Cross-Shelf Habitat Occupancy Probabilities for Juvenile Groupers in the Florida Keys Coral Reef Ecosystem

W. J. Harford,* S. G. Smith, J. S. Ault, and E. A. Babcock
Rosenstiel School of Marine and Atmospheric Science, University of Miami,
4600 Rickenbacker Causeway, Miami, Florida 33149, USA

Abstract

In the Florida Keys coral reef ecosystem, delineation of reef fish distributions in relationship to habitat patterns is important for improving the design characteristics of fishery-independent surveys. Efficient survey design depends on analysis of fish distribution patterns to inform and improve the precision of future surveys. We used a diver visual survey to quantify occupancy patterns of preexploitation-size Black Grouper Mycteroperca bonaci and Red Grouper Epinephelus morio. The survey was based on a stratified random sampling design with strata reflecting cross-shelf coral reef habitat types. A multiple spatial scale modeling approach confirmed a cross-shelf occurrence gradient for Red Grouper, with higher nearer-to-shore occupancy probability and lower offshore occupancy probability. Black Grouper occurrence followed a latitudinal gradient, with higher occurrence probabilities in the lower Florida Keys than in the upper Keys. Local habitat characteristics measured within reef strata suggested that occupancy relationships for both species varied according to vertical relief. Our analysis also included multilevel slope coefficients (random effects), which revealed unforeseen variance structure in Black Grouper occurrence probability among cross-shelf reef strata. Our study improves on previous qualitative observations of juvenile grouper distributions in the Florida Keys and highlights the use of multilevel models in revealing variance structures of fish distributions not revealed by fixed-effects models. Our analysis contributes to a discussion about foraging characteristics in producing the observed distributional patterns, and we suggest that examining the links between the distributions of forage fishes and larger predators (i.e., groupers) would be a useful step in improving survey stratification schemes.

In the Florida Keys coral reef ecosystem, delineation of fish and shellfish distribution patterns is critically important for efficient fishery-independent monitoring of reef resources (Ault et al. 1999, 2005a, 2009; Lindeman et al. 2000). The iterative process of designing and implementing fish surveys can capitalize on analyses of fish distributional patterns as a means to inform and improve survey design characteristics (Ault et al. 2005a; Smith et al. 2011a, 2011b). A common approach to delineating distribution patterns is to use resource selection functions (RSFs; Manly et al. 2002). Although RSFs are used increasingly to address a wide range of ecological questions, their application to marine ecosystems remains relatively rare (Robinson et al. 2011). This apparent underutilization is unfortunate because RSFs hold great potential for addressing a range of ecological concerns, including gauging responses to reef degradation, informing marine reserve design, and improving survey stratification schemes (Meester et al. 2004; Grober-Dunsmore et al. 2006; Smith et al. 2011a).
In this study, we had two objectives aimed at evaluating RSFs for preexploitation-size Red Grouper Epinephelus morio (<50 cm) and Black Grouper Mycteroperca bonaci (<60 cm) within the Florida Keys reef tract. Our focus on Red and Black groupers was motivated by the fact that descriptions of habitat use have been mainly qualitative for preexploitation or juvenile life stages, whereas statistical analyses have more frequently focused on surveys of adult groupers (Alevizon et al. 1985; Sluka and Sullivan 1996; Sluka et al. 1996, 2001; Newman et al. 1997; Connell and Kingsford 1998; Lopez-Rivera and Sabat 2009). Our first objective was to determine whether grouper distribution patterns varied in relation to cross-shelf habitat characteristics at two spatial scales. In riverine ecosystems, nested spatial hierarchies can be used to describe fish distributions in relation to watershed-scale climate patterns, catchment-scale surface runoff and stream flow, and stream-scale substrate characteristics (Harig and Fausch 2002; Harford and McLaughlin 2007; Tisseuil et al. 2013). In coral reef ecosystems, fish distributions can be described in relation to cross-shelf categorization of reef structures and local habitat characteristics within reef strata (Done 1983; Sluka et al. 1994; Connell and Kingsford 1998; Grober-Dunsmore et al. 2008; Kendall et al. 2011; Pittman and Brown 2011; Yeager et al. 2011).

Our second objective was to determine whether multilevel models (also known as hierarchical, random-effects, or mixed models) could reveal variation in grouper distribution patterns that would probably remain undetected by simpler fixed-effects models (Kutner et al. 2005; Gelman and Hill 2007). Multilevel linear models allow slope coefficients to vary among cross-shelf reef strata, which could reveal disparate strata-level responses to habitat covariates. Terrestrial applications of multilevel models have been demonstrably beneficial in revealing how habitat use patterns can vary between geographic regions and times of year (Boyce et al. 2002; Gillies et al. 2006; McLoughlin et al. 2010). In revealing whether and how habitat relationships vary spatially, previously unforeseen variance structures can in turn inform stratification schemes that generate cost-effective allocations of sampling effort in future surveys (Xu et al. 2015).

METHODS

Coral reef surveys.—The Florida Keys coral reef extends 400 km southwest along an island archipelago from Key Biscayne near Miami to the Dry Tortugas region 113 km west of Key West (Figure 1). Unique topographic and oceanographic conditions help sustain the highly productive ecosystem (Ault et al. 2005a). The coral reef tract consists of a series of parallel low ridges and connected valleys that are situated parallel to the Florida current and Florida Bay (Hoffmeister 1974). This coastal marine ecosystem consists of estuaries, lagoons, mangrove stands, coral islands, seagrass beds, and coral reefs. The cross-shelf formations of coral reefs consist of inner-shelf patch reefs that form discontinuous linear clusters or irregularly scattered clusters and outer-shelf fore reefs that occur along the edge of the reef tract (Hoffmeister 1974; Shinn et al. 1977; Lidz et al. 2006; Smith et al. 2011b). Within reef formations, habitats are varied, with dramatic changes in topographic relief, substrate type, coral density, flow patterns, and wave action (Hoffmeister 1974; Geister 1977; Shinn et al. 1977; Smith et al. 2011b).

Since 1979, a multispecies fisheries-independent reef fish survey has been conducted in the Florida Keys coral reef ecosystem (Ault et al. 1998, 2005b; Smith et al. 2011a). We used scuba diver visual observations from these surveys that were conducted in depths <18 m during 2003 and annually between 2005 and 2011 (n = 2,225; Smith et al. 2011a). During this time period, divers began collecting detailed habitat measurements within observation plots. Surveys were conducted between May and September in the 885-km² domain of the Florida Keys coral reef tract using a two-stage stratified random sampling design that employed a spatial hierarchy of cross-shelf and site-scale habitat characteristics (Ault et al. 1998; Smith et al. 2011a). Annual allocation of sampling effort among strata was determined using a Neyman allocation to achieve estimates of a desired precision for multispecies reef fish density, abundance at length, and community composition (Smith et al. 2011a).

The two-stage survey design employed primary sampling units (PSUs; 200-m × 200-m mapped grids) and second-stage sampling units (SSUs; 15-m-diameter plots of 177 m²). Within chosen PSUs, two randomly selected SSUs were visited by divers; thus, the sampling design consisted of replicate visits to PSUs in which the observation plots (SSUs) were unique and covered only a small spatial extent of each PSU. At each PSU, closely spaced pairs of scuba divers conducted a standardized observation process that involved listing all observed fish species during 5-min sampling periods before recording abundance and fork length. Cross-shelf reef strata were classified according to reef type, rugosity, and cross-shelf position (Table 1). Because the observed abundances of Red and Black groupers were relatively low at each SSU, we opted to recode the observed counts as binary occurrence indicators (positive = 1 and none-observed = 0). Variables considered as plausible correlates of grouper occurrence (hereafter, habitat variables) were bottom depth (m), maximum vertical relief (m), percent coral cover, percent hard bottom, and latitude. The first four variables were measured by divers during each visit to an SSU. All habitat variables were averaged to produce PSU-level observations and were subsequently standardized by subtracting the mean and dividing by the standard deviation. Variable standardization produced scale-independent variables, which helped to facilitate convergence of the Markov chain–Monte Carlo algorithm and allowed the magnitudes of the slope coefficients to be compared directly.
FIGURE 1. Map of the Florida Keys coral reef tract. Reef strata are described in Table 1.

TABLE 1. Physical and biological habitat characteristics of primary sampling units (PSUs) during 2003–2011. Abbreviations are as follows: \( n \) = the number of PSUs sampled, \( \text{min} \) = the minimum value, \( \text{max} \) = the maximum value, and \( \text{avg} \) = the mean value.

<table>
<thead>
<tr>
<th>Habitat stratum type</th>
<th>Rugosity</th>
<th>( n )</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
<th>Min</th>
<th>Avg</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Inshore patch reef</td>
<td>Low–med</td>
<td>99</td>
<td>1.8</td>
<td>3.5</td>
<td>7.8</td>
<td>0.10</td>
<td>1.16</td>
<td>3.12</td>
<td>0</td>
<td>13.8</td>
<td>45.6</td>
<td>0</td>
<td>17.5</td>
<td>75.6</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ii. Midchannel patch reef</td>
<td>Low–med</td>
<td>390</td>
<td>1.5</td>
<td>5.3</td>
<td>12.6</td>
<td>0.12</td>
<td>1.37</td>
<td>4.27</td>
<td>0</td>
<td>10.7</td>
<td>52.2</td>
<td>0</td>
<td>10.7</td>
<td>77.1</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iii. Offshore patch reef</td>
<td>Low–med</td>
<td>228</td>
<td>2.1</td>
<td>7.0</td>
<td>13.5</td>
<td>0.10</td>
<td>1.08</td>
<td>4.00</td>
<td>0.2</td>
<td>6.2</td>
<td>32.2</td>
<td>0.2</td>
<td>24.4</td>
<td>24.9</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iv. Offshore patch reef</td>
<td>High</td>
<td>45</td>
<td>2.2</td>
<td>6.2</td>
<td>12.4</td>
<td>0.90</td>
<td>2.41</td>
<td>3.97</td>
<td>2.2</td>
<td>9.0</td>
<td>26.0</td>
<td>2.2</td>
<td>25.1</td>
<td>25.1</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v. Outer fore reef</td>
<td>Low–med</td>
<td>1,184</td>
<td>2.1</td>
<td>9.0</td>
<td>18.0</td>
<td>0.10</td>
<td>0.79</td>
<td>3.20</td>
<td>0</td>
<td>3.3</td>
<td>25.0</td>
<td>0</td>
<td>24.4</td>
<td>24.9</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vi. Outer fore reef</td>
<td>High</td>
<td>279</td>
<td>2.8</td>
<td>6.6</td>
<td>17.1</td>
<td>0.32</td>
<td>2.19</td>
<td>4.42</td>
<td>0.6</td>
<td>5.3</td>
<td>36.8</td>
<td>0.6</td>
<td>24.4</td>
<td>24.7</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Resource selection functions.**—We evaluated logistic regression and more complex zero-inflated binomial models in our preliminary development of RSFs using occurrence data. Logistic occurrence probability was modeled using the binomial density function

\[ Z_i \sim \text{binomial}(J_i, \psi_i), \tag{1} \]

where \( Z_i \), the number of positive detections at \( PSU_i \), is a function of the occurrence probability, \( \psi_i \), and the number of SSUs visited at \( PSU_i \), \( J_i \). Hall (2000) introduced the zero-inflated binomial model for bounded count data to account for excess zeros, which was later extended to site occupancy modeling of animal distributions (MacKenzie et al. 2002; Tyre et al. 2003). Site occupancy models consist of several approaches similar to logistic regression but in which two binomial processes are jointly estimated (MacKenzie et al. 2002; Royle and Dorazio 2009):

\[ Z_i \sim \begin{cases} 0, & \text{with probability } 1 - \psi_i, \\ \text{binomial}(J_i, p_i), & \text{with probability } \psi_i. \end{cases} \tag{2} \]

The number of positive detections, \( Z_i \), is a function of the probability of occurrence, \( \psi_i \), and the conditional probability of observation, \( p_i \), that arises from the \( J_i \) samples (visits) in the second binomial process. The probability of a zero observation at a particular PSU can be thought of as the sum two possible outcomes: (1) no individuals were present (which has the probability \( 1 - \psi_i \)) or (2) one or more individuals were present but went unobserved (which has the probability \( \psi_i(1 - p_i)^{J_i} \)).

A logit-linear occurrence submodel \( \psi_i' \) was used to describe occurrence as a function of habitat variables \( (X_{ik})_j \):

\[ \psi_i' = \mu + \alpha_{d[i]} + \sum_{k=1}^{3} \beta_{k,s[i]} X_{ik}, \tag{3} \]

with probability of occurrence \( (\psi_i) \) written as

\[ \psi_i = \exp(\psi_i')/(1 + \exp(\psi_i')). \tag{4} \]

In equation (3), the notation \( s[i] \) refers to the cross-shelf reef stratum containing \( PSU_i \), \( \mu \) is the grand mean, and \( \alpha_d \) refers to the categorical reef strata coefficients. Slope coefficients \( \beta_{k,s} \) were modeled as random effects to enable responses to vary among reef strata:

\[ \beta_{k,s[i]} = \mu_{p[k]} + \delta_{k,s[i]}, \tag{5} \]

where \( \mu_{p[k]} \) is the mean slope for coefficient \( k \) and \( \delta_{k,s[i]} \) are random coefficients describing stratum-level deviations with variances \( \sigma_{\delta k}^2 \).

While our primary interest was in occupancy probabilities, we also wanted to explore how the observation process \( p \), expressed as a function of the habitat variables, could affect model fit. Thus, \( p \) was expressed as a logit-linear observation submodel of PSU-level habitat variables:

\[ p_i' = v + \sum_{k=1}^{3} \omega_k X_{ik}, \tag{6} \]

with the mean intercept \( v \) expressed as a probability \( (p_i) \):

\[ p_i = \exp(p_i')/(1 + \exp(p_i')). \tag{7} \]

The coefficients \( \omega_k \) corresponded to maximum vertical relief, depth, and coral cover.

The Bayesian approach was used to fit logistic regression and site occupancy models in OpenBUGS (Lunn et al. 2009; Kéry and Schaub 2012). In RSF development, some model variants were considered that were not well supported by the data and were excluded from our final model formulations. These models included ones with quadratic terms \( (X_{ik}^2) \) to screen for curvilinear responses to habitat variables and a categorical year coefficient to account for interannual variation in occurrence probability. For the logistic regression and site occupancy model formulations, a Gibbs sampler–based variable selection technique was used to identify and select habitat variables with the highest posterior probabilities (Kuo and Mallick 1998; Congdon 2003; Ntzoufras 2009). To implement this Bayesian variable selection technique, the logit-linear responses were modified such that coefficients were multiplied by a binary indicator parameter. When the indicator parameters took on a value of 1, their associated coefficients were included in the model; when they took on a value of 0, however, the associated coefficients were excluded. The indicator parameters had Bernoulli priors with probabilities of 0.5 to give each variable an equal prior probability of inclusion. Posterior means of binary indicators determined the inclusion probabilities of their associated parameters, with those having probabilities >0.50 being retained in each RSF (Ntzoufras 2009).

Logit-scale intercept and categorical reef stratum coefficients were assigned diffuse normal priors with means zero and variances 2.7 (Lunn et al. 2012). Mean slope coefficients were assigned diffuse normal priors, and among-strata slope variances, \( \sigma_{s[j]}^2 \), were assigned priors of

\[ \sqrt{\sigma_{s[j]}^2} \sim \text{uniform}(0, 5). \]

We also conducted a sensitivity analysis for the choice of priors for the logit-scale parameters using a diffuse \( t \)-distribution (Dorazio et al. 2011). After we discarded an initial 150,000 iterations, the Markov chain–Monte Carlo algorithm converged (based on Geweke and Gelman-Rubin criteria) for all models (Geweke 1992; Congdon 2003; Gelman et al. 2004). Approximation of the posterior distribution was obtained from a subsequent 150,000 samples from two parallel chains. Model adequacy was assessed by calculating squared Pearson residuals to...
compare the lack of model fit to the data against the lack of fit that would be expected from replicated data sets produced using the model’s assumptions and estimated parameters (Brooks et al. 2000; Gelman and Hill 2007; Kéry 2010). The Bayesian $P$-value was calculated as the proportion of times that the replicated residuals were greater than the observed residuals, with values near 0.5 indicating a good fit (Gelman et al. 2004; Ntzoufras 2009).

**RESULTS**

**Coral Reef Surveys**

Diver samples at PSUs varied between 213 and 457 during 2003–2011. The PSUs sampled were located between latitudes 24.431°N and 25.749°N, which corresponds approximately to the coral reef habitats occurring between Key West and Miami, Florida (Table 1). Across all PSUs, bottom depths ranged between 1.5 and 18.0 m and maximum vertical relief ranged between 0.10 and 4.42 m. The prevalence of hard-bottom habitats varied from 0% to 100%, but the percentage of coral cover never exceeded 52.2%. There was some evidence of correlation between the standardized variables used in the site occupancy models (Pearson’s $|r| < 0.38$ [all pairwise comparisons]). Average abundance at the scale of the SSU ranged between 0.0 and 8.0 individuals per sample for Black Grouper and between 0.0 and 2.5 individuals per sample for Red Grouper. Thus, there was relatively little information loss concerning habitat use patterns in recoding relative abundance as binary observations.

**Red Grouper Resource Selection Functions**

For Red Grouper occurrence, the logistic regression formulation (equation 1) had a poor fit to the data (Bayesian $P$-value of 0.99) and thus was considered to have low support as a plausible descriptor (Figure 2). The site occupancy formulation (equation 2) fit appreciably better, with a Bayesian $P$-value of 0.61. In each model formulation, inclusion probabilities supported reef stratum–level intercepts, which are interpreted as conditional categorical responses ($\alpha_i$; Table 2). We calculated odds ratios for the intercept coefficients using reef stratum i (low–medium relief inshore patch reefs) as a reference category (Figure 3). Odds ratios greater than 1 indicated better odds of Red Grouper occurrence relative to the reference category, and odds ratios less than 1 indicated worse odds of occurrence. While the stratum–level intercepts were mostly nonsignificant, with the exception of high-rugosity offshore fore reefs (habitat stratum vi), a cross-shelf decline in occurrence odds from inshore to offshore was evident (Figure 3).

Posterior inclusion probabilities also supported a negative occupancy relationship with maximum vertical relief in both the simple logistic and site occupancy model formulations (Table 2). Both formulations suggested negative relationships with vertical relief, but the site occupancy formulation supported vertical relief–induced heterogeneity in observation probability rather than in occupancy probability. The predicted responses from zero-inflated submodel components $p$ and $\psi$ were plotted against maximum vertical relief (Figure 4). The trends illustrate how the predicted source ($p$ or $\psi$) of a negative response to vertical relief changed based on whether heterogeneity in the observation process
was included in model structure. In addition, there was weak support in the inclusion probabilities for differences in occurrence responses to latitude among habitat strata, but the latitude slope coefficients were nonsignificant and no systematic trends were apparent. We examined whether these results (and those for Black Grouper below) were sensitive to the form of the diffuse logit-scale priors used in the analysis. Using an alternative $t$-distribution prior, we found posterior distributions of model parameters to be quite similar between priors (Tables 2, 3).

**Black Grouper Resource Selection Functions**

The fit of the logistic regression model to the data was poor, with a Bayesian $P$-value of 1.0. The site occupancy

---

**TABLE 2. Variable inclusion probabilities and coefficient means for Red Grouper statistical models. Asterisks indicate variables included in resource selection functions; values in parentheses are standard errors; na = not applicable.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Logistic regression</th>
<th>Site occupancy</th>
<th>Base diffuse priors</th>
<th>$t$-distribution diffuse priors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occurrence submodel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean slope ($\mu_j$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>0.02</td>
<td>0.13</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Vertical relief</td>
<td>0.91*</td>
<td>0.27</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>0.01</td>
<td>0.03</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Latitude</td>
<td>0.01</td>
<td>0.03</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>% Hard bottom</td>
<td>0.00</td>
<td>0.02</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Random slopes ($\delta_j$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>0.14</td>
<td>0.16</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Vertical relief</td>
<td>0.16</td>
<td>0.27</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>0.11</td>
<td>0.44</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Latitude</td>
<td>0.61</td>
<td>0.57</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>% Hard bottom</td>
<td>0.02</td>
<td>0.07</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>Reef stratum ($\alpha_s$)</td>
<td>1.00*</td>
<td>1.00*</td>
<td>0.71 (0.75)</td>
<td>0.66 (0.75)</td>
</tr>
<tr>
<td>Strata i to vi</td>
<td></td>
<td></td>
<td>0.92 (0.68)</td>
<td>0.80 (0.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.86 (0.70)</td>
<td>0.75 (0.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.35 (1.05)</td>
<td>0.41 (1.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>−0.49 (0.65)</td>
<td>−0.58 (0.65)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>−2.31 (0.72)</td>
<td>−2.44 (0.75)</td>
</tr>
<tr>
<td>Grand mean ($\mu$)</td>
<td>na</td>
<td>na</td>
<td>−0.06 (0.65)</td>
<td>0.04 (0.66)</td>
</tr>
<tr>
<td><strong>Conditional submodel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean slope ($\omega_j$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>na</td>
<td>0.13</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Vertical relief</td>
<td>na</td>
<td>0.78*</td>
<td>−0.52 (0.08)</td>
<td>−0.53 (0.07)</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>na</td>
<td>0.01</td>
<td>na</td>
<td></td>
</tr>
<tr>
<td>Grand mean ($\nu$)</td>
<td>na</td>
<td>na</td>
<td>−0.88 (0.11)</td>
<td>−0.89 (0.11)</td>
</tr>
</tbody>
</table>

---

**FIGURE 3.** Intercept odds ratios of Red Grouper occurrence in each reef stratum. Odds ratios were calculated relative to reference stratum i, defined as low–medium relief inshore patch reefs. The circles denote means, and the lines are centered 95% confidence intervals.
model fit appreciably better, with a Bayesian $P$-value of 0.67. Posterior inclusion of maximum vertical relief and latitude were supported by logistic regression (Table 3). As with Red Grouper, when the observation process ($p$) was allowed to vary in relation to habitat variables, the inclusion probabilities supported vertical relief–induced heterogeneity in $p$ and vertical relief no longer corresponded to occurrence probability (Table 3). However, unlike in the case of Red Grouper, $p$ varied positively with maximum vertical relief. The inclusion probabilities did not support reef stratum intercepts, but there was support for reef stratum–level differences in the response of occurrence probability to depth (Table 3).

For the logistic regression and site occupancy models for Black Grouper, predicted occurrence probability was plotted against latitude with maximum vertical relief fixed at its observed mean, and vice versa (Figure 5A, B). In each formulation, a negative occurrence probability response to latitudinal change was predicted. The analysis also revealed support for cross-shelf differences in the response of occurrence probability to depth (Table 3; Figure 6). The slope coefficients for depth in the occurrence submodel varied among reef strata and were consistently negative for the strata nearest to shore and positive for the outer-shelf fore reef strata (Figure 6).

DISCUSSION

The fishery-independent diver visual survey (Smith et al. 2011a) was designed from principles of probability-based statistical sampling, thus providing a sound foundation for occurrence modeling (Cochran 1977; Hayek and Buzas 1996; MacKenzie et al. 2006). The large spatial extent of the Florida Keys sampling domain and the innovations in survey stratification in relation to relevant habitat and environmental features enabled investigation at appropriate ecological scales (Wiens 1989; Ault et al. 2005b, 2013; Johnson et al. 2013). This survey has been tailored to the objectives of multispecies stock assessment, evaluation of the effectiveness of no-take marine reserves, and measurement of ecosystem and reef fish community condition (Smith et al. 2011a; Ault et al. 2014). Its design reflects a trade-off in sampling effort allocation between repeating site visits (to improve understanding of observation processes) and visiting many unique sites (to improve occurrence and abundance estimation) (Smith et al. 2011a). This trade-off can be complicated (for more on survey design for occupancy and detection estimation, see Field et al. 2005; MacKenzie and Royle 2005; Guillera-Arroita et al. 2010; Wintle et al. 2012; Monk 2014). In our analysis, simple logistic regression produced poor fits to the data because the number of zeroes in the data set exceeded the binomial expectation. The site occupancy models produced improved fits, so we expected that this formulation would be more reliable in evaluating trends in the probability of occupancy.

Applying the site occupancy model to the survey data did, however, introduce additional complexity in interpreting the results of the observation submodel. It may seem intuitive to interpret the observation submodel as providing detection probabilities (as is commonly done in site occupancy modeling), but the particular characteristics of grouper ecology and the fishery-independent survey make this conclusion...
problematical for several reasons. Excess zeroes could have occurred if individual fish inhabiting a PSU moved into or out of the SSU, lowering the probability of their being observed by a diver. While many reef-associated fishes tend to exhibit low mobility, larger grouper species maintain home ranges considerably larger than the 200-m × 200-m PSU (Chapman and Kramer 2000; Farmer and Ault 2011). The areal extent of an SSU in relation to that of a PSU also suggests that fine-scale spatial environmental heterogeneity could influence the propensity for not being observed by divers. A separate (but not mutually exclusive) possibility is that individual fish that are present within an SSU when it is observed by a diver are not actually detected. Thus, it is possible that in the observation submodel $p$ is a confounding of the probability of movement into/out of an SSU, fine-scale differences in fish distribution between SSUs, and detection probability given presence in the SSU.

Accounting for observational processes in marine ecosystems remains a challenging issue in occupancy and abundance estimation. Although the surveys that we analyzed did entail repeated visits to PSUs, the observation plots themselves were not repeatedly visited and covered only small portions of the much larger PSUs. Designs that achieve replication by conducting surveys in different observation plots within a larger sampling unit have been used for occupancy modeling (MacKenzie and Royle 2005; MacKenzie et al. 2006), but these designs may exaggerate abundance-induced heterogeneity in the observation process (Figures 4, 5). Abundance-induced observation heterogeneity arises in many sampling situations, as higher local (site) abundance is expected to yield more net detections (Royle and Dorazio 2009). With respect to detection probability, several previous studies indicate that it can vary with body size, schooling behavior, cryptic nature, distance from divers, and survey method (Byerly and Bechtol 2005; MacNeil et al. 2008a, 2008b; Bozec et al. 2011; Dickens et al. 2011). Clearly, accounting for observational processes, including detection probability, is an

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variable inclusion probability</th>
<th>Site occupancy</th>
<th>Site occupancy coefficient means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Logistic regression</td>
<td>Site occupancy</td>
<td>Base diffuse priors</td>
</tr>
<tr>
<td><strong>Occurrence submodel</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean slope ($\mu_j$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>0.02</td>
<td>0.09</td>
<td>na</td>
</tr>
<tr>
<td>Vertical relief</td>
<td>0.90*</td>
<td>0.04</td>
<td>na</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>0.29</td>
<td>0.04</td>
<td>na</td>
</tr>
<tr>
<td>Latitude</td>
<td>0.98*</td>
<td>0.80*</td>
<td>-1.10 (0.15)</td>
</tr>
<tr>
<td>% Hard bottom</td>
<td>0.01</td>
<td>0.10</td>
<td>na</td>
</tr>
<tr>
<td>Random slopes ($\delta_j$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>0.99*</td>
<td>1.00*</td>
<td>-1.02 (0.66)</td>
</tr>
<tr>
<td>Strata i to vi</td>
<td></td>
<td></td>
<td>-1.21 (0.32)</td>
</tr>
<tr>
<td>Vertical relief</td>
<td>0.36</td>
<td>0.11</td>
<td>na</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>0.05</td>
<td>0.05</td>
<td>na</td>
</tr>
<tr>
<td>Latitude</td>
<td>0.05</td>
<td>0.39</td>
<td>na</td>
</tr>
<tr>
<td>% Hard bottom</td>
<td>0.03</td>
<td>0.18</td>
<td>na</td>
</tr>
<tr>
<td>Reef stratum ($\alpha_s$)</td>
<td>0.31</td>
<td>0.05</td>
<td>na</td>
</tr>
<tr>
<td>Grand mean ($\mu$)</td>
<td>na</td>
<td>na</td>
<td>0.17 (0.19)</td>
</tr>
<tr>
<td><strong>Observation submodel</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean slope ($\omega_j$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth</td>
<td>na</td>
<td>0.06</td>
<td>na</td>
</tr>
<tr>
<td>Vertical relief</td>
<td>na</td>
<td>1.00*</td>
<td>0.61 (0.06)</td>
</tr>
<tr>
<td>% Coral cover</td>
<td>na</td>
<td>0.04</td>
<td>na</td>
</tr>
<tr>
<td>Grand mean ($v$)</td>
<td>na</td>
<td>na</td>
<td>-1.23 (0.09)</td>
</tr>
</tbody>
</table>
Both Red and Black grouper demonstrated occurrence responses that varied with regional topographic features of the Florida Keys reef tract. The occurrence of preexploitation-size Red Grouper appears to vary systematically with the proximity of reef habitat to the coastline. Occurrence probability was predicted to be higher in

**FIGURE 5.** Black Grouper responses to (A) latitude and (B) maximum vertical relief from the occurrence submodel ($\psi$) and (C) vertical relief from observation submodel ($\rho$). The solid (1) and dashed (2) lines correspond to the logistic regression and site occupancy formulations, respectively.

**FIGURE 6.** Conditional responses in Black Grouper occurrence probability ($\psi$) to depth within each reef stratum for the logistic regression and site occupancy model formulations. Reef strata (i–vi) are described in Table 1. Only reef strata with statistically significant slope coefficients are shown.

important step in statistical occupancy and abundance estimation, but one that is currently not receiving sufficient attention in the analysis of marine ecosystems (Monk 2014).
nearthshore patch reef habitats than in offshore fore reef habitats. The strength of this pattern was somewhat inconsistent, but inshore-to-offshore occurrence patterns were observed in all model formulations. Positive occurrence nearer to shore confirms previous observations that preexploitation-stage Red Grouper inhabit inshore waters of the Florida Keys (Moe 1969; Sluka et al. 1994). The probability of occurrence of preexploitation-phase Black Grouper tended to decrease from Key West (24.60°N) to Key Largo (25.07°N) and toward Miami. This southwest-to-northeast decline is intriguing in the context of the related distribution of Black Grouper spawning abundance. Although latitude (a north–south gradient) was used in our analysis, spawning abundance is highly correlated with longitude (Pearson’s rho = 0.93) given the geographic orientation of the Florida Keys reef tract (Figure 1). Using the same intensive diver surveys analyzed in this study, Ault et al. (2013) reported on the recovery of exploited reef fishes in marine reserves of the Dry Tortugas. While the Dry Tortugas region contains about 22% of the broader Florida Keys–Dry Tortugas coral reef habitat, it accounts for over 50% of Black Grouper spawning abundance (Ault et al. 2013). Ault et al. (2013) raise the question whether the Dry Tortugas region functions as an important recruitment source. This possibility is also supported by the directionality of the regional currents that may transport larval ﬁsh across the Florida Keys (Lee et al. 1994; Domeier 2004).

Red and Black grouper also exhibited occurrence responses that varied with the local habitat characteristics of the PSUs. Both species responded to vertical relief, but in opposite ways (Figures 4, 5). We note that the ﬁnal site occupancy model formulations included vertical relief in the observation submodel rather than the occupancy submodel. However, the directionality of occurrence in relation to vertical relief is more consistent with abundance-induced observation heterogeneity (occurrence patterns) than with observation complexities (detection heterogeneity). For instance, the declining detection of Red Grouper with increasing vertical relief could reﬂect the difﬁculty of detecting cryptic ﬁshes in highly rugose habitats, but this is inconsistent with the gregarious and possibly territorial responses of Red Grouper reported by approaching divers. Instead, the declining occupancy response with respect to vertical relief probably reﬂects the species’ use of benthic substrates for foraging and shelter. Like other members of the genus *Epinephelus*, Red Grouper probably utilize crevices in hard substrates for shelter and for ambushing benthic prey (Smith 1961; Cailliet et al. 1986; Parrish 1987; Brulé and Rodriguez Canché 1993; López-Rocha and Arreguin-Sanchez 2008; Coleman et al. 2010). Conversely, Black Grouper tend to forage above the bottom, are slender and have a tapering body form, and appear to be more agile swimmers (Parrish 1987). Like other mycteroperid groupers, they maintain a ﬁsh-dominated diet that includes fast-swimming and pelagic prey species (Randall 1967; Parrish 1987; Brulé et al. 2005). The positive association between Black Grouper and vertical relief could be indicative of the diversity of this species’ forage base and the larger habitat volume created for forage ﬁsh by vertical structures. In support of this possibility, observational studies of the Florida Keys reef tract report the occurrence of Black Grouper near a diversity of reef formation types (Sullivan and Sluka 1996).

Our analysis of multilevel slope coefﬁcients suggests that the occurrence response of Black Grouper to depth differed among reef strata. Within inshore and midchannel patch reefs, Black Grouper occurred with higher probability in the shallowest of available depths, while occurrence at outer fore reefs was highest in the deepest of available habitats (Figure 6). This result could mean that Black Grouper are not cueing on depth at all but that the variation in occurrence associated with depth reﬂects the inﬂuence of more ecologically meaningful variables. For instance, the occurrence of Black Grouper may reﬂect their forage base, which could be correlated with depth in some of the reef strata. Black Grouper have a diverse forage base, including pelagic and demersal species (e.g., Pomacentridae, Carangidae, Scaridae, and Labridae) that occur in both nearshore habitats and deeper-water fore reef habitats (Williams 1991; Overholtzer and Motta 1999; Brulé et al. 2005; Aguilar-Perera and Appeldoorn 2008; Collins and McBride 2011). This highlights a potentially unexplored aspect of Black Grouper ecology and the usefulness of multilevel models for identifying unforeseen variance structure in ﬁsh distribution patterns. More broadly, the occurrence patterns of both species suggest foraging characteristics that reﬂect a rather incomplete aspect of our analysis. Given our interest in improving survey design, it seems prudent that subsequent analyses move beyond simply linking ﬁsh distributions to abiotic surrogates of the underlying processes that result in occupancy. Rather, examining the spatial variance structures of predatory species (groupers) should move toward analysis of forage base habitat use and subsequently link forage base distributions to predator distributions. Such an approach could be inﬂuential in improving the precision of surveys that are essential for resource management (Smith et al. 2011a; Xu et al. 2015).

**ACKNOWLEDGMENTS**

We thank two anonymous reviewers for comments that led to the improvement of this manuscript. Financial support for this research was provided by the Cooperative Institute for Marine and Atmospheric Studies of the University of Miami (NOAA cooperative agreement NA10OAR4320143), the NOAA Coral
HABITAT OCCUPANCY PROBABILITIES FOR JUVENILE GROUPERS

Reef Conservation Program grant NA17RJ1226, NOAA MARFIN Grant NA11NMF4330129, and a Natural Sciences and Engineering Research Council of Canada postgraduate scholarship to W.J.H.

REFERENCES


