation in the sub-population model. Our example data set was zero inflated with 75% and 70% zeros in the trap catches and camera counts, respectively. We approximated the structural component of these zeros with the log-normal random effect ε^{abun}; however, this method makes explicit the assumption that these potential zeros are actually very small non-zero values. Our model fit test suggested that this model structure provided adequate fit to our example data; however, another option is to model a zero-inflated spatial abundance process by including a shared Bernoulli variable in both (or all) observation models as:

$$c_{s,t}^{trap} \sim \text{Poisson}\left(z_{s,t} e^{\log(n_{s,t}) + cov_{t,j}^{trap} + \varepsilon_{s,t}^{trap}}\right) \tag{8}$$

$$c_{s,t}^{cam} \sim \text{Poisson}\left(z_{s,t} e^{\log(n_{s,t}) + cov_{t,j}^{cam} + \varepsilon_{s,t}^{cam}}\right) \tag{9}$$

where $z_{s,t}$ is a latent random variable distributed as, $z_{s,t} \sim \text{Bernoulli}(\psi_{s,t})$. The Bernoulli probability of a non-zero abundance could be modeled independently for each year, as a function of a set of spatial covariates with a logit link, or as a function of $co\ {}_{s,t}^{n}$ to create a formal relationship between the spatial abundance and occurrence processes (e.g. Smith et al.

2012). Additionally, a zero-inflated observation process could be modelled by specifying unique Bernoulli processes for each observation sub-model.

Another prominent extension would be to model spatiotemporal variation in patterns in abundance more explicitly. For example, applying a multivariate normal prior to $\varepsilon_{s,t}^{abun}$ to explicitly model spatial auto-correlation could be used to improve the predictive potential of the model and to better account for changes in the spatial distribution of sampling among years. Furthermore, specifying covariate effects of the spatial abundance process as random effects across years could be used to evaluate and account for non-stationarity in these relationships through time. These are only a few examples of potentially useful extensions to our model that could improve its application to various settings. Thus, we see this model as a foundation that could be easily extended to accommodate the nuances of a variety of data structures and contexts.

4.2 Management implications

The application of our model to *R. aurorubens* revealed a systematic decline in abundance across the time period of 1990-2016. This decline was similar to, but stronger than the decline described by the Conn index (Conn et al., 2010). This discrepancy between the two indices is in the direction that would be predicted given the systematic expansion of the sampling design into latitudes, longitudes, and depths of greater abundance, and given that the Conn index does not account for this systematic expansion, while the SSM does (Figure 7d). The difference also suggests that the *R. aurorubens* stock may have a lesser ability to compensate for the reductions in density due to harvest (i.e. lower productivity) than would be indicated by the Conn index. It is difficult to predict the effect this would have on management recommendations; however, we may expect the use of the SSM in a formal stock assessment to result in more conservative harvest regulations to meet management targets such as Maximum Sustainable Yield (Beverton and Holt, 1957) and maintain acceptable levels of risk of overfishing (Zhou et al. 2016). Although an explicit comparison between the outcomes of formal stock assessments with each index would be necessary to know this for sure.

Acknowledgements

We thank SERFS staff members (SCDNR-MARMAP and SEAMAP-SA and NMFS/SEFSC/Beaufort) and volunteers for field assistance and the captains and crew of the NOAA Ships *Pisces* and *Nancy Foster*, R/V *Palmetto*, R/V *Sand Tiger*, and R/V *Savannah* for sampling platforms. We also thank J. Ballenger, R. Cheshire, T. Kellison, M. Reichert, and T. Smart for providing comments on earlier versions of this manuscript, as well as J. Thorson and two anonymous reviewers who's comments substantially improved this manuscript. Funding was provided by the US National Marine Fisheries Service. The use of trade, product, industry, or firm names, products, software, or models, whether commercially available or not, is for informative purposes only and does not constitute an endorsement by the US government or NOAA. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the position of any government agency.

References

Andersen, M., Jacobsen, L., Grønkjær, P., and Skov, C. 2008. Turbidity increases behavioural diversity in northern pike, *Esox lucius* L. during early summer. Fisheries Management and Ecology 15, 5-6.

Authier, M., Saraux, C., and Peron, C. 2016. Variable selection and accurate predictions in habitat modelling: a shrinkage approach. Ecography 39, 1-12.

Bacheler, N.M., and Ballenger, J.C. 2018. Decadal-scale decline of scamp (Mycteroperca phenax) abundance along the southeast United States Atlantic coast. Fisheries Research 204, 74-87.

Bacheler, N.M., and Carmichael, J.T. 2014. Southeast reef fish survey video index development workshop. Final Report, National Marine Fisheries Service and South Atlantic Fisheries Management Council.

Bacheler, N.M., Schobernd, C.M., Schobernd, Z.H., Mitchell, W.A., Berrane, D.J., Kellison, G.T., and Reichert, M.J.M. 2013a. Comparison of trap and underwater video gears for indexing reef fish presence and abundance in the southeast United States. Fisheries Research 143, 81-88.

Bacheler, N.M., Bartolino, V., and Reichert, M.J.M. 2013b. Influence of soak time and fish accumulation on catches of reef fishes in a multispecies trap survey. Fishery Bulletin 111, 218-232.

Bacheler, N.M., Schobernd Z.H., Berrane, D.J., Schobernd, C.M., Mitchell, W.A., and Geraldi, N.R. 2013c. When a trap is not a trap: converging entry and exit rates and their effect on trap saturation of black sea bass (*Centropristis striata*). ICES Journal of Marine Science 70, 873-882.

Bacheler, N.M., Berrane, D.J., Mitchell, W.A., Schobernd, C.M., Schobernd, Z.H., Teer, B.Z., Ballenger, J.C. 2014. Environmental conditions and habitat characteristics influence trap and video detection probabilities for reef fish species. Marine Ecology Progress Series 517, 1-14.

Barbieri M.M., and Berger J.O. 2004. Optimal predictive model selection. Annals of Statistics 32, 870–897.

Barker, R. J., Schofield, M. R., Link, W. A., and Sauer, J. R. 2017. On the reliability of N-mixture models for count data. Biometrica, doi:10.1111/biom.12734

Beverton, R. and Holt, S. 1957. On the Dynamics of Exploited Fish Populations. Chapman and Hall, London.

Bubley, W.J. and Smart, T.I. 2017. Vermilion snapper fishery-independent index of abundance in US South Atlantic waters based on a chevron trap survey (1990-2016). SEDAR55-WP02. SEDAR, North Charleston, SC. 15 pp.

Campbell, M.D., Pollack, A.G., Gledhill, C.T., Switzer, T.S., and DeVries, D.A. 2015. Comparison of relative abundance indices calculated from two methods of generating video count data. Fisheries Research 170, 125-133.

Cao, J., Thorson, J.T., Richards, R.A., and Chen. Y. 2017. Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. Canadian Journal of Fisheries and Aquatic Sciences 74, 1781-1793.

Cheshire, R., Bacheler, N., and Shertzer, K. 2017. Standardized video counts of Southeast U.S. vermilion snapper (*Rhomboplites aurorubens*) from the Southeast Reef Fish Survey. SEDAR55-WP01. SEDAR, North Charleston, SC 21 pp.

Clark, J.S., and Bjornstad, O.N. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. Ecology 85, 3140-3150.

Coggins, L.G, Bacheler, N.M., and Gwinn, D.C. 2014. Occupancy models for monitoring marine fish: a Bayesian hierarchical approach to model imperfect detection with a novel gear combination. PLoS ONE 9, e108302. doi:10.1371/journal.pone.0108301.

Conn, P.B. 2010. Hierarchical analysis of multiple noisy abundance indices. Canadian Journal of Fisheries and Aquatic Sciences 67, 108–120.

Conn, P. B. 2011. An Evaluation and Power Analysis of Fishery Independent Reef Fish Sampling in the Gulf of Mexico and U. S. South Atlantic. NOAA Tech. Memorandum NMFS-SEFSC-610.

Conn, P.B., Thorson, J.T., and Johnson, D.S. 2017. Confronting preferential sampling when analyzing population distributions: diagnosis and model-based triage. Methods in Ecology and Evolution 8, 1535-1546.

Dennis, D., Plaganyi, E., Van Putten, I., Hutton, T., and Pascoe, S. 2015. Cost benefit of fishery-independent surveys: Are they worth the money? Marine Policy 58, 108-115.

Ducharme-Barth, N.D., Shertzer, K.W., and Ahrens, R.N.M. 2018. Indices of abundance in the Gulf of Mexico reef fish complex: A comparative approach using spatial data from vessel monitoring systems. Fisheries Research 198, 1-13.

George, E.I., and McCollock, R.E. 1993. Variable selection via Gibbs sampling. Journal of the American Statistical Association 88, 881-889.

Gibson-Reinemer, D.K., Ickes, B.S., and Chick, J.H. 2017. Development and assessment of a new method for combining catch per unit effort dta from different fish sampling gears: multigear mean standardization (MGMS). Canadian Journal of Fisheries and Aquatic Sciences 74, 8-14.

Gwinn, D.C., Beesley, L.S., Close, P., Gawne, B., and Davies, P.M. 2016. Imperfect detection and the determination of environmental flows for fish: challenges, implications and solutions. Freshwater Biology 61, 172-180.

Hilborn R., and Walters, C. J. 1992. Quantitative Fisheries Stock Assessment. Chapman & Hall, New York.

Hooten, M. B., and Hobbs, N. T. 2015. A guide to Bayesian model selection for ecologists. Ecological Monographs 85, 3-28.

Ishwaran, H., and Rao, J.S. 2005. Spike and slab variable selection: frequentist and Bayesian strategies. The Annals of Statistics 33, 730-773.

Jiao, Y., Hayes, C., and Cortes, E. 2009. Hierarchical Bayesian approach for population dynamics modelling of fish complexes without species-specific data. ICES Journal of Marine Science 66, 367-377.

Kimura, D.K., and D.A. Somerton. 2006. Review of statistical aspects of survey sampling for marine fisheries. Reviews in Fisheries Science 14, 245-283.

Kotwicki, S., Ressler, P.H., Ianelli, J.N., Punt, A.E., and Horne, J.K. 2018. Combining data from bottom trawl and acoustic surveys to estimate an index of abundance for semipelagic species. Canadian Journal of Fisheries and Aquatic Sciences 75, 60-71.

Langseth, B.J., Schueller, A.M., Shertzer, K.W., Craig, J.K., and Smith, J.W. 2016. Management implications of temporally and spatially varying catchability for the Gulf of Mexico menhaden fishery. Fisheries Research 181, 186-197.

Maunder, M.N., 2001. A general framework for integrating the standardization of catch-per-unit-of-effort into stock assessment models. Canadian Journal of Fisheries and Aquatic Sciences 58, 795–803.

Maunder, M.N., and Piner, K.R. 2015. Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science 72, 7-18.

Maunder, M.N., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70, 141-159.

Maunder M. N., Sibert J. R., Fonteneau A., Hampton J., Kleiber P., and Harley S. 2006. Interpreting catch-per-unit-of-effort data to assess the status of individual stocks and communities, ICES Journal of Marine Science 63, 1373-1385.

McMahon, T.E., and Holanov, S.H. 1995. Foraging success of largemouth bass at different light intensities: implications for time and depth of feeding. Journal of Fish Biology 46, 759-767.

Monk J. 2013. How long should we ignore imperfect detection of species in the marine environment when modelling their distribution. Fish and Fisheries 15, 352–358.

Ntzoufras, I. 2009. Bayesian modeling using WinBUGS. Wiley, Hoboken, New Jersey.

O'Hara, R.B., and Sillanpaa, M.J. 2009. A review of Bayesian variable selection methods: what, how and which. Bayesian Analysis 4, 85-118.

Ono, K., Ianelli, J.N., McGilliard, C.R., and Punt, A.E. 2018. Integrating data from multiple surveys and accounting for spatio-temporal correlation in index the abundance of juvenile Pacific halibut in Alaska. ICES Journal of Marine Science 75, 572-584.

Parker, D., Winker, H., Bernard, A.T.F., Heyns-Veale, E.R., Langlois, T.J., Harvey, E.S., and Götz, A., 2016. Insights from baited video sampling of temperate reef fishes: how biased are angling surveys? Fisheries Research 179, 191–201.

Pennino, M.G., Paradinas, I., Illian, J.B., Muñoz, F., Bellido, J.M., López-Quílez, A., and Conesa, D. 2018. Accounting for preferential sampling in species distribution models. Ecology and Evolution 9, 653-663.

Plummer, M. 2003. "JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling." In *Proceedings of the 3^rd^ International Workshop on Distributed Statistical computing (DSC 2003). March*, pp. 20-22.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Reineking, B., and Schroder, B. 2006. Constrain to perform: regularization of habitat models. Ecological Modelling 193, 675-690.

De Robertis, A., Ryer, C.H., Veloza, A., and Brodeur, R.D. 2003. Differential effects of turbidity on prey consumption of piscivorous and planktivorous fish. Canadian Journal of Fisheries and Aquatic Sciences 60, 1517-1526.

Schobernd, Z.H., Bacheler, N.M., and Conn, P.B. 2014. Examining the utility of alternative video monitoring metrics for indexing reef fish abundance. Canadian Journal of Fisheries and Aquatic Science 71, 464-471.

SEDAR. 2018. SEDAR 55 – South Atlantic Vermilion Snapper Assessment Report. SEDAR, North Charleston SC. 170 pp. available online at: http://sedarweb.org/sedar-55.

Shertzer, K.W., Bacheler, N.M., Coggins, L.G., and Fieberg, J. 2016. Relating trap capture to abundance: a hierarchical state-space model applied to black sea bass (*Centropristis striata*) ICES Journal of Marine Science 73, 512–519.

Shmueli, G. 2010. To explain or to predict. Statistical Science 25, 289-310.

Smith, A.N.H., Anderson, M.J., and Millar, R.B. 2012. Incorporation the intraspecific occupancy-abundance relationship into zero-inflated models. Ecology 93, 2526-2532.

Streich, M.K., Ajemian, M.J., Wetz, J.J. and Stunz, G.W. 2018. Habitat-specific performance of vertical line gear in the western Gulf of Mexico: A comparison between artificial and natural habitats using a paired video approach. Fisheries Research 204, 16-25.

Su,Y. and Yajima, M., 2015. R2jags: Using R to Run 'JAGS'. R package version 0.5-6. http://CRAN.R-project.org/package=R2jags.

Tenan, S., O’Hara, R.B., Hendriks, I., and Tavecchia, G. 2014. Bayesian model selection: The steepest mountain to climb. Ecological Modeling 283, 62-69.

Thorson, J.T., and Ward, E.J. 2014. Accounting for vessel effects when standardizing catch rates from cooperative surveys. Fisheries Researech 155: 168-176.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society 58, 267-288.

Walters, C.J. 2003. Folly and fantasy in the analysis of spatial catch rate data. Canadian Journal of Fisheries and Aquatic Science 60, 1433-1436.

Walters, C.J. and Maguire, J. 1996. Lessons for stock assessment from the northern cod collapse. Reviews in Fish Biology and Fisheries 6, 125-137.

Walters, C.J. and Martell, S.J.D. 2004. Fisheries Ecology and Management. Princeton University Press, Princeton, New Jersey.

Ward, P. 2008. Empirical estimates of historical variations in the catchability and fishing power of pelagic longline fishing gear. Reviews in Fish Biology and Fisheries 18, 409-426.

Zhou, S., Hobday, A.J., Dichmont, C.M., and Smith, A.D.M. 2016. Ecological risk assessments for the effects of fishing: a comparison and validation of PSA and SAFE. Fisheries Research 183, 518-529.

Tables

Table 1

Covariate descriptions and definitions.

Variable	Abbreviation	Class	Definition
Latitude	*lat*	continuous	The latitude of the sample location.
Longitude	*lon*	continuous	The longitude of the sample location.
Depth	*dep*	continuous	A continuous variable indicating the water depth at the trap location.
Soak time	*E*	continuous	A continuous variable indicating the length of time the trap was set before retrieval.
Temperature	*temp*	continuous	The water bottom temperature at the trap locations during sampling.
Turbidity	*turb*	categorical	A dummy variable indicating the level of turbidity (1 = level 2, 0 = level 1).
Substrate	*sub*	continuous	The percent of the substrate visible with the camera that is hard-bottom.
Relief	*rel*	categorical	A dummy variable with value of 1 indicating that the relief was "high".
Current away	*dir1*	categorical	A dummy variable that is 1 when the current direction is flowing away from the camera lens.
Current side	*dir2*	categorical	A dummy variable that is 1 when the current direction is flowing perpendicular to the camera lens.

Table 2

Simulation structure and inputs. The equations represent the structure of the data-generating model and the Inputs are the parameter values used in the simulation.

Data-generating model	Description	Inputs
<u>Process model</u>		
$\log\ I_t\ \sim \text{Normal}\ \ \mu, \sigma$	Temporal abundance model	$\mu = -3, \sigma = 0.93$
$\log(n_{s,t}) = \log\ I_t\ \ + co\ {}^{n}_{s,t}$	Site-level abundance model	
$co\ {}^{n}_{s,t} = \theta_1 x1_{s,t} + \theta_2 x2_{s,t} + \theta_3 x3_{s,t}$	Spatial covariates	$\theta_1 = 1, \theta_2 = -0.5, \theta_3 = 0$
<u>Trap observation model</u>		
$c_{s,t}^{trap} \sim \text{Poisson}\left(e^{\log(n_{s,t}) + cov_{s,t}^{trap}}\right)$	Trap observation model	
$co\ {}^{trap}_{s,t} = \theta_4 x4_{s,t} + \theta_5 x5_{s,t} + \theta_6 x6_{s,t}$	Trap catchability covariates	$\theta_4 = 1, \theta_5 = -0.5, \theta_6 = 0$
<u>Camera observation model</u>		
$c_{s,t}^{cam} \sim \text{Poisson}\left(e^{\log(n_{s,t}) + cov_{s,t}^{cam}}\right)$	Camera observation model	
$co\ {}^{cam}_{s,t} = \theta_7 x7_{s,t} + \theta_8 x8_{s,t} + \theta_9 x9_{s,t}$	Camera catchability covariates	$\theta_7 = 1, \theta_8 = -0.5, \theta_9 = 0$

Table 3

Bayesian p-values for model fit evaluation. Each model includes or excludes site level random effects in the abundance (ε^{abun}), trap (ε^{trap}), and camera (ε^{cam}) sub-models. The Bayesian p-value is the output metric of a posterior-predictive check where a value of 0.5 indicates perfect fit of the model to the data and values approaching zero or one indicate under- or over-dispersion of the data relative to model predictions, respectively.

#	Random effects included in model	Model deviance	Bayesian p-value	
			Camera	Trap
1	ε^{abun}, ε^{trap}, ε^{cam}	31511.1	0.33	0.49
2	ε^{abun}, ε^{trap}	31487.2	0.11	0.67
3	ε^{abun}, ε^{cam}	31281.3	0.33	0.48
4	ε^{trap}, ε^{cam}	31741.2	0.99	0.71
5	ε^{abun}	120003.7	0.02	0.67
6	ε^{trap}	3081507.0	1.00	0.64
7	ε^{cam}	183413.1	1.00	1.00
8	None	3218795.0	1.00	1.00

Table 4

Covariate parameter posterior summaries. Posterior means and credible intervals are derived from posterior samples of the full model prior to model reduction. The grey text indicates covariates that are not statistically different than zero at $\alpha = 0.05$. Variable definitions are presented in Table 1. The column labeled 'Mean (SSVS)' is the mean of the posterior distribution with induced shrinkage via the Stochastic Search Variable Selection procedure (SSVS) and the column labeled 'I_j' is the parameter inclusion indicator variable.

Variable	Parameter	Mean	95% Credible intervals	Mean (SSVS)	I_j
Abundance					
lat	β_1	-1.38	-2.03, -0.71	-1.86	1.00
lat^2	β_2	-1.02	-1.27, -0.76	-1.00	1.00
lon	β_3	0.80	0.24, 1.35	1.24	1.00
lon^2	β_4	-1.03	-1.33, -0.74	-0.64	1.00
lat:lon	β_5	0.79	0.29, 1.32	0.09	**0.12**
dep	β_6	0.12	-0.04, 0.29	0.00	0.00
dep^2	β_7	-0.15	-0.19, -0.1	-0.14	1.00
rel	β_8	0.04	-0.33, 0.39	0.00	0.00
sub	β_9	0.61	0.50, 0.73	0.60	1.00
Trap					
E	η_1	4.65	2.86, 6.47	4.83	1.00
E^2	η_2	-2.92	-4.28, -1.6	-3.05	1.00
temp	η_3	0.81	0.71, 0.91	0.72	1.00
temp^2	η_4	0.11	0.09, 0.13	0.10	1.00
turb	η_5	0.04	-0.21, 0.28	0.00	0.00
*dir*1	η_6	0.63	0.32, 0.97	0.10	**0.23**
*dir*2	η_7	0.35	0.06, 0.64	0.00	0.00
Camera					
turb	φ_1	0.78	0.49, 1.03	0.79	1.00
*dir*1	φ_2	1.06	0.72, 1.43	0.59	1.00
*dir*2	φ_3	0.41	0.06, 0.76	0.00	0.00
temp	φ_4	0.28	0.13, 0.42	0.02	**0.11**
temp^2	φ_5	0.05	0.00, 0.11	0.00	0.00

Figures

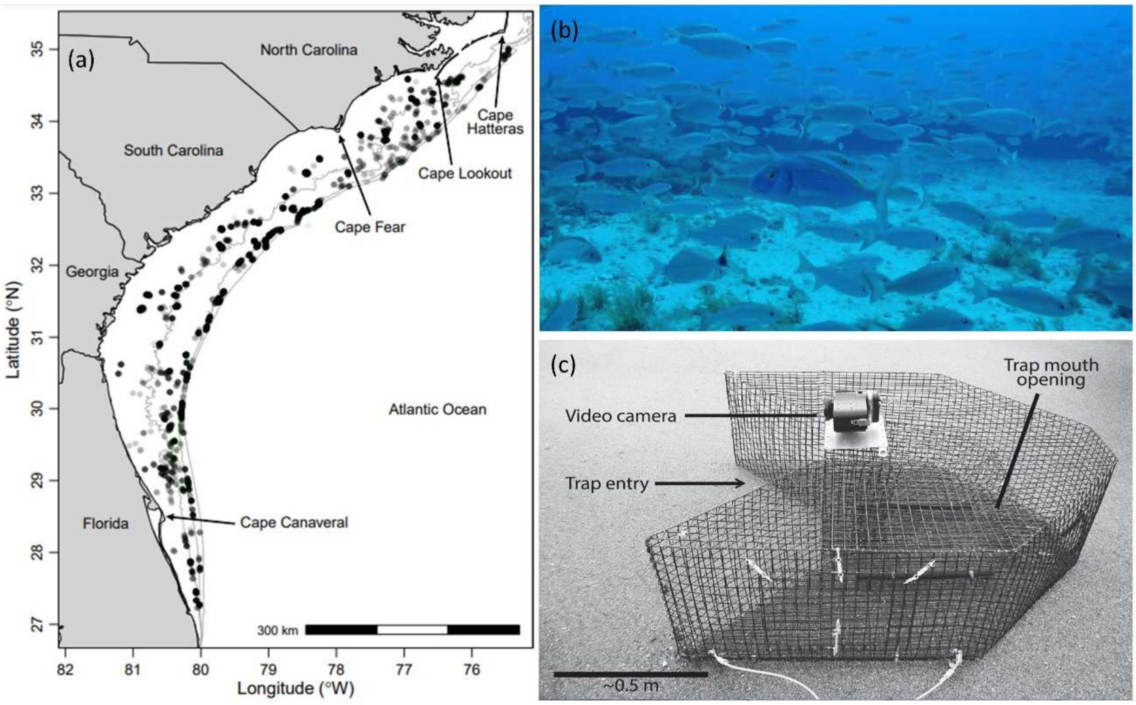

Fig. 1. Study area (a), sample video frame (b) and a Chevron fish trap outfitted with an outward-looking Canon high-definition video camera over the mouth of the trap (c). The points on panel (a) represent sample locations.

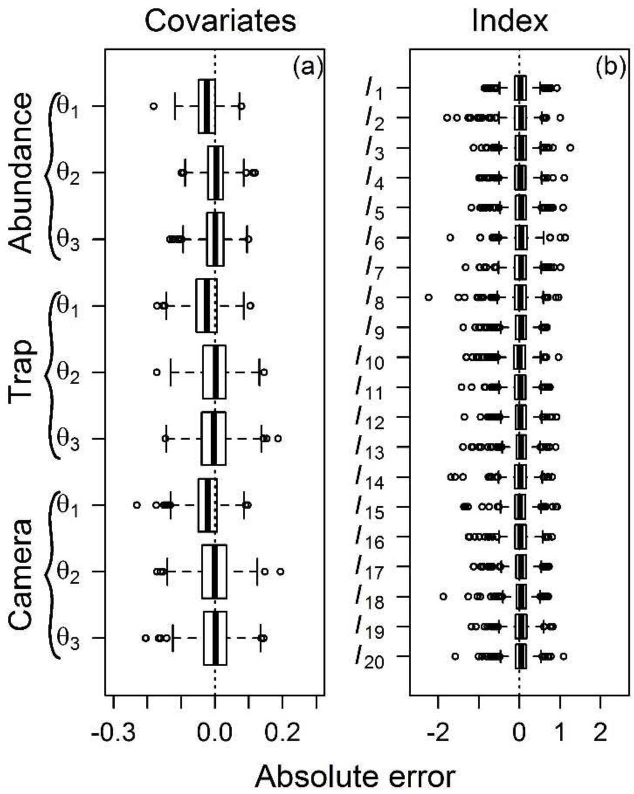

Fig. 2. The bias of simulated covariate effect parameters (a) and fishery index (b) posterior distributions. Posterior samples were derived by fitting the data-generating model to 200 simulated 20-year data sets.

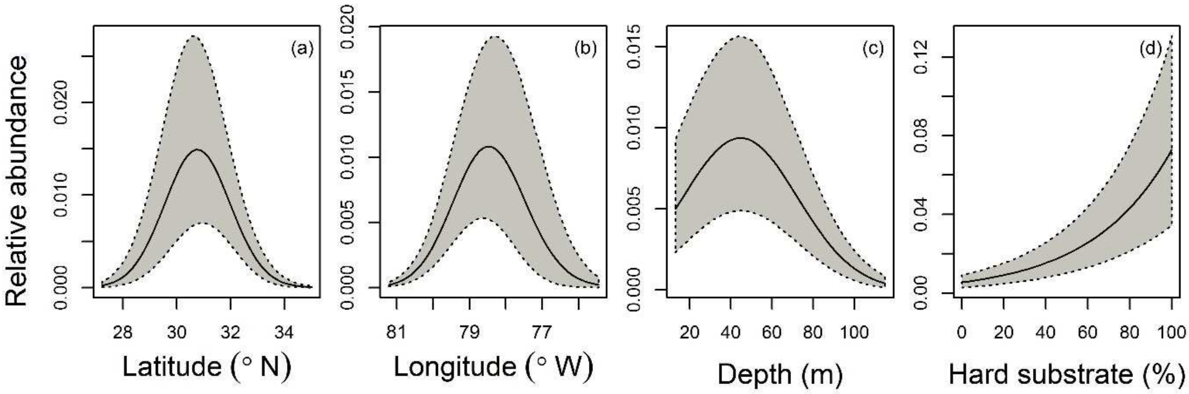

Fig. 3. The predicted response of sub-population abundance to spatial covariates. The grey region represents 95% Bayesian credible intervals.

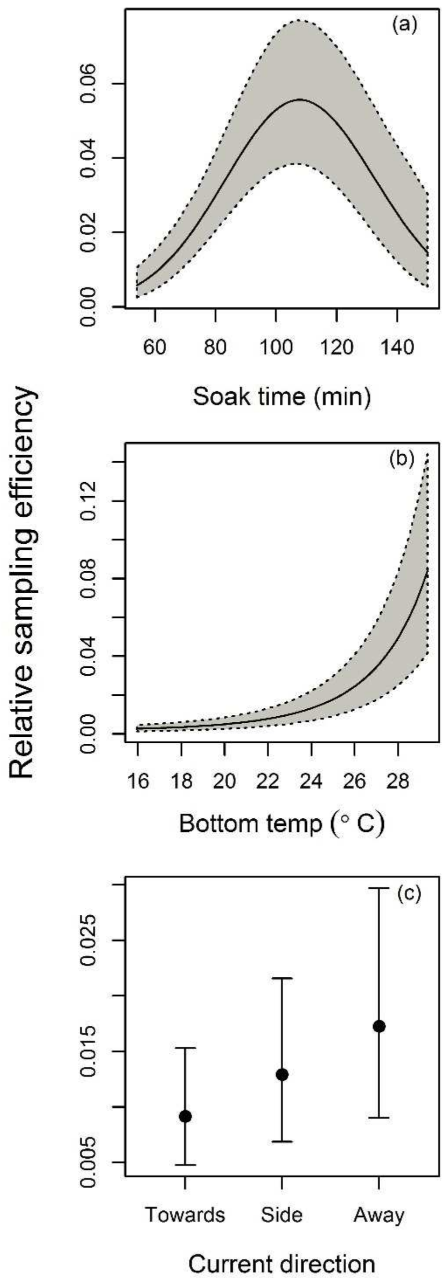

Fig. 4. The predicted response of chevron trap sampling efficiency to sampling and environmental covariates. The grey region represents 95% Bayesian credible intervals.

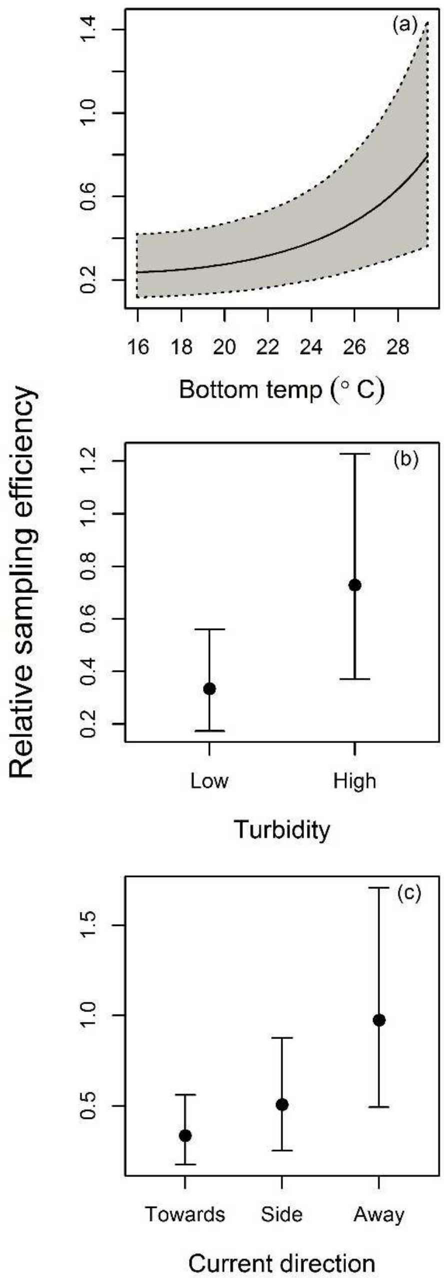

Fig. 5. The predicted response of video camera sampling efficiency to environmental covariates. The grey region represents 95% Bayesian credible intervals.

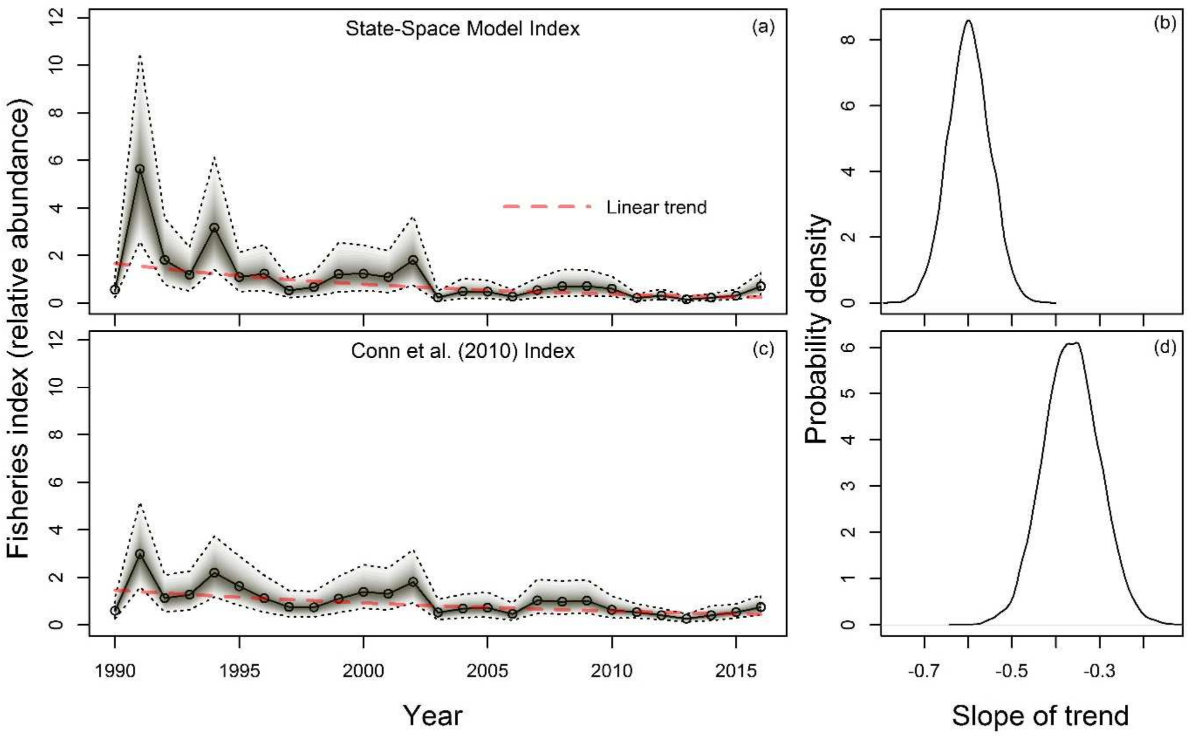

Fig. 6. The predicted annual relative abundance of the vermilion snapper meta-population using our State-Space Model (a) and the methods of Conn et al. (2010) (c). The grey region represents 95% Bayesian credible intervals. The dashed line represents the estimated linear trend. Panels (b) and (d) represent the probability distributions of the bootstrapped linear trend for each index.

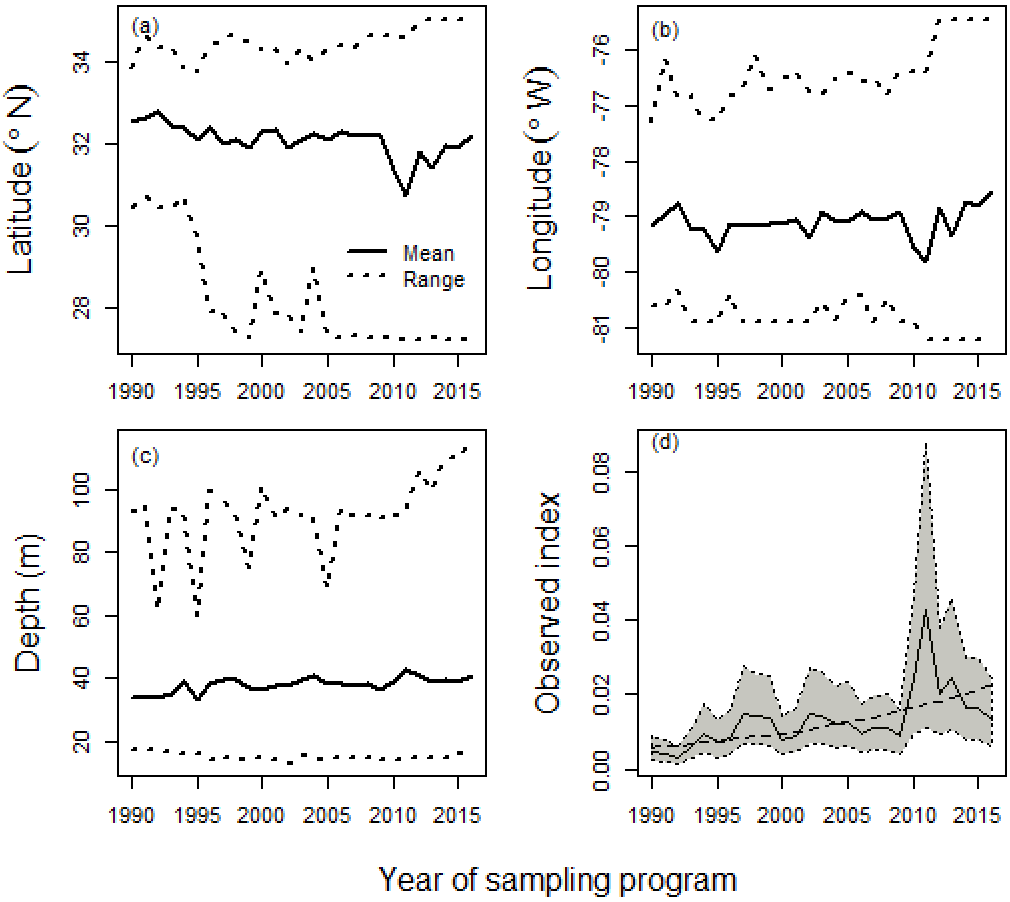

Fig. 7. Impact of changing sampling frame on the predicted relative abundance at the meta-population level. Panel (a), (b), and (c) represent the mean and range in latitude, longitude, and depth across samples collected each year. Panel (d) is the model predicted increases to the fishery index expected when the spatial changes on panel (a), (b), and (c) are not accounted for. The grey region on panel (d) represents 95% Bayesian credible intervals of the predictions and the dashed line is the log-linear trend of the predictions.