Technical Report Series HCFRU-001

ESTIMATING DETECTION AND OCCUPANCY COEFFICIENTS FOR THE PACIFIC ISLANDS CORAL REEF FISH SPECIES

Bobbie Suarez[1] and Timothy B. Grabowski[2]

[1]Tropical Conservation Biology and Environmental Science, Hawai'i Cooperative Fishery Research Unit, University of Hawai'i at Hilo, Hilo, Hawai'i 96720 USA

[2]U.S. Geological Survey, Hawai'i Cooperative Fishery Research Unit, University of Hawai'i at Hilo, Hilo, Hawai'i 96720 USA

U.S. Geological Survey, Hawai'i Cooperative Fishery Research Unit
University of Hawai'i at Hilo
200 West Kāwili Street
Hilo, Hawai'i 96720, USA

JANUARY
2021

About this report series: The Hawai'i Cooperative Fishery Research Unit Technical Report Series was initiated in 2021 to enable the archiving and retrieval of research project data and reports resulting from work conducted by the students and scientists of affiliated with the U.S. Geological Survey Hawai'i Cooperative Fishery Research Unit. All contributions to this series with a U.S. Geological Survey author have been peer reviewed and approved for publication consistent with U.S. Geological Survey Fundamental Science Practices (http://pubs.usgs.gov/circ/1367/). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

The Hawai'i Cooperative Fishery Research Unit is jointly sponsored by the U.S. Geological Survey, the University of Hawai'i System, Hawai'i Department of Land and Natural Resources, and the U.S. Fish and Wildlife Service.

Contractual references: Funding for this work was provided by the Western Pacific Regional Fishery Management Council through its Coral Reef Conservation Program Cooperative Agreement (NA17NMF4410251). Previously published documents that partially fulfilled any portion of this contract are referenced within, when applicable. (USGS IPDS #: IP-124358).

Recommended citation: Suarez, B., and T. B. Grabowski. 2021. Estimating detection and occupancy coefficients for the Pacific Islands coral reef fish species. Hawai'i Cooperative Fishery Research Unit Technical Report Series HCFRU-001. University of Hawai'i at Hilo. Available online at: http://dspace.lib.hawaii.edu/handle/10790/5552.

For additional copies or information contact:

Dr. Timothy Grabowski
U.S. Geological Survey
Hawai'i Cooperative Fishery Research Unit
University of Hawai'i at Hilo
Hilo, Hawai'i 96720, USA
Phone: 808-932-7575
e-mail: tgrabowski@usgs.gov | tbg@hawaii.edu

ESTIMATING DETECTION AND OCCUPANCY COEFFICIENTS FOR THE PACIFIC ISLANDS CORAL REEF FISH SPECIES

Bobbie Suarez[1] and Timothy B. Grabowski[2]

[1]*Tropical Conservation Biology and Environmental Science, Hawai'i Cooperative Fishery Research Unit, University of Hawai'i at Hilo, Hilo, Hawai'i 96720 USA*

[2]*U.S. Geological Survey, Hawai'i Cooperative Fishery Research Unit, University of Hawai'i at Hilo, Hilo, Hawai'i 96720 USA, email: tgrabowski@usgs.gov*

Final project report presented to the Western Pacific Regional Fishery Management Council, Honolulu, Hawai'i

25 January 2021

Summary

The data-limited stock assessment models used to monitor the status of coral reef fish species in the Western Pacific region are dependent upon accurate estimates of standing stock biomass generated from underwater visual surveys of reefs. However, the imperfect detection of and variable occupancy of habitat by reef fishes are not currently accounted for in these estimates. Therefore, the objective of this project was to estimate detection and occupancy coefficients for the species listed in the Western Pacific Regional Fishery Management Council's Fishery Ecosystem Plans by analyzing the Pacific Island Fishery Science Center-Coral Reef Ecosystem Program Reef Fish Dataset. These detection and occupancy coefficients would then be applied to refine standing stock biomass estimates. In general, species with higher detection probabilities and/or lower occupancy rates tended to exhibit the greatest differences in the estimates of standing stock biomass calculated with and without accounting for detection and occupancy. The standing stock biomass of most reef fish species seem to be underestimated when detection and occupancy are not accounted for. However, the standing stock biomass of larger-bodied targeted species, such as jacks, snappers, and groupers, seem to be over-estimated relative to the estimates generated when accounting for occupancy and detection. While there are still issues to resolve regarding how well the current data collection methods meet the underlying assumptions of the detection and occupancy modeling approach, the inclusion of detection and occupancy coefficients seems likely to improve estimates of standing stock biomass of coral reef fish species.

Acknowledgements

We thank M. Henderson, J. Hines, T. Kindinger, D. MacKenzie, R. Tabandera, and I. Williams for their technical assistance during the development of this project. M. Henderson and C. Teague provided comments and suggestions on an earlier draft of this report that greatly improved the final product. Funding for this work was provided by the Western Pacific Regional Fishery Management Council through its Coral Reef Conservation Program Cooperative Agreement (NA17NMF4410251). The Hawai'i Cooperative Fishery Research Unit is jointly sponsored by the U.S. Geological Survey, the University of Hawai'i System, Hawai'i Department of Land and Natural Resources, and the U.S. Fish and Wildlife Service. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Introduction

Coral reef fisheries are critical to the food security, economic well-being, and cultural practices of people throughout the Pacific Basin (Aswani and Sabetian 2010; Sabater and Kleiber 2014). For example, the coral reef fisheries of the main Hawaiian Islands are estimated to provide over 7 million meals annually and can add major financial contributions to an island's or region's economy by the way of tourism (van Beukering and Cesar 2004; Newton et al. 2007). Furthermore, many of the species targeted in reef fisheries play important roles in maintaining the health and resiliency of coral reef ecosystems (Jones et al. 2004). Herbivorous fishes, such as acanthurid surgeonfishes and scarid parrotfishes, provide a check on algal growth and ensure the coral colonies are not overgrown (Aswani and Sabetian 2010; Lindfield et al. 2014). Yet despite their importance, most coral reef fisheries are declining or at low levels due to a combination of habitat degradation, increased demand and effort associated with the growing populations on Pacific Islands, and improvements to fishing methods and gears (Wolanski et al. 2009; Houk et al. 2011; Keith et al. 2018).

There are numerous challenges to the effective management of these fisheries. In particular, the large number of species that comprise a typical coral reef fishery and the difficulty in effectively surveying them in structurally complex and ecologically sensitive habitats leaves managers relying on data-limited stock assessment models to evaluate the status of reef fish populations, as in the Western Pacific Regional Fishery Management Council's Fishery Ecosystem Plans. While the catch-maximum sustainable yield (MSY) approach utilizes a catch time series to estimate MSY with a plausible combination of rate of population increase (r) and carrying capacity (K; Martell and Froese 2012), most data-limited stock assessment models applied to Pacific reef fishes require reliable estimates of standing stock biomass generated from underwater visual census surveys. For example, Sabater and Kleiber (2014) augmented the catch-MSY approach with an estimated standing stock biomass from Williams et al. (2011) to anchor the biomass projections to a data-linked level. The length-based mortality estimation method developed by Nadon et al. (2015) also requires population abundance estimates from underwater visual census surveys to determine the proportion of the standing stock biomass that can be sustainably harvested. In most cases, estimates of standing stock biomass within a management unit are based on extrapolating species biomass density estimated from underwater visual census surveys to areas of known hard bottom substrate within 0 – 30 m depth (i.e., areal expansion method; Williams 2010).

This process of expanding standing stock biomass estimates from underwater visual census surveys conducted at discrete sites to generate standing stock biomass estimates at larger spatial scales includes implicit assumptions of uniform distribution and perfect detection that are unlikely to be fully met. For example, variations in fish behavior, particularly in response to the presence of divers; varying exposure among survey sites to fishing pressure; habitat type; conditions at the time of the surveys, such as turbidity and light levels; and diver survey experience affect the ability to detect species and individuals and have repercussions to the population estimators (Williams et al. 2006; MacNeil et al. 2008; Gray et al. 2016), especially for fished areas (Lindfield et al. 2014). Presence or absence of a species from a survey site at any given time can be attributed to detection or non-detection and/or occupancy or non-occupancy and/or a combination of both (MacKenzie et al. 2018). Therefore, understanding the effects of species detection and occupancy is essential for estimating population abundance based on expanding standing stock biomass estimates generated from underwater stationary point count (SPC) surveys.

Most underwater visual census surveys focus on estimating relative biomass and abundance, ignoring the effects of imperfect detection and variable occupancy. Accounting for detection errors through

distance sampling or multiple independent observers has been demonstrated to reduce the bias resulting from imperfect detection inherent to underwater visual census data (Elphick 2008; Bernard et al. 2013). The current biomass density estimate for many fish species in the Western Pacific is provided by the SPC surveys conducted by the Pacific Islands Fisheries Science Center-Coral Reef Ecosystem Program (PIFSC-CREP). The SPC is a stratified random design with paired adjacent divers, which presents an opportunity to explore species-level detectability and occupancy across a range of reef fishes and evaluate how accounting for detection and occupancy influences estimates of standing stock biomass. The PIFSC-CREP Pacific Islands coral reef underwater visual census dataset encompass nine years of surveys encompassing a total of 12,532 independent SPC surveys at 5,177 sites from 44 islands across the Western Pacific region (Heenan et al. 2017). These surveys have recorded almost 4 million individual fishes representing about 700 species. Therefore, the overall aim of this project was to analyze the PIFSC-CREP Reef Fish Dataset to estimate detection and occupancy coefficients for species listed in the Council's Fishery Ecosystem Plans to refine the standing stock biomass estimates used in stock assessment.

Methods

Fish survey data

All survey data were collected as part of the reef assessment and monitoring conducted by the PIFSC-CREP during 2010-2017 across five regions in the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa (Appendix I). Heenan et al. (2017) provided a comprehensive description of this data set, but a brief summary follows. Surveys were conducted using the paired-diver SPC method (Williams 2010; Ayotte et al. 2011; Smith 2011). The SPC method consists of two divers establishing abutting 7.5-m diameter survey cylinders extending from the substrate to the surface along a transect line. Each diver slowly rotates within the center of this cylinder and identifies, counts, and estimates the lengths of the fishes within or entering their cylinder. Fishes moving between survey cylinders are counted by each diver, but individuals that leave a cylinder and later return to the same cylinder are only counted once.

Survey sites were selected randomly within three strata: depth bins (shallow: 0-6 m; mid-depth: 6-18 m; and deep: 18-30 m), benthic habitat classification, and reef zone. Benthic habitat was classified using backscatter benthic area data provided by PIFSC-CREP and has the estimated area in hectares (ha) for each habitat type. These patches of benthic habitat types are divided into sectors as they are the smallest statistical spatial unit NOAA uses for stock assessment purposes, but each survey also included site specific-data. Habitat classification followed that of the National Centers for Coastal Ocean Science: 1) aggregate reef; 2) aggregated patch reefs; 3) individual aggregated patch reef; 4) pavement; 5) pavement with patch reefs; 6) pavement with sand channels; 7) rock/boulder; 8) reef rubble; 9) spur and groove; and 10) sand with soft coral or rock (Heenan et al. 2017). Reef zones are broader classifications of habitat type on a reef and included backreef, forereef, protected slope, and lagoon.

Construction of candidate models

Candidate models were constructed from incorporating different combinations of site-specific and survey-specific variables as covariates. Variables separate from those influenced by a surveyor were considered site-specific variables (MacKenzie et al. 2018) and included latitude, longitude, depth bin, benthic habitat classification, and the percent coverage of hard coral, macro-algae, crustose coralline algae, sand, and other (Heenan et al. 2017). The only non-continuous site-specific variable was the benthic habitat classification, i.e. spur and groove, pavement, or aggregate reef (Ayotte et al. 2011). Percent coverage was averaged between the two surveyors at each site when used as a covariate for occupancy, due to the subjectivity of the estimates as well as the fact that divers did not occupy the same volume of water during the survey. Percent coverage estimates made by each diver were used as covariates for detection.

In contrast, survey-specific variables were those that were directly related to the surveyor (MacKenzie et al. 2018). These variables included effort, diver experience, and each diver's estimate of benthic coverage within their survey cylinder. Surveyor specific benthic coverage was defined as the subjective estimate of hard coral, macro algae, crustose coralline algae, sand, or other (Heenan et al. 2017). Diver experience can potentially bias survey results (Williams et al. 2006); however, diver experience outside of the dataset was not available. Therefore, experience was represented as the number of surveys a diver conducted in a specific region rather than across the entire data set, because only about 18% of the divers in the dataset participated in surveys across multiple regions. As such, the increasing experience level of a diver that participated across multiple years of surveying was accounted for in the analysis. Furthermore, a diver that had accrued significant experience in one region would be treated as inexperienced if they conducted surveys in a different region. Only a small number of divers (< 5%) conducted surveys in multiple regions. Surveys were not standardized in terms of effort, e.g., survey duration, and the length of the dive was not recorded as part of the dataset. Therefore, effort of each survey was calculated using the theoretical maximum no-decompression time at the maximum depth recorded in the survey using the Conformité Européenne (CE) standard for breathing rate and assuming the divers were using 11.1 L air cylinders (EN14143:2003 1994), equivalent to the ubiquitous 80-cubic foot aluminum air cylinder used by NOAA.

Estimation of Detection Coefficients

We used a single-species single season occupancy modeling framework for each species in each region of the survey. Season is defined as the unit of time to which the survey was conducted. Tropical coral reef ecosystems may experience slight seasonality in the abundance of fish likely attributable to seasonal patterns in water turbidity and visibility (Madduppa et al. 2012), but species composition is relatively stable. For the purpose of this project, seasonality was not considered in our analyses. While surveys were usually conducted during the same month of every survey years (± 2 – 4 months), the same survey sites were not visited multiple times across the dataset.

We estimated the occupancy coefficient for each species using the detection history from the SPC survey time series by relating the occupancy state of species *i* to its detectability using a

two-part conditional model (MacKenzie et al. 2006). First, if a species has not been detected or observed ($y_i = 0$) then its occupancy state is unknown and can be described by the sum of two conditional probabilities where the species is either not present ($\psi_i = 0$) or present ($\psi_i = 1$) but undetected (Pi <1). This can be mathematically be represented as:

$$p(y_i = 0) = \psi_i(1-p_i)^k + (1-\psi_i) \text{ [Equation 6]}$$

where i indicates the reef fish species being modeled; y_i is the observed number of species detections; k is the number of SPC surveys per site; ψ_i is the probability of species i occupancy at a specific site; and p_i is the probability of detection for species i given it is present at a particular site. However, if species i has been detected at least once over k SPC surveys ($y_i > 0$), then the detection history is equivalent to modeling the number of detections at each site as a binomial random variable with an inflated zero class:

$$p(y_i > 0) = \psi_i\binom{k}{yi}P_i^{yi}(1-p_i)^{k-yi} \text{ [Equation 7]}$$

These models served as the basic structure for the detections of individual fish species, but covariates for detection and occupancy were incorporated using linear models and logit link:

$$\frac{\psi i}{(1-\psi i)} = \beta_0 + \beta_1 x_1 + \ldots + \beta_j x_j \text{ [Equation 8]}$$

$$\frac{pi}{(1-pi)} = \gamma_0 + \gamma_1 x_1 + \ldots + \gamma_j x_j \text{ [Equation 9]}$$

where the β's are covariate parameters for occupancy and γ's are covariate parameters for detectability. All covariates were assumed to be normally distributed.

Data analysis

We used the R-Presence v 13.05 package (Hines 2006) implemented in R v. 3.6.3 (R Core Team 2020). The R-Presence package allows for the use of multiple covariates in a single season, independent double-observer survey regime, with the assumptions that 1) the state of occupancy does not change, 2) probability of occupancy for a species is equal across surveys, 3) the probability of detecting a species in a survey is equal given its presence, 4) detection of a species is independent of other surveys of the unit, 5) detection history of each location is independent, and 6) there is no misidentification of target species (MacKenzie et al. 2018). All count data were converted to binomial presence/absence data with an observed species indicated by 1 and a species not observed indicated as 0. Continuous variables used as covariates were standardized using a mean of zero and standard deviation of 1. A set of 20 candidate models were developed and tested for all species observed in each region (Table 1). Models were assessed for fit using the Akaike's Information Criterion (AIC) and the relative fit of each model was assessed using ΔAIC (Anderson and Burnham 2004). Top models, i.e., models with a ΔAIC score ≤ 2.0, were averaged in R-Presence and the output modeled averaged estimates of both occupancy and

detection were used to re-estimate standing stock biomass. All of the R code and data used for modeling are available in the electronic supplementary material provided with this report.

Biomass re-estimation

Estimates of standing stock biomass ($\boldsymbol{SSB_i}$) were derived from the biomass density data collected from SPCs and the estimated area of each benthic habitat class. Biomass density ($\boldsymbol{b'_i}$) was calculated from the total weight of a target species that was estimated using published length-weight conversion factors where A_{lw} and B_{lw} are species-specific constants (Froese and Pauly 2000; Kulbicki et al. 2005):

$$\boldsymbol{W} = (\boldsymbol{a\ x\ L^B}) \text{ [Equation 10a]}$$

$$\boldsymbol{W} = (\boldsymbol{a\ x\ L^B})(\boldsymbol{y}) \text{ [Equation 10b]}$$

and then summed across all size bins to estimate the total weight in grams for each species:

$$\boldsymbol{Wt} = \sum_{\boldsymbol{W=i}}^{\boldsymbol{n}} \boldsymbol{i_1 + i_2 + i_3 + \cdots + i_n} \text{ [Equation 11a]}$$

$$\boldsymbol{Wt_p} = \sum_{\boldsymbol{W=i_p}}^{\boldsymbol{n}} \boldsymbol{i_{p_1} + i_{p_2} + i_{p_3} + \cdots + i_{p_n}} \text{ [Equation 11b]}$$

Biomass density (b'_i) was estimated as the quotient of the total weight ($wt_{(t)}$) divided by the area of the stationary point count (A=176.71 m^2), then doubled to account for the two cylinders in each survey:

$$\boldsymbol{b'i} = \boldsymbol{Wt}/(\boldsymbol{A} \times \boldsymbol{2}) \text{ [Equation 12a]}$$

$$\boldsymbol{b'i_p} = \boldsymbol{Wt_p}/(\boldsymbol{A} \times \boldsymbol{2}) \text{ [Equation 12b]}$$

Biomass density was estimated for each benthic habitat class by multiplying with the estimated area for hardbottom substrate per depth strata (A_i). Mean biomass densities were calculated from all sites within a depth bin and benthic habitat class and multiplied by the area of each benthic habitat class to derive standing stock biomass. Details of the expansion method are described in Williams (2010). The standing stock biomass expansion for a single habitat class can be represented by:

$$\boldsymbol{SSB_i} = \overline{\boldsymbol{b'\imath}}\ \boldsymbol{x}\ \boldsymbol{A_i} \text{ [Equation 13]}$$

The values of SSB_i for each habitat class in each depth strata were then be summed to generated an overall estimate of standing stock biomass. The model-averaged species- and site-specific detection probabilities were included in the re-estimation of the standing stock biomass by multiplying the biomass density and the model-averaged species-site specific detection coefficients with the estimated for hardbottom substrate per depth strata with the model-averaged species- and site-specific occupancy coefficient:

$$\boldsymbol{SSB_i} = (\boldsymbol{b'i})(\boldsymbol{Ai} \times \boldsymbol{\psi}) \text{ [Equation 14a]}$$

$$\boldsymbol{SSB_{i_{(\frac{p}{\psi})}} = (p_i \times b'i)(Ai\ x\ \psi)}\ [Equation\ 14b]$$

As with *SSBi*, the values of $SSB'i_{(p/\psi)}$ were summed across each habitat class within each depth strata to generate an estimate of standing stock biomass corrected for $p < 1.0$ and $\psi < 1.0$. Finally, the percent difference between SSB_i *and* $SSB'i_{(p/\Psi)}$ was calculated as:

$$\%_{\boldsymbol{difference}} = [\boldsymbol{SSB_i - SSB_{i_{(\frac{p}{\psi})}}}] \times \mathbf{100}\ \text{[Equation 15]}$$

Results

We were able to model 549 species representing 55 families of coral reef fishes from across the Pacific Basin. The full set of model selection results, the parameter estimates and associated error measures, and re-estimated standing stock biomass for each species by region are provided in the electronic supplementary materials submitted as the primary deliverable for this project. As a general rule, species observed during less than about 25% of surveys within a region yielded models that were inconclusive with poor explanatory power. The detection and occupancy coefficients generated for these species that were rarely observed in the SPC surveys yielded corrected estimates of standing stock biomass that were ≥ 1000% larger than the uncorrected estimates and possessed very wide confidence intervals. While a high proportion of the species within the families Apogonidae, Muraenidae, and Scorpaenidae fit this pattern, there were representatives in almost every family that were rare in the surveys and resulted in unreliable corrected standing stock biomass estimates.

Regardless of region and species, candidate model 15 – ψ (latitude*longitude*depth bin); p (diver 1 experience*diver 2 experience*effort), was identified among the top-ranked models (ΔAIC ≤ 2.0) most frequently (15.0%; Figure 1). Candidate models where detection was held constant, such as models 2, 3, and 4, and model 12 – ψ (.); p (effort), also appeared relatively frequently as top-ranked models. Most of the models incorporating habitat classification or percent substrate composition as covariates were ranked as top models ≤ 5.0% of the time. The null (model 1) and global (model 20) models were included among the top-ranked models for a species in a region only about 7.5% of the time. This pattern of the likelihood of models to be highly ranked was generally consistent within families of coral reef fishes (Figure 2); however, models incorporating longitude, latitude, and depth as covariates for occupancy, i.e., models 16 and 17, also were ranked highly in Acanthuridae, Balistidae, Chaetodontidae, Labridae, Mullidae and Pomacentridae. At the level of individual species that occurred across multiple regions, the top-ranked model usually differed between regions.

The degree of difference between spawning stock biomass estimated directly from the PIFSC-CREP SPC survey data and the same estimates corrected by accounting for imperfect detection and occupancy was strongly correlated to the species-specific estimates for detection probability and occupancy. Detection probability was negatively correlated with the absolute difference between the two estimates (Spearmans's *rho* = -0.60; $P < 0.001$), while occupancy exhibited a positive correlation to the absolute difference between the two estimates (Spearmans's *rho* =

0.27; $P < 0.001$). Indicating that surveys of species with higher detection probabilities and/or lower occupancy coefficients tended to produce corrected standing stock biomass estimates more similar to the uncorrected ones. However, these relationships were reversed when the directionality of the true value of the difference between the standing stock biomass estimates was considered, with detection exhibiting a positive correlation (Spearmans's *rho* = 0.60; $P < 0.001$) and occupancy exhibiting a negative correlation (Spearmans's *rho* = -0.66; $P < 0.001$). This pattern indicates that the original standing stock biomass of species with high detection probabilities tended to be underestimates relative to the corrected standing stock biomass estimate and that the original estimate of biomass of species with low occupancy rates tended to be overestimates relative to the corrected one (Figure 3).

This relationship between detection, occupancy, and degree of difference between the corrected and uncorrected estimates of standing stock biomass did not vary by region, but resulted in an interesting trend across reef fish families (Figure 4). Families consisting of species with relatively low detectability, but moderate to high occupancy, such as Chaetodontidae, Holocentridae, Labridae, Pomacentridae, and Serranidae, had a large proportion, approximately 72-75%, of their species with corrected standing stock biomass estimates that were greater than the uncorrected estimates. In contrast, families consisting of highly detectable species with low to moderate occupancy, such as Carangidae (88%) and Lutjanidae (51%), tended to have a high proportion of species for which the uncorrected biomass estimates were higher than the corrected ones. Approximately 40% of the species from the remaining families had uncorrected standing stock biomass estimates that were higher than those corrected for occupancy and imperfect detection.

Conclusions

Incorporating detection and occupancy coefficients has the potential to improve estimates of standing stock biomass using areal expansion methods. In terms of how correcting for imperfect detection and occupancy being < 1.0 would influence estimates of biomass, our results are consistent with those of other, primarily terrestrial studies (Kroll et al. 2008; Hines et al. 2010; Bozec et al. 2011; Walls et al. 2011; Green et al. 2013). For example, Bozec et al. (2011) described the effect of both species-specific intrinsic factors, such as body size, schooling behavior, and reaction to divers, as well as environmental factors, such as visibility, on the detection probabilities of reef fishes in New Caledonia from distance sampling. Detection probabilities varied among families, species within families, and species across sites (Bozec et al. 2011). However, there was no attempt to incorporate these findings into the estimation of biomass of these species in New Caledonia. In contrast, Green et al. (2013) estimated the detection probabilities of invasive Indo-Pacific lionfishes, *Pterois* spp., from conventional underwater visual census methods employed on Caribbean reefs. Lionfish biomass estimated from conventional survey methods was approximately 200% less than more specialized survey methods and that detection covaried with lionfish size and substrate rugosity. However, we were unable to find reports detailing attempts to incorporate occupancy coefficients into estimates of

reef fish biomass at broad spatial scales. This effort represents one of the first, large-scale efforts to incorporate detection and occupancy into the stock assessment of coral reef fishes and as such, there are several aspects of this approach that require validation before this approach is widely applied to SPC data collected using the current standard protocol implemented by PIFSC-CREP (Williams 2010; Ayotte et al. 2011; Smith 2011).

Detection probabilities and occupancy coefficients calculated from double observer surveys are generated based on the assumption that the two observers are searching the same area/volume for the target species, i.e., all of the individuals seen by the first observer are available to be seen by the second and vice versa. While the SPC data are temporally paired, the two divers are surveying two separate volumes of water that likely exhibit some level of independence depending on the population density and behavior of the fishes within or passing through them. This independence is problematic from the standpoint of estimating detection probability coefficients. For example, if diver 1 and diver 2 both report sighting three individuals of species X, it is not clear whether the two divers saw a total of three individuals, six individuals or some number in between. This is likely ameliorated to some extent through the use of presence-absence data and incorporating an unreconciled double-observer approach (Riddle et al. 2010). However, it does not fully avoid the issue that the two divers are surveying different volumes of water. Assessing the effects of violating this underlying assumption of the occupancy model could be assessed empirically by conducting paired surveys that combine a two independent observer survey using the standard PIFSC-CREP SPC protocols and a two independent observer wherein two divers conduct a SPC in the same volume of water. We were hoping to report the results of our assessment using this method, but the COVID-19 pandemic delayed the initiation of fieldwork during 2020. However, data collection is currently underway at the time of writing. Secondarily, the effects of the two divers surveying adjacent volumes of water on estimates of detection and occupancy coefficients could also be assessed by using the actual counts generated during the surveys in an *N*-mixture model framework using the Royle-Nichols (2003) approach. This approach would allow for the possibility of modeling finer scale occupancy coefficients for rare or cryptic species with variation in the abundance over a greater study area because the area size being sampled is controlled (Royle and Nichols 2003). Until such time that the effects that the currently employed survey protocol has on estimates of detection and occupancy coefficients are better understood, caution in the interpretation of detection and occupancy corrected estimates of standing stock biomass is warranted.

Interestingly, the habitat classes and substrate composition variables employed in our models did not comprise a particularly important component of the top candidate models. The lack of strong support for models incorporating habitat covariates for detection and occupancy is contrary to the general perception of reef fishes as being habitat and/or dietary specialists (Ferry-Graham et al. 2002; Wainwright and Bellwood 2002). It is possible that the low explanatory power of the habitat variables employed in our models is due to the scale, accuracy, or precision at which they were measured. However, it seems more likely that these variables simply are not particularly influential on the detection probabilities and occupancy patterns of Pacific reef fishes and that other factors not incorporated into these analyses, such as availability of recruits or assemblage composition are more important. However, Red Snapper *Lutjanus campechanus* on

temperate/subtropical reefs off the east coast of Florida and Georgia exhibited occupancy patterns there were more strongly influenced by latitude and depth than any finer scale habitat metric assessed, suggesting a greater degree of generality in habitat use than expected (Coggins et al. 2014). Further, recent research seems to suggest that reef fishes may exhibit a lower degree of dietary and habitat specificity than previously assumed (Brandl et al. 2015; MacDonald et al. 2018) supporting the results of our models that habitat class or composition are not the primary determinants of occupancy for most Pacific reef fishes.

Assuming that the effects of divers surveying adjacent volumes of water on the estimation of detection and occupancy coefficients are minimal, our results have several important implications to the management of reef fishes in the Pacific Basin. In particular, the degree to which standing stock biomass of many of the more heavily-targeted reef fish species; particularly carangids and lutjanids, but also a large proportion of acanthurid surgeonfishes, mullid goatfishes and scarid parrotfishes; may be overestimated by the current areal expansion method is important to note and warrants further examination. Consistent overestimation of standing stock biomass could lead to a failure to detect overfishing or adoption of management actions that are ineffective or even counterproductive to maintaining sustainable populations of these species (Sabater and Kleiber 2014; Gray et al. 2016). However, most reef fish species seem to be underestimated when detection and occupancy are not accounted for, as seen in many of the chaetodontids, holocentrids, labrids, and serranids. While this potential underestimation would likely result in a more conservative management strategy than would be necessary to ensure sustainable reef fish populations, there is the risk of unintended socio-economic impacts if regulations impact subsistence or small-scale commercial fishing or reduce the credibility of management agencies (Gu and Swihart 2004; Gray et al. 2016).

Literature cited

Anderson, D., and K. Burnham. 2004. Model selection and multi-model inference, 2nd edition. New York: Springer-Verlag.

Aswani, S., and A. Sabetian. 2010. Implications of urbanization for artisanal parrotfish fisheries in the Western Solomon Islands. Conservation Biology 24:520-530.

Ayotte, P. K., I. McCoy, I. Williams, and J. Zamzow. 2011. Coral Reef Ecosystem Division standard operating procedures: data collection for Rapid Ecological Assessment fish surveys. Pacific Islands Fisheries Science Center, H-11-08.

Bernard, A. T. F., A. Götz, S. E. Kerwath, and C. G. Wilke. 2013. Observer bias and detection probability in underwater visual census of fish assemblages measured with independent double-observers. Journal of Experimental Marine Biology and Ecology 443:75-84.

Bozec, Y.-M., M. Kulbicki, F. Laloë, G. Mou-Tham, and D. Gascuel. 2011. Factors affecting the detection distances of reef fish: implications for visual counts. Marine Biology 158:969-981.

Brandl, S. J,, W. D. Robbins, and D. W. Bellwood. 2015. Exploring the nature of ecological specialization in a coral reef fish community: morphology, diet and foraging microhabitat use. Proceedings of the Royal Society B. 282: 20151147.

Coggins Jr., L. G., N. M. Bacheler, and D. C. Gwinn. 2014. Occupancy models for monitoring marine fish: a Bayesian hierarchical approach to model imperfect detection with a novel gear combination. PLoS ONE 9(9): e108302.
Elphick, C. S. 2008. How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology 45:1313-1320.
EN14143:2003. 1994. Personal Protective Equipment (EC Directive) Regulations S.I. 1992/3139. U.S.D.O.T. a. Industry, editor.
Ferry-Graham, L. A., D. I. Bolnick, and P. C. Wainwright. 2002. Using functional morphology to examine the ecology and evolution of specialization. Integrative and Comparative Biology 42: 265-277.
Froese, R., and D. Pauly. 2000. "FishBase". R. Froese, and D. Pauly, editors.
Gray, A. E., I. D. Williams, K. A. Stamoulis, R. C. Boland, K. C. Lino, B. B. Hauk, J. C. Leonard, J. J. Rooney, J. M. Asher, K. H. Lopes Jr, and R. K. Kosaki. 2016. Comparison of reef fish survey data gathered by open and closed circuit SCUBA divers reveals differences in areas with higher fishing pressure. PLoS One 11(12):e0167724.
Green, S. J., N. Tamburello, S. E. Miller, J. L. Akins, and I. M. Côté. 2013. Habitat complexity and fish size affect the detection of Indo-Pacific lionfish on invaded coral reefs. Coral Reefs 32:413-421.
Gu, W., and R. K. Swihart. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. Biological Conservation 116:195-203.
Heenan, A., I. D. Williams, T. Acoba, A. DesRochers, R. K. Kosaki, T. Kanemura, M. O. Nadon and R. E. Brainard. 2017. Long-term monitoring of coral reef fish assemblages in the Western Central Pacific. Scientific Data 4:170176-170176.
Hines, J. E. 2006. PRESENCE - Software to estimate patch occupancy and related parameters.
Hines, J. E., J. D. Nichols, J. A. Royle, D. I. MacKenzie, A. M. Gopalaswamy, N. Samba Kumar, and K. U. Karanth. 2010. Tigers on trails: occupancy modeling for cluster sampling. Ecological Applications 20:1456-1466.
Houk, P., K. Rhodes, J. Cuetos-Bueno, S. Lindfield, V. Fread, and J. L. McIlwain. 2011. Commercial coral-reef fisheries across Micronesia: A need for improving management. Coral Reefs 31:13-26.
Jones, G. P., M. I. McCormick, M. Srinivasan, and J. V. Eagle. 2004. Coral decline threatens fish biodiversity in marine reserves. Proceedings of the National Academy of Sciences 101:8251-8253.
Keith, S. A., A. H. Baird, J. -P. A. Hobbs, E. S. Woolsey, A. S. Hoey, N. Fadli, and N. J. Sanders. 2018. Synchronous behavioural shifts in reef fishes linked to mass coral bleaching. Nature Climate Change 8:986-991.
Kroll, A. J., K. Risenhoover, T. McBride, E. Beach, B. J. Kernohan, J. Light, and J. Bacha. 2008. Factors influencing stream occupancy and detection probability parameters of stream-associated amphibians in commercial forests of Oregon and Washington, USA. Forest Ecology and Management 255:3726-3735.
Kulbicki, M., N. Guillemot, and M. Amand. 2005. A general approach to length-weight relationships for New Caledonian lagoon fishes. Cybium 29:235-252.
Lindfield, S. J., E. S. Harvey, J. L. McIlwain, and A. R. Halford. 2014. Silent fish surveys: bubble-free diving highlights inaccuracies associated with SCUBA-based surveys in heavily fished areas. Methods in Ecology and Evolution 5:1061-1069.

MacDonald, C., M. I. Tauati, and G. P. Jones. 2018. Depth patterns in microhabitat versatility and selectivity in coral reef damselfishes. Marine Biology 165:1-14.
MacKenzie, D. I., J. Nichols, J. Royle, K. Pollock, L. Bailey, and J. Hines. 2018. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence, 2nd edition. Burlington, Massachusetts: Elsevier/Academic Press.
MacNeil, A., C. Fonnesbeck, and T. McClanahan. 2008. Occupancy models for estimating the size of reef fish communities. Pages 785-789 *in*: B. Riegle and R. Dodge, editors. Proceedings of the 11th International Coral Reef Symposium. Nova Southeastern University: National Coral Reef Institute.
Madduppa, H., S. Ferse, U. Aktani, and H. Palm. 2012. Seasonal trends and fish-habitat associations around Pari Island, Indonesia: Setting a baseline for environmental monitoring. Environmental Biology of Fishes 95:383-398.
Nadon, M. O., J. S. Ault, I. D. Williams, S. G. Smith, and G. T. DiNardo. 2015. Length-based assessment of coral reef fish populations in the Main and Northwestern Hawaiian Islands. PLoS One 10(8):e0133960.
Newton, K., I. M. Côté, G. M. Pilling, S. Jennings, and N. K. Dulvy. 2007. Current and future sustainability of island coral reef fisheries. Current Biology 17:655-658.
Nichols, J. D., J.E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. The Auk 117:393-408.
R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
Riddle, J., K. Pollock, and T. Simons. 2010. An unreconciled double-observer method for estimating detection probability and abundance. The Auk 127:841-849.
Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology 84:777-790.
Sabater, M., and P. Kleiber. 2014. Augmented catch-MSY approach to fishery management in coral-associated fisheries. Pages 199-218 *in* Bortone, S. A., ed. Interrealtionships between coral and fisheries. Boca Raton, Florida: Taylor and Francis.
Smith, S. G. 2011. Multispecies survey design for assessing reef-fish stocks, spatially explicit management performance, and ecosystem condition. Fisheries Research 109:25-41.
van Beukering, P., and H. Cesar. 2004. Economic analysis of marine managed areas in the Main Hawaiian Islands. University of Hawai'i, USA : H.C.R.I. Research Program. Available online at: https://www.nodc.noaa.gov/archive/arc0017/0001756/1.1/data/1-data/paper-economicanalysis_final_30apr04.pdf [accessed 06 December 2020]
Wainwright, P. C., and D. R. Bellwood. 2002. Ecomorphology of feeding in coral reef fishes. Pages 33-55 *in* Coral reef fishes: dynamics and diversity in a complex ecosystem (ed. Sale, P. E.). San Diego, CA: Academic Press.
Walls, S. C., J. H. Waddle, and R. M. Dorazio. 2011. Estimating occupancy dynamics in an anuran assemblage from Louisiana, USA. Journal of Wildlife Management 75:751-761.
Williams, I. 2010. US Pacific reef fish biomass estimates based on visual survey data. NOAA, National Marine Fishery Service, Pacific Island Fishery Science Center (PIFSC). PIFSC Internal Report: IR-10-024.
Williams, I., W. Walsh, B. Tissot, and L. Hallacher. 2006. Impact of observers' experience level on counts of fishes in underwater visual surveys. Marine Ecology Progress Series 310:185-191.

Williams, I. D., B. L. Richards, S. A. Sandin, J. K. Baum, R. E. Schroeder, M. O. Nadon, B. Zgliczynski, P. Craig, J. L. McIlwain, and R. E. Brainard 2011. Differences in reef fish assemblages between populated and remote reefs spanning multiple archipelagos across the central and western Pacific. Journal of Marine Biology 2011:826234.

Wolanski, E., J. A. Martinez, and R. H. Richmond. 2009. Quantifying the impact of watershed urbanization on a coral reef: Maunalua Bay, Hawaii. Estuarine, Coastal and Shelf Science 84:259-268.

Table 1. Candidate model set evaluated to generate occupancy (ψ) and detection (p) coefficients for Pacific reef fish in five regions of the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa, using data from paired stationary point count fish surveys conducted by the Pacific Island Fisheries Science Center – Coral Reef Ecosystem Program during 2010-2017.

Model Number	Model
1	ψ(.); p(.)
2	ψ(latitude); p(.)
3	ψ(longitude); p(.)
4	ψ(latitude, longitude); p(.)
5	ψ(habitat); p(.)
6	ψ(habitat, depth); p(.)
7	ψ(latitude, longitude, habitat, depth); p(.)
8	ψ(.); p(HC, MA, CCA, sand, other)
9	ψ(.); p(D1_exp, D2_exp)
10	ψ(.); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other)
11	ψ(.); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, D1_exp, D2_exp)
12	ψ(.); p(effort)
13	ψ(.); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, D1_exp, D2_exp, effort)
14	ψ(latitude, longitude, habitat); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, effort)
15	ψ(latitude, longitude, depth); p(D1_exp, D2_exp, effort)
16	ψ(latitude, longitude, habitat, depth); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, D1_exp, D2_exp)
17	ψ(latitude, longitude, habitat, depth); p(effort)
18	ψ(latitude, longitude, habitat); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, D1_exp, D2_exp, effort)
19	ψ(habitat, depth); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other)
20	ψ(latitude, longitude, habitat, depth); p(D1_HC, D1_MA, D1_CCA, D1_sand, D1_other, D2_HC, D2_MA, D2_CCA, D2_sand, D2_other, D1_exp, D2_exp, effort)

latitude = Distance in degrees from the Equator, **longitude** = distance from the 144˚E, **habitat** = categorical three letter designation by PIFSC-CREP for overall habitat of the survey site, **HC** = mean hard coral coverage of both survey cylinders, **MA** = mean macro algae coverage of both survey cylinder, **CCA** = mean crustose coralline algae coverage of both survey cylinders, **sand** = mean sand coverage of both survey cylinders, **other** = mean coverage of both survey cylinders that is made up of something other than the aforementioned categories, **exp** = experience measured as the cumulative number of surveys that a diver has done in a region, **D_1** & **D_2** = these are designated as the more experienced diver being D_1 and the lesser experienced being D_2; there is also subjective percent benthic coverage by each divers survey area, **depth** = maximum depth in meters for each site, **effort** = the theoretical maximum no-decompression limit based on depth and CE gas consumption rates

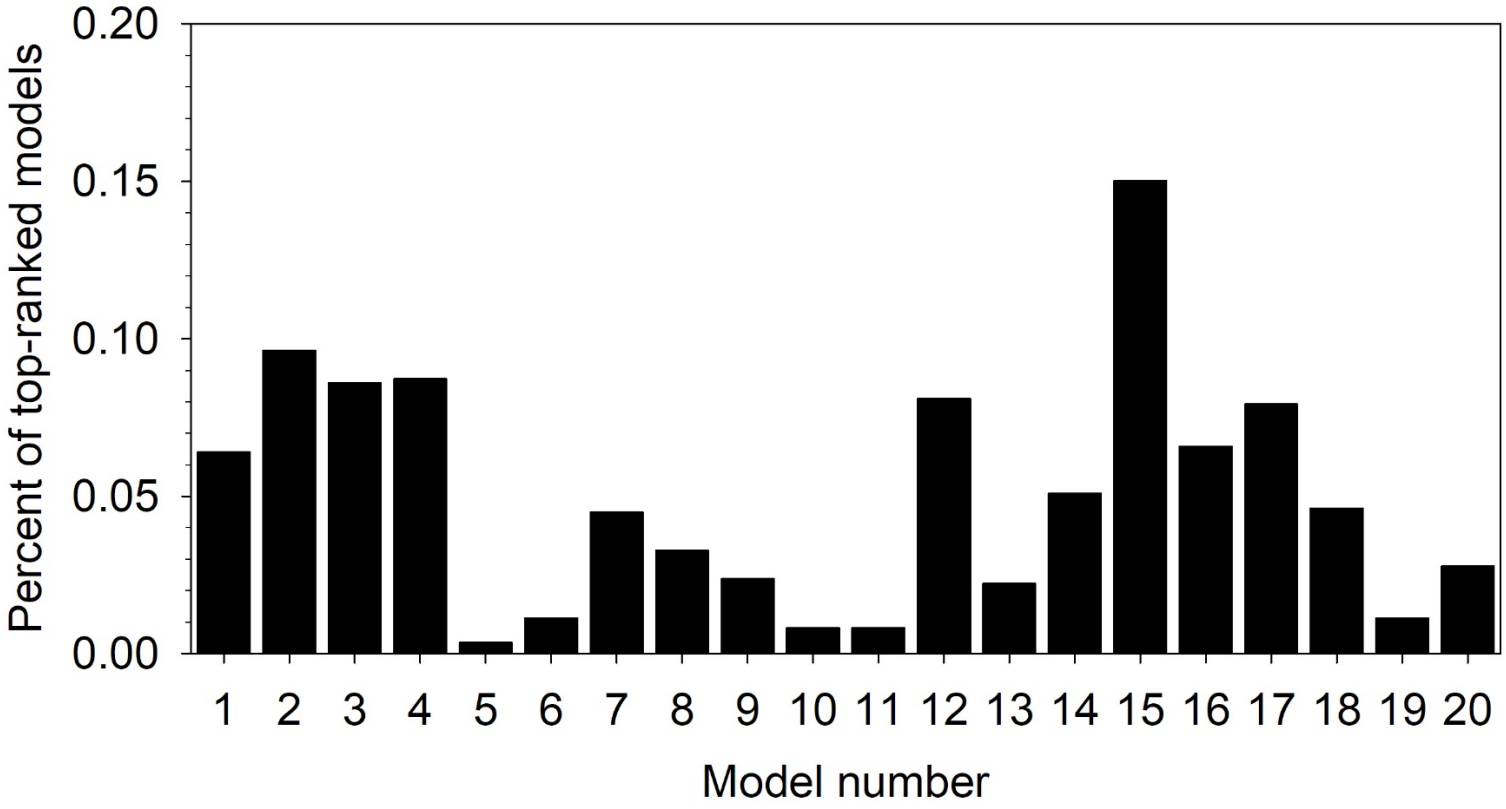

Figure 1. Frequency of occurrence of candidate models as top-ranked models ($\Delta AIC \leq 2.0$) when modeling the detection probability and occupancy coefficients of coral reef fishes in five regions of the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa, using data from paired stationary point count fish surveys conducted by the Pacific Island Fisheries Science Center – Coral Reef Ecosystem Program during 2010-2017. Model descriptions are in Table 1.

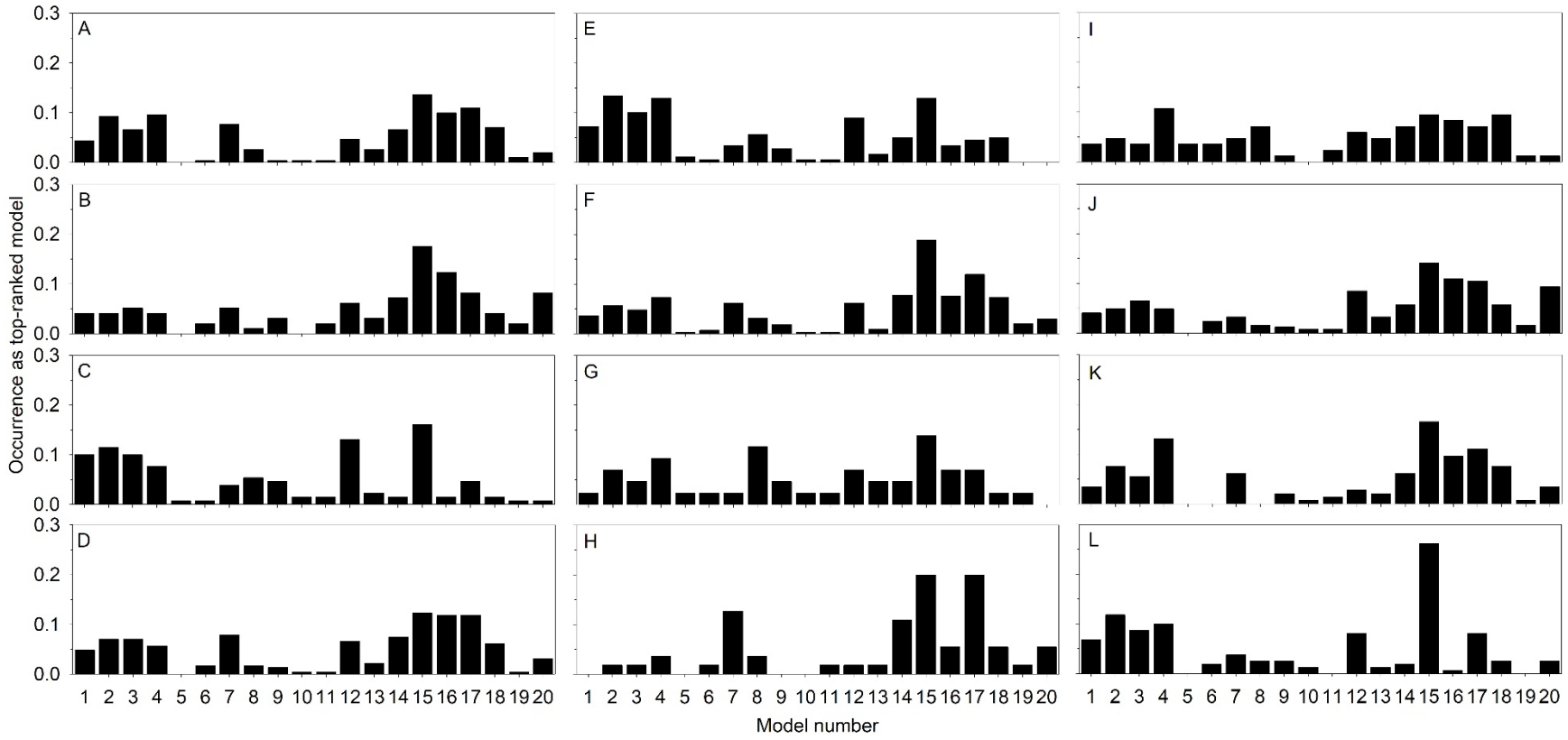

Figure 2. Frequency of occurrence of candidate models as top-ranked models ($\Delta AIC \leq 2.0$) when modeling the detection probability and occupancy coefficients of coral reef fishes in the family Acanthuridae (A), Balistidae (B), Carangidae (C), Chaetodontidae (D), Holocentridae (E), Labridae (F), Lethrinidae (G), Lutjanidae (H), Mullidae (I), Pomacentridae (J), Scaridae (K), and Serrandiae (L) across five regions of the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa, using data from paired stationary point count fish surveys conducted by the Pacific Island Fisheries Science Center – Coral Reef Ecosystem Program during 2010-2017. Model descriptions are in Table 1.

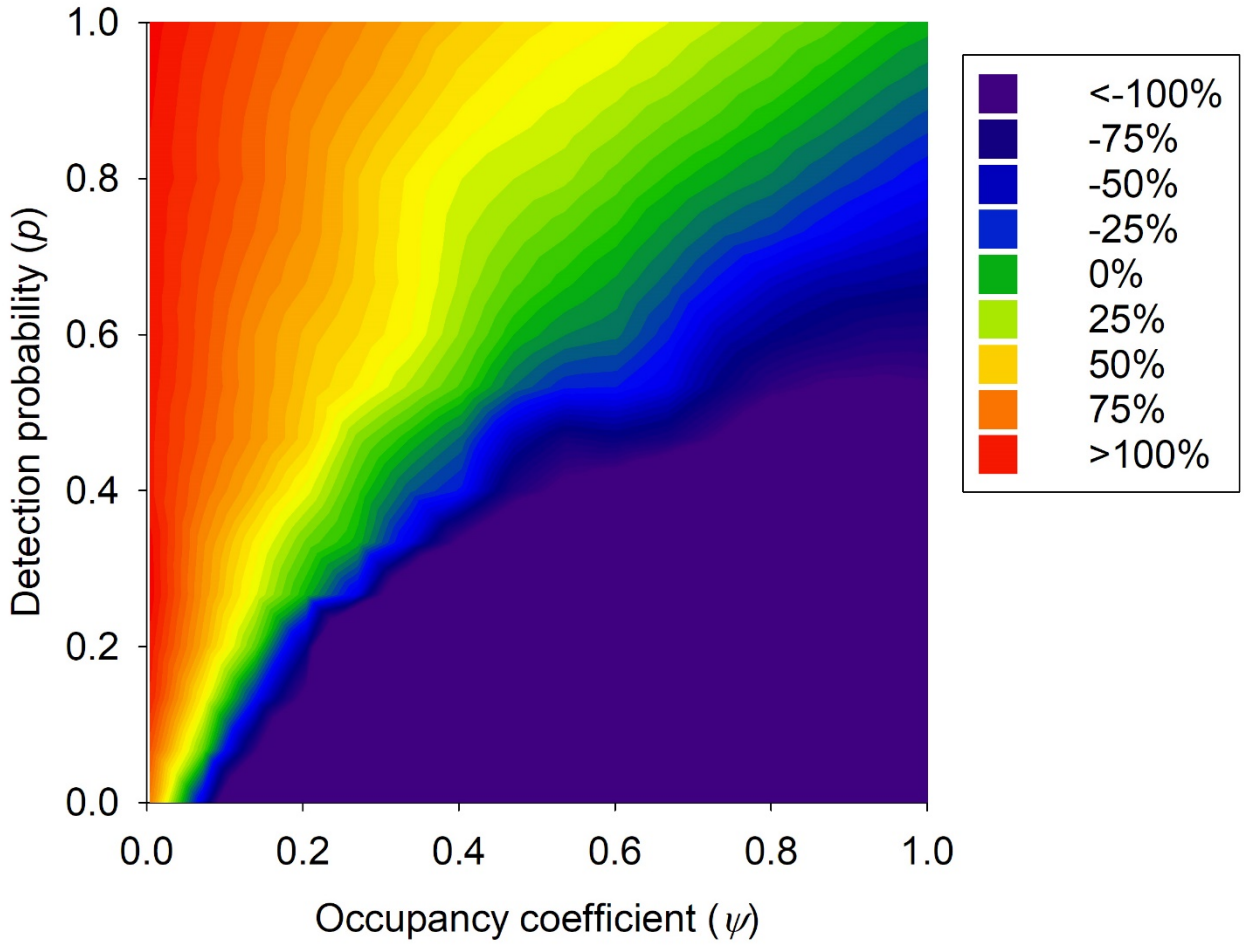

Figure 3. Relationship between detection probability (p), occupancy coefficient (ψ), and the percent difference between standing stock biomass estimated by the areal expansion method with and without accounting for p and ψ of coral reef fishes in five regions of the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa, using data from paired stationary point count fish surveys conducted by the Pacific Island Fisheries Science Center – Coral Reef Ecosystem Program during 2010-2017.

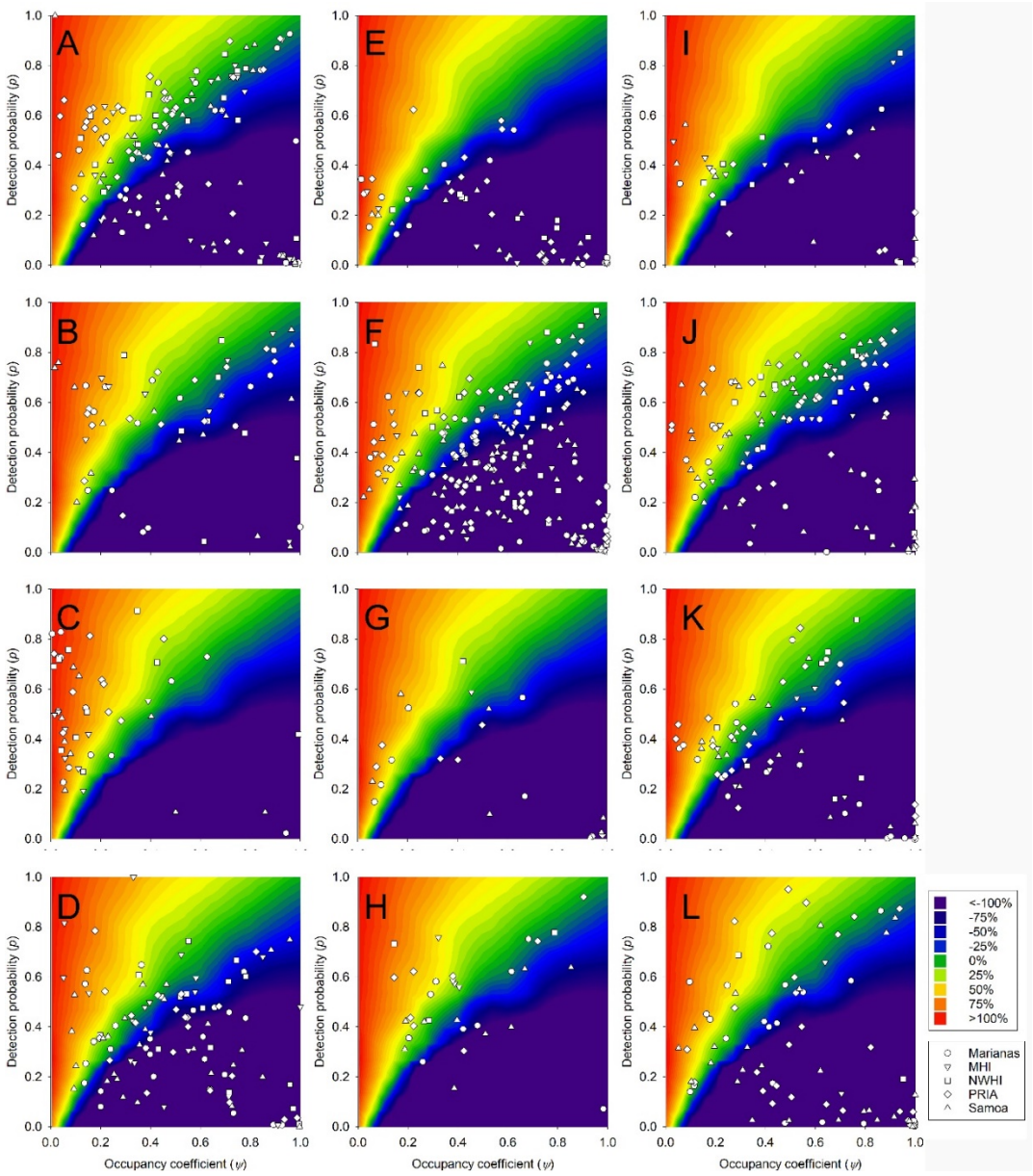

Figure 4. Relationship between detection probability (p), occupancy coefficient (ψ), and the percent difference between standing stock biomass estimated by the areal expansion method with and without accounting for p and ψ of Acanthuridae (A), Balistidae (B), Carangidae (C), Chaetodontidae (D), Holocentridae (E), Labridae (F), Lethrinidae (G), Lutjanidae (H), Mullidae (I), Pomacentridae (J), Scaridae (K), and Serranidae (L) across five regions of the Pacific Basin: Main Hawaiian Islands, Northwest Hawaiian Islands, Marianas Islands, Pacific Remote Island Area, and Samoa, using data from paired stationary point count fish surveys conducted by the Pacific Island Fisheries Science Center – Coral Reef Ecosystem Program during 2010-2017. The contour plot shown on all panels is the same as that shown in Figure 3 and represents the relationship of p and ψ on the percent difference between standing stock biomass estimated by the areal expansion method with and without accounting for p and ψ of the entire 2010-2017 dataset.