A Simulation Framework for Evaluating Multiple Fishery-independent Survey Strategies in the Main Hawaiian Islands

Kisei R. Tanaka and Thomas A. Oliver

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration

National Marine Fisheries Service
Pacific Islands Fisheries Science Center

NOAA Technical Memorandum NMFS-PIFSC-134
https://doi.org/10.25923/dqaf-wa04

October 2022

A Simulation Framework for Evaluating Multiple Fishery-independent Survey Strategies in the Main Hawaiian Islands

Kisei R. Tanaka[1] and Thomas A. Oliver[1]

[1] Pacific Islands Fisheries Science Center
National Marine Fisheries Service
1845 Wasp Boulevard
Honolulu, HI 96818

NOAA Technical Memorandum NMFS-PIFSC-134

October 2022

U.S. Department of Commerce

Gina Raimondo, Secretary

National Oceanic and Atmospheric Administration
Richard W. Spinrad, Ph.D., NOAA Administrator

National Marine Fisheries Service
Janet Coit, Assistant Administrator for Fisheries

About this report

The Pacific Islands Fisheries Science Center of NOAA's National Marine Fisheries Service uses the NOAA Technical Memorandum NMFS-PIFSC series to disseminate scientific and technical information that has been scientifically reviewed and edited. Documents within this series reflect sound professional work and may be referenced in the formal scientific and technical literature.

Recommended citation

Tanaka KR, Oliver TA. 2022. A simulation framework for evaluation multiple fishery-independent survey strategies in the main Hawaiian Islands. U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-PIFSC-134, 30 p. doi:10.25923/dqaf-wa04

Acknowledgements

The authors would like to acknowledge the NOAA Coral Reef Conservation Program for their generous funding and support of this project (#31295).

Copies of this report are available from

Pacific Islands Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
1845 Wasp Boulevard, Building #176
Honolulu, Hawaii 96818

Or online at

https://repository.library.noaa.gov/

Table of Contents

List of Tables

List of Figures

List of supplemental figures

Executive Summary

Selecting an effective and cost-efficient sampling strategy in scientific surveys is a major concern in the management of living marine resources. This is particularly true when the target populations are highly structured over space and time, and the allocation of survey effort and resources are subject to logistical limitations and uncertainties. We developed a simulation framework to evaluate sampling strategies within the data-limited main Hawaiian Island region based on 10 years of *in situ* ecological fishery-independent surveys. Specifically, we compare quantitative precision and bias of the spatiotemporal distribution of reef fish biomass estimates among functional levels using three contrasting stratified random survey designs; (1) geographically comprehensive ("traditional"), (2) ecologically homogeneous ("zone-based"), and (3) ecologically homogeneous but geographically reduced ("zone-triaged") stratified designs.

We applied a stratified survey strategy simulation to functional reef fish biomass, modeled with the observed biogeographical characteristics. The simulation scheme allowed us to incorporate varying sampling efforts to analyze the sensitivity of estimated biomass to the different strategies. Beyond this specific application, the simulation framework that we develop here will be a useful tool for evaluating the efficacy of fishery-independent surveys.

We found that across all three survey scenarios, the ecologically homogeneous but geographically reduced ("zone-triaged") stratified design outperformed in terms of accuracy when sampling efforts were much lower than historical averages (2013, 2016, 2019). However, differences between the surveys' performance became negligible when the scale of sampling efforts were matched to historical averages. This result indicates that measuring "fewer, regionally representative, ecologically homogeneous strata" (i.e. zone-triaged stratified designs) could be an acceptable strategy when survey resources are limited. This study showed that ecologically focused, sub-island stratification schemes could mitigate the loss of precision of survey estimates due to severe reduction of sampling efforts. However, under "normal", historical allocation of effort and quality standards, this methodological shift would not lead to substantial improvement in survey error characteristics.

Introduction

Marine populations exhibit complex spatiotemporal dynamics, presenting various logistical challenges in deriving accurate and precise estimates of their management-relevant parameters (Katsanevakis et al. 2012; Murphy and Jenkins 2010; Tommasi et al. 2017). Designing and implementing effective survey protocols are critical prerequisites for many living marine resource (LMR) monitoring programs that are needed to quantify the risk of overexploitation and to evaluate the efficacy and precautionary levels of alternate management plans (Murphy and Jenkins 2010). Implementing effective and efficient surveys with a solid basis in statistical sampling schemes has become a core component of many science-based LMR monitoring programs worldwide (Gunderson 1993).

A common logistical problem of scientific survey design is the balance between the cost and the efficacy of the survey design and sampling protocol (Gunderson 1993; Tyre et al. 2003). Low-cost, reduced sampling survey efforts can lead to reduced precision in the estimate of target population size and structure. Expensive but statistically inappropriate survey efforts can lead to both over-sampling and potentially correlated metrics. These issues are often difficult to identify and avoid on-site as survey practitioners may target multiple populations under various spatial scales and sampling efforts. Furthermore, fisheries-independent surveys can be hindered or reduced by temporary or permanent reductions in resources or funding (e.g., vessel breakdowns, unexpected or unfavorable weather conditions; ICES, 2020; Zimmermann & Enberg, 2017). The ongoing COVID-19 pandemic provides a clear example of how severe an unexpected reduction can occur to ongoing survey efforts, as scientists and managers are challenged with revised planning and quantitative evaluations to minimize information gaps and uncertainties in subsequent assessments or management advice (ICES, 2020). There is a need to better understand the uncertainties associated with changes to established surveys with standardized survey protocols.

The fishery-independent biomass estimates for many reef fish species in the main Hawaiian Islands (MHI) region are available through the stratified random stationary point count (SPC) surveys as a part of the NOAA Coral Reef Conservation Program's National Coral Reef Monitoring Program (NCRMP) (Heenan et al. 2017; Towle et al. 2022). Since 2010, the NCRMP has focused on collecting geographically comprehensive in situ data and providing periodic assessments of coral reef ecosystems in the United States. Most LMR monitoring programs throughout the western-central Pacific are limited in their survey resources and often lack ecological inferences of spatiotemporal LMR trends (e.g., annual and bi-annual fishery-independent surveys). Due to the large size of its survey domain and logistical constraints, the Pacific NCRMP survey protocol employs a geographically comprehensive but temporally sparse sampling scheme (i.e., "wide-but-thin"). While the insufficiency of LMR monitoring data can potentially hinder effective management, a variety of alternate methods has been proposed to derive unbiased estimates of LMR status across time and space (Oliver et al. 2020).

Simulation studies are commonly used to quantify and compare the efficacy of alternate sampling strategies (e.g., Li et al. 2015; Regular et al. 2020). The NCRMP underwater visual census data set presents an ideal opportunity to explore the impacts of alternative survey designs on the functional group-level estimates of reef fish biomass at an island scale. Stakeholders involved in existing survey programs are often skeptical of hypothesized changes to existing sampling protocols until the potential benefits of alternative designs have been empirically demonstrated. To this end, we developed a generalized simulation-based framework to examine alternative statistical LMR survey designs at different levels of sampling effort within a data-poor environment.

We hypothesized that replacing the existing geographically comprehensive survey (i.e., covering all reef areas) with two alternate survey designs that employ "ecologically homogeneous strata" and "geographically representative strata" can maintain or improve our ability to quantify reef fish biomass and spatial distribution and infer temporal trends (Oliver et al. 2020). Our simulation framework follows these steps: 1) reconstruct a spatiotemporally structured biomasses of four defined functional groups across the MHI survey domain, 2) sample each functional group biomass correlated across space and time under three different randomized survey designs, and 3) obtain and compare biomass estimates at various sampling effort levels. With this framework, we aimed to balance accuracy, simplicity, generality, and computational feasibility. This study contributes to resilience-based management in an era of uncertain support for field surveys.

Methods

Fishery-independent data

The most reliable estimates of marine fish biomasses and spatial distributions at broad scales are generally derived from fishery-independent surveys. The in situ biomass estimates for approximately 300 reef fish taxa in the main Hawaiian Islands (MHI) region were collected through the National Coral Reef Monitoring Program (NCRMP) stratified random stationary point count (SPC) surveys led by the NOAA Pacific Islands Fisheries Science Center (PIFSC). The statistically randomized SPC survey used in this study encompasses 10 years (2010, 2012–2013, 2015–2016, 2019; NCEI Accession 0162472, 0157591, 0210958, 0157589, 0157590, 0211063, 0210958) and 1,798 survey locations across 7 islands (Niihau, Kauai, Oahu, Molokai, Lani, Maui, and Hawaii). The detailed descriptions of these data sets can be found in Heenan et al. (2017), McCoy et al. (2019), and Suarez and Grabowski (2021). Briefly, these fishery-independent surveys were based on a stratified random sampling design using the paired-diver SPC method (Heenan et al. 2017). The stationary diver-based sampling method records fish species, size, and abundance in visually estimated paired stationary 15-m diameter survey cylinders (353 m^2) extending from the seafloor to the surface. These fishery-independent surveys have recorded approximately 2 million fishes and provide site-level biomass records across a range of fish species and functional groups. Fishes were categorized into four functional groups (primary, secondary, planktivores, and herbivores; (Friedlander et al. 2018; McCoy et al. 2019)), for subsequent statistical and power analyses. The "primary" group consists of primary consumers such as herbivores and detritivores, while the "secondary" group are secondary consumers that represent omnivores and benthic invertivores.

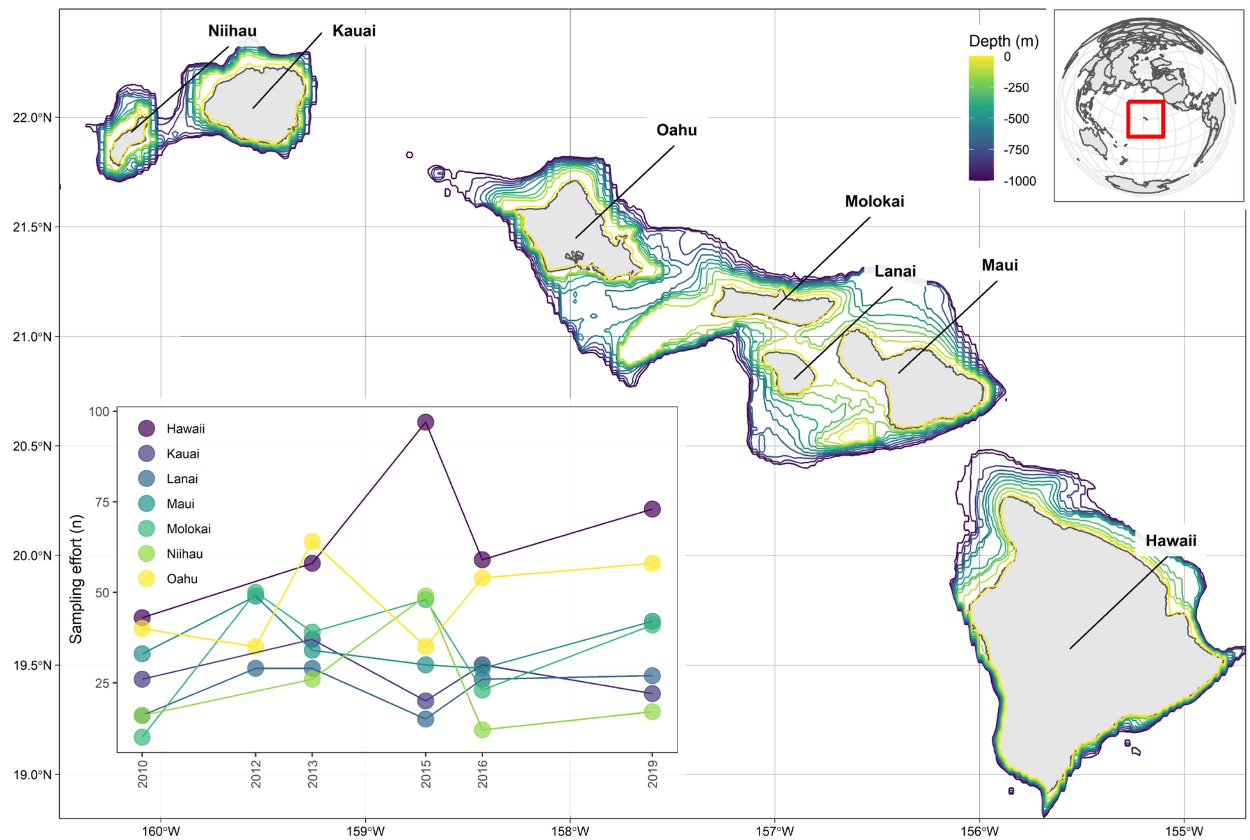

Figure 1. Study region encompassing the main Hawaiian Islands and seven islands featured in this study. The study area ranges from 15 to 100 m in depth, in latitude from 18.9° to 22.3,° and in longitude from-160.5° to -154.8°. Bathymetry lines represent 100-m contour lines. Bathymetric data were drawn from ETOPO1 1 arc-minute global relief model (Amante and Eakins 2009). The inset panel shows the National Coral Reef Monitoring Program sampling efforts (number of surveyed sites) between 2010 and 2019.

Spatiotemporal modeling of functional group biomass

We developed a statistical model to derive spatiotemporal estimates of four functional group biomass (primary, secondary, planktivore, and piscivore) that serves as the putative true survey populations within our simulation framework. We applied a mixed-modeling approach (Zuur et al. 2009) which incorporates a spatially explicit temporal trend (i.e., local trend) alongside spatial (temporally constant) and spatiotemporal (time-varying) components. It imposes correlation across space and time in the estimates of functional group biomasses and spatiotemporal distributions across the 7 islands within the MHI region. Accounting for spatial autocorrelation between spatially referenced observations proximate in both space and time can facilitate subsequent survey simulations with sufficient biogeographical accuracy for evaluating the efficacy of various sampling strategies.

Using the R *sdmTMB* package (Anderson et al. 2022), we fitted a full spatiotemporally explicit generalized linear mixed model (GLMM) with a local trend to each size-aggregated functional group biomass (piscivore, planktivore, primary, secondary). The *sdmTMB* package provides a flexible mixed modeling framework that incorporates an automatic differentiation platform, which fits models by finding the minimum log likelihood based on the *nlminb* optimization routine (Kristensen et al. 2015). We included both spatial and spatiotemporal components. Depth was modeled as a quadratic effect (Barnett et al. 2021), and year was included as a factor. The full model incorporating the local trend was fitted with a Tweedie distribution and a log link (Tweedie 1984). This setting has been shown to perform well with zero-inflated data (Barnett et al. 2021; Tanaka et al. 2018). The full model can be written as:

$$y_{s,t} \sim Tweedie(\mu_{s,t}, \rho, \Phi), 1 < \rho < 2,$$

$$\mu_{s,t} = exp(\alpha_t + \beta_1 D_{s,t} + \beta_2 D^2{}_{s,t} + \omega_s + \epsilon_{s,t} + \zeta_s t),$$

$$\omega \sim MVNormal\left(0, \sum\nolimits_{\omega}\right),$$

$$\epsilon_t \sim MVNormal\left(0, \sum\nolimits_{\varepsilon}\right),$$

$$\zeta \sim MVNormal\left(0, \sum\nolimits_{\zeta}\right),$$

Equation 1

where $y_{s,t}$ is functional biomass (g per m^2) at location s and time t, $\mu_{s,t}$ is the mean functional biomass at location *s* and time *t*, ρ is the Tweedie power parameter, and φ is the dispersion parameter. The α_t is estimated independently for each year. The β_1 and β_2 are two depth covariates included as log-transformed depth and log-transformed depth squared. ω_s and $\epsilon_{s,t}$ are spatial and spatiotemporal random effects, respectively, derived from Gaussian Markov random fields (Cressie and Wikle 2015) with respective covariance matrices $\sum_{\varepsilon}$ and $\sum_{\omega}$. The $\zeta_s t$ are the spatially varying coefficients that capture local trends through time (i.e., 2010-2019), also derived from Gaussian Markov random fields. Time *t* (i.e., year) is incorporated after multiplying with ζ_s and centered by its mean value. All random fields incorporate covariance matrices constrained by anisotropic Matérn covariance functions with independent scales but share the same κ parameters controlling the decay rate of spatial correlation as a function of distance (Cressie and Wikle 2015; Thorson 2019). All functional biomass models were fitted with the same covariates (year and two depth covariates) (Barnett et al. 2021). Using the R *INLA* package (Arab 2015), the continuous random fields with triangulated mesh were prepared with vertices at 500 knots (Fig S2). Conventional diagnostic plots (e.g., quantile-quantile plots; Fig S3) and spatial patterns in residuals were examined to analyze model fits. We predicted functional biomass at each grid location defined by NOAA Coastal Relief Model vol.10 bathymetry data (3 arc seconds, ~90m) to develop a smooth surface of functional biomass estimates across the MHI region. Spatiotemporal-GLMM outputs were spatially aggregated at the resolution of the NOAA CRM vol. 10 bathymetry grid (~8,100 m^2), which is the spatial resolution used for subsequent survey simulations.

Simulation design

The final steps in the analysis were (1) to conduct realistic simulations of fisheries-independent SPC surveys and (2) to evaluate the efficacy of alternate sampling protocols by assessing the deviation of stratified estimates of biomass from the true biomass available to the SPC survey. We developed a simple simulation framework using the R *SimSurvey* package (Regular et al. 2020) to run a series of surveys and stratified analyses over the reconstructed functional group biomasses that varied in space and time. These modeled biomasses were considered true populations for applying three alternative sampling protocols, and the stratified analyses provided estimates of total biomass for each functional group.

The first sampling protocol (hereafter "traditional survey") simulates the existing geographically comprehensive stratified random design as part of NCRMP. The traditional survey domain includes all hard-bottomed reef habitat within 1–30 m depth and is stratified into three depth zones (shallow [1–6 m], moderate [6–18 m], and deep [18–30 m]), three reef zones (forereef, backreef, and lagoon), and designated island sectors. The number of sites in each stratum is allocated based on the area of strata and the observed variance of fish biomass. **The second sampling protocol (hereafter "zone-based")** follows the survey domain defined by ecologically homogeneous strata proposed by Oliver et al. (2020). Briefly, these ecologically homogeneous strata are based on statistically downscaled ecological data from the historical NCRMP Pacific Reef Assessment and Monitoring Program surveys. Oliver et al. (2020) developed a statistical framework, which used contiguous clustering was used to identify "natural" divisions between observations while maintaining the optimal balance between fine spatial resolution and statistical robustness of the identified clusters. **The third sampling protocol (hereafter "zone-triaged")** also follows the ecologically homogeneous strata proposed using methods from Oliver et al. (2020) but mimics a hypothetical resource-constrained situation where survey protocols randomly drop ⅓ of total strata at each island (i.e., every stratum from an island is assigned a 2/3 probability of being surveyed).

The sampling sites in each stratum are allocated based on the stratum's proportional area and the variance structure of the observed reef fish populations within the stratum. The allocated number of cells is randomly selected in each stratum. The biomass of fish "recorded" at each sampling site is calculated by applying binomial sampling of the estimated fish biomass in each sampled cell (~8,100 m^2) by the proportion of the area covered by the survey area (353 m^2 from two 15-m diameter survey cylinders) and the simplified detectability defined as 0.5 (i.e., catchability, Regular et al. 2020). We tested survey performance using 50 levels of sampling efforts, where the simulated sampling effort started at 10 sampling sites per island and increased to 500 sampling sites in intervals of 10 sites for computational efficiency. The simulated survey efforts capture the range of observed survey efforts from past NCRMP SPC survey data (2013, 2016, and 2019; Table 1).

Table 1. Summary of historical National Coral Reef Monitoring Program sampling efforts (allocated number of sites) by islands based on surveys conducted in 2013, 2016, and 2019. High: 75th percentile: Low: 25th percentile.

	High (75th percentile)	Median	Low (25th percentile)
Hawaii	115.5	102	98
Kauai	71.5	50	44.5
Lanai	48	46	45.5
Maui	61	48	46
Molokai	66	65	52
Niihau	36.5	28	23.5
Oahu	105.5	100	98.5

We used root-mean-squared error (RMSE) to measure the precision and bias of the biomass estimates from each survey (Regular et al. 2020). The combinations of survey settings are based on four functional groups, seven islands, ten sampling efforts, and three survey designs. Every individual survey setting was simulated ten times, resulting in 420,000 simulations to evaluate the overall survey performances. All analyses were performed in the R programming environment (ver. 4.0.1.; R Core Team 2021s; <www.r-project.org>). Reproducible R scripts and data can be found in https://github.com/krtanaka/ncrmp_power_analysis. Original survey data can be found at https://www.fisheries.noaa.gov/inport/item/24447

Results

Spatiotemporal dynamics of MHI functional reef fish biomass distributions

Predictions of the spatially explicit temporal trend from the spatiotemporal model show fine-scale spatial structures in rates of changes of functional group biomass across the MHI region (Fig. 2). Analysis of the predicted localized trends (slope of biomass across years) and the median biomass revealed several geographic patterns that are difficult to detect at broader scale such as increases in reef fish biomass across all functional groups around Oahu (Fig. 2a) and relatively higher predicted biomass in Niihau over 2010-2019 (Fig. 2b). The predicted functional group biomass varied from 0.1 to 92.4 g per m^2 during 2010-2019 (Fig. 3). It is important to note that the distributions of these functional group biomasses extend into regions deeper than are surveyed, and findings from these analyses only describe the dynamics of their biomass distribution within the MHI NCRMP survey domain. The lowest mean biomass was among the planktivore, and the highest was the primary group species (Fig. 3).

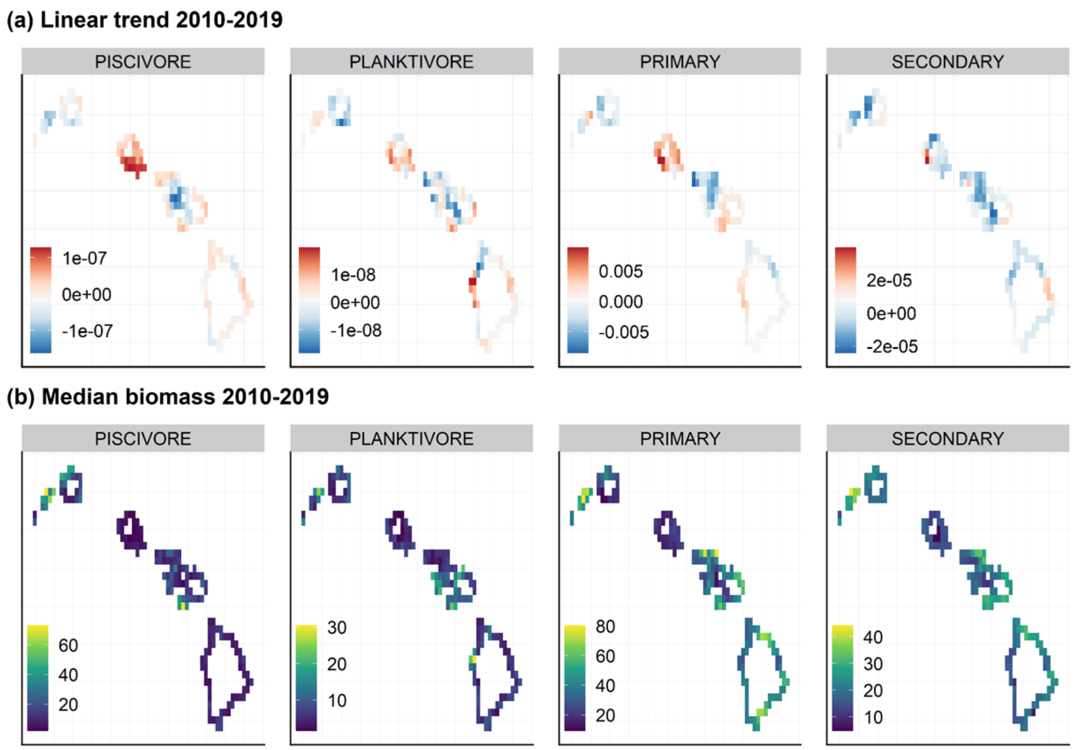

Figure 2. Model predicted spatial and temporal patterns of reef fish biomass among functional levels across the main Hawaiian Islands from 2010 to 2019; (a) predicted local trends (slope coefficients; year^{-1}), (b) spatial distribution of median predicted biomass (g per m^2). Model predictions were originally at 3 arc seconds (~90 m) but aggregated to 0.1 decimal degrees.

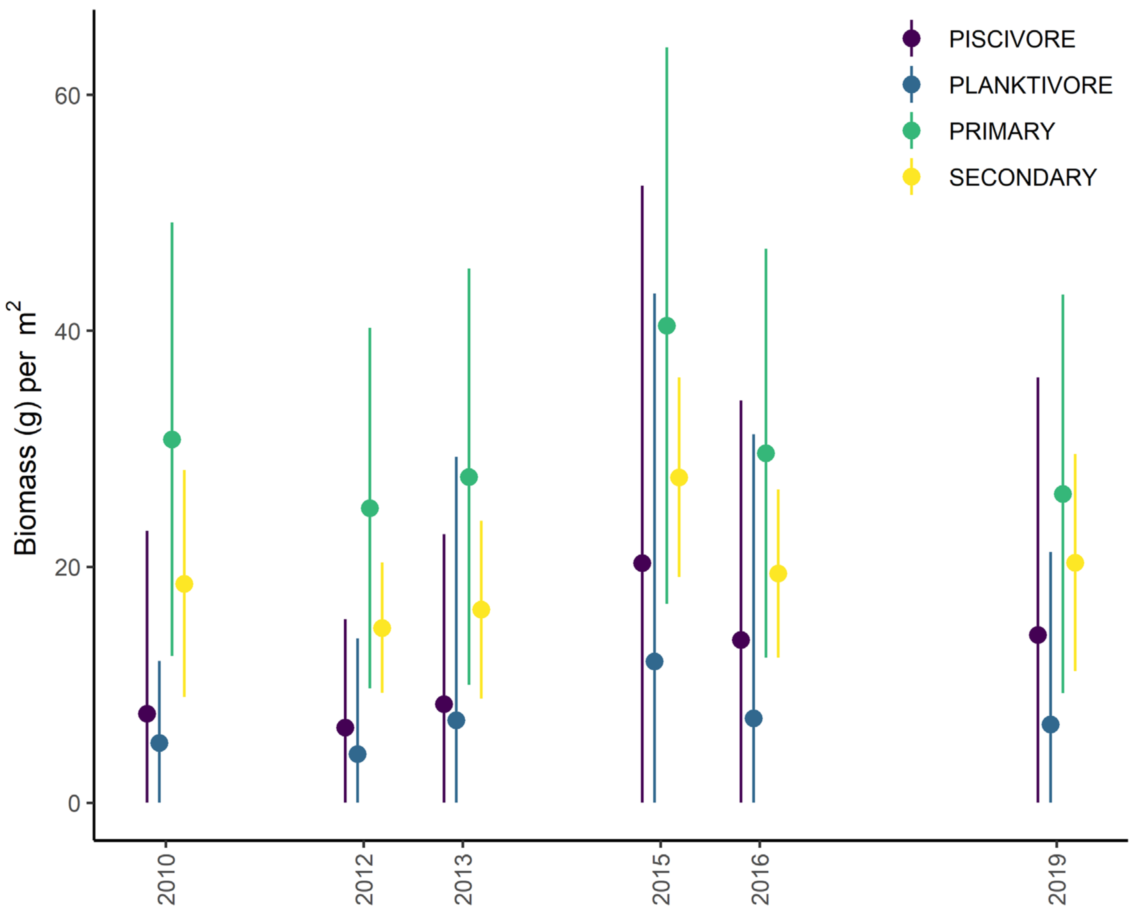

Figure 3. The time series of median predicted functional reef fish biomass with 95% confidence intervals across the main Hawaiian Islands.

Simulation testing

We first provide a summary of results from a single combination of simulation settings. This simulation quantified the efficacy of three survey designs (traditional, zone-based, and zone-triaged) by assessing the deviation of stratified estimates of piscivore biomass around Oahu from 2010 through 2019 by applying the median historical sampling effort over this period (number of sampling sites set at 100; Table 1) (Fig. 4, Fig. 5). All survey designs captured the temporal trends in piscivore biomass reasonably well (Fig. 5). Based on the RMSE, the traditional survey provided the most accurate biomass estimates (RMSE = 8.4^10^7). The effect of heterogeneous spatial allocation of sampling sites on biomass estimates is clear in the zone-triaged survey design, where systematic bias appears to be a problem with stratified biomass estimates that were consistently biased low across all years (RMSE = 1.6×10^8). Traditional (RMSE = 8.4×10^7) and zone-based (RMSE = 9.7×10^7) survey designs, in contrast, exhibited relatively smaller biases.

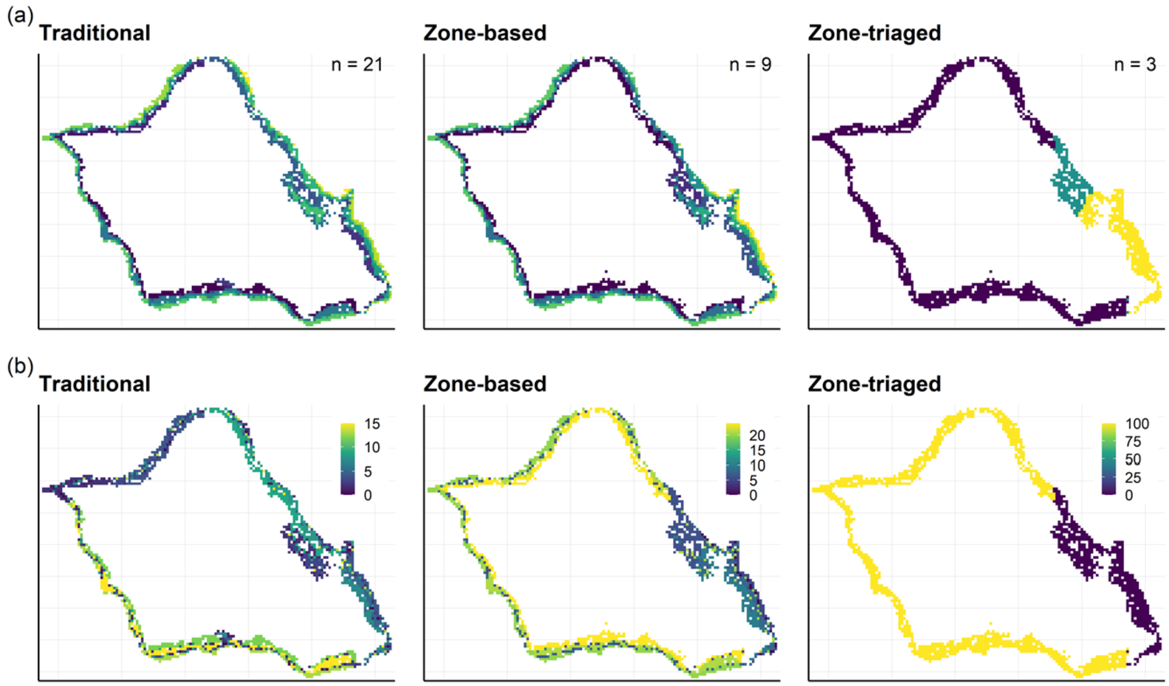

Figure 4. Three survey designs in Oahu. (a) Number of strata under each sampling design, (b) Spatial allocations of sampling sites under the median sampling efforts (n = 100: see Table 1). The color legend in panel (b) represents number of sites per strata.

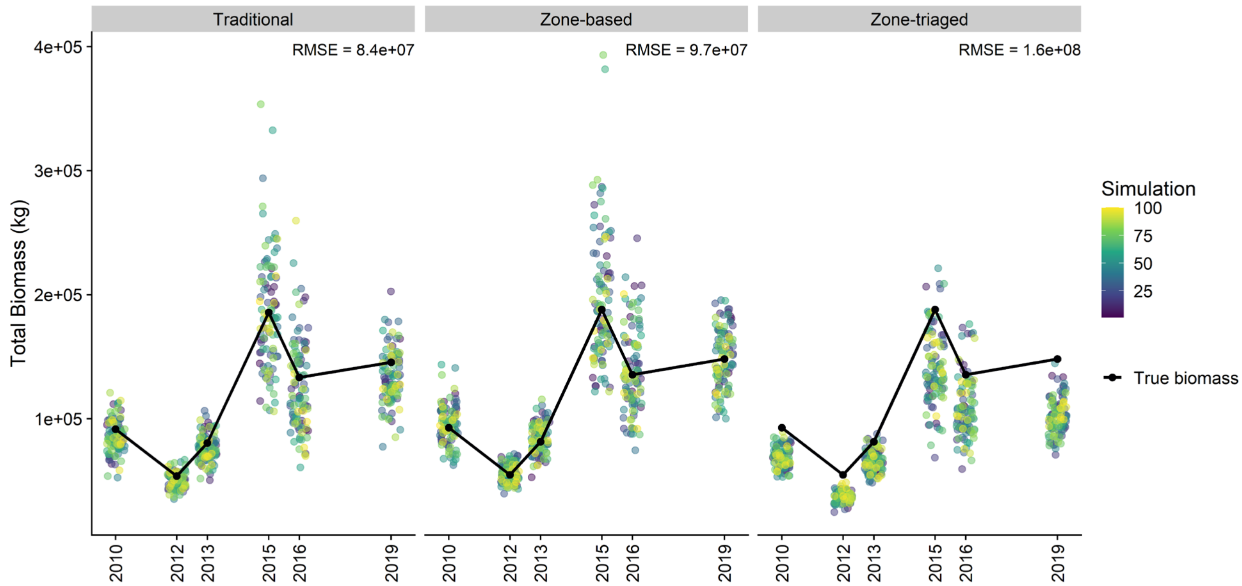

Figure 5. Example results of 100 simulated survey efforts on piscivore reef fish biomass (2010–2019) under three survey designs from Oahu. Thick solid lines in panel b represent true biomass trends.

To compare simulation results, we standardized RMSE across 7 islands and 4 functional reef fish groups by calculating z-scores (i.e., each island * functional group result has equal means and SDs but a different range). The full results with 4.2×10^5 simulations show clear declines in standardized RMSE in biomass estimates at higher sampling efforts (Fig. 6). Increasing the sampling efforts above 100 sites per island results in negligible RMSE differences between the three survey designs. Further increases in sampling efforts result in diminishing improvement in survey performance. The zone-triaged design outperformed the other survey designs at lower sampling efforts than historical averages (e.g., 2013, 2016, 2019; Fig. 6). This result indicates that measuring fewer ecologically homogeneous strata is more beneficial than measuring all strata when survey resources are limited (e.g., days allocated for each island).

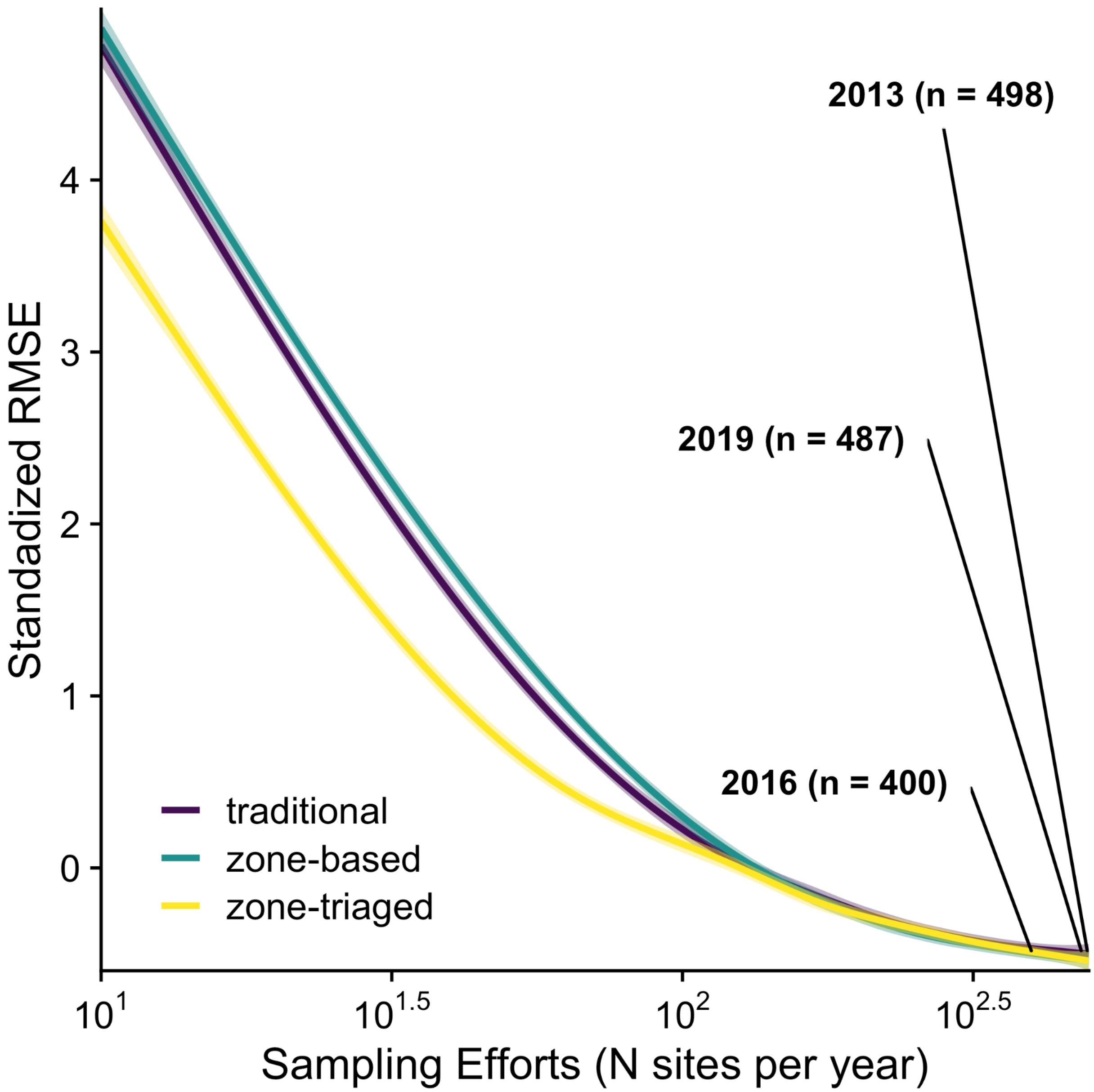

Figure 6. Changes in the standardized precision and accuracy (root-mean-squared error; RMSE) from an array of surveys with different island-level sampling protocols (traditional, zone-based, and zone-triaged). The x-axis represents the hypothetical stratified sampling efforts that varied from 10 to 500 sites per year. Note that the x-axis has been log10 transformed.

Comparison of individual simulation results (7 islands * 4 functional groups) did not identify any consistently "winning" survey designs as survey performance varied among islands and functional groups (Fig. 7, Fig. 8). Surveys from Lanai showed the best overall performance (mean RMSE $2.06\times10^7 \pm$ SD2.38×10^7), while surveys targeting piscovores returned the lowest RMSE overall (mean RMSE $3.03\times10^7 \pm$ SD 2.88×10^7). Among the survey designs targeted toward piscivores, the zone-triaged survey design outperformed the other designs in Hawaii, Kauai, Maui, Niihau, and Oahu. Traditional and zone-based survey designs showed the best performance in Lanai and Molokai, respectively.

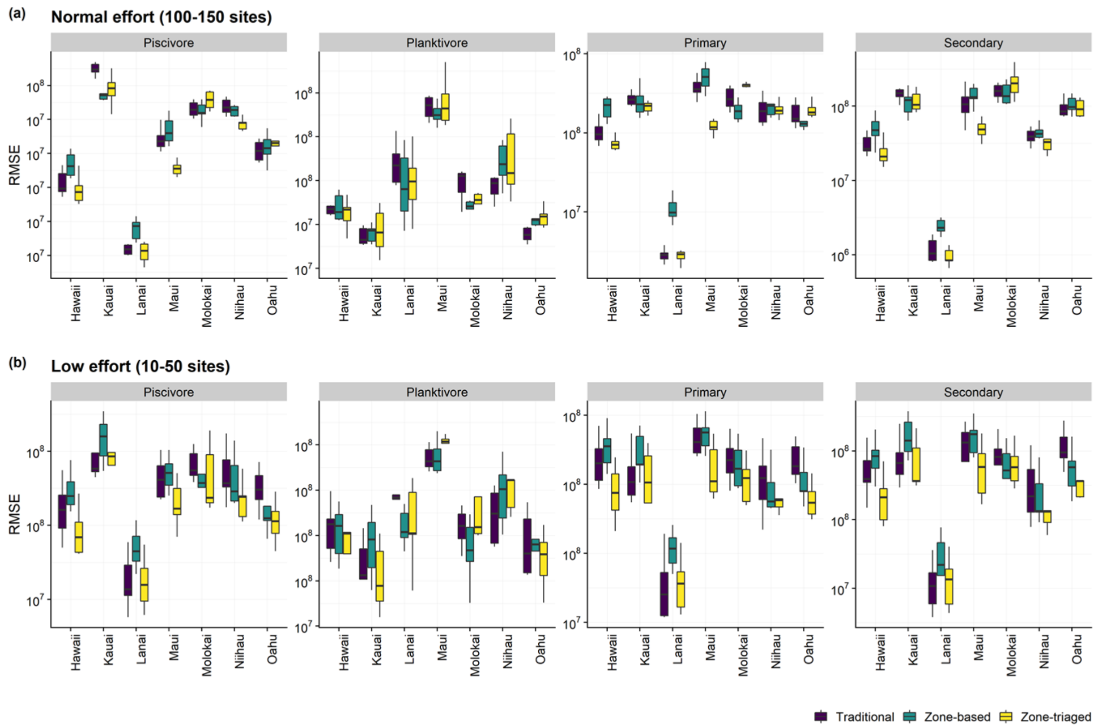

Figure 7. Divergent boxplots of the precision and accuracy (root-mean-squared error; RMSE) of stratified four functional group biomass estimates under three alternative sampling protocols and two different sampling efforts (2010–2019).

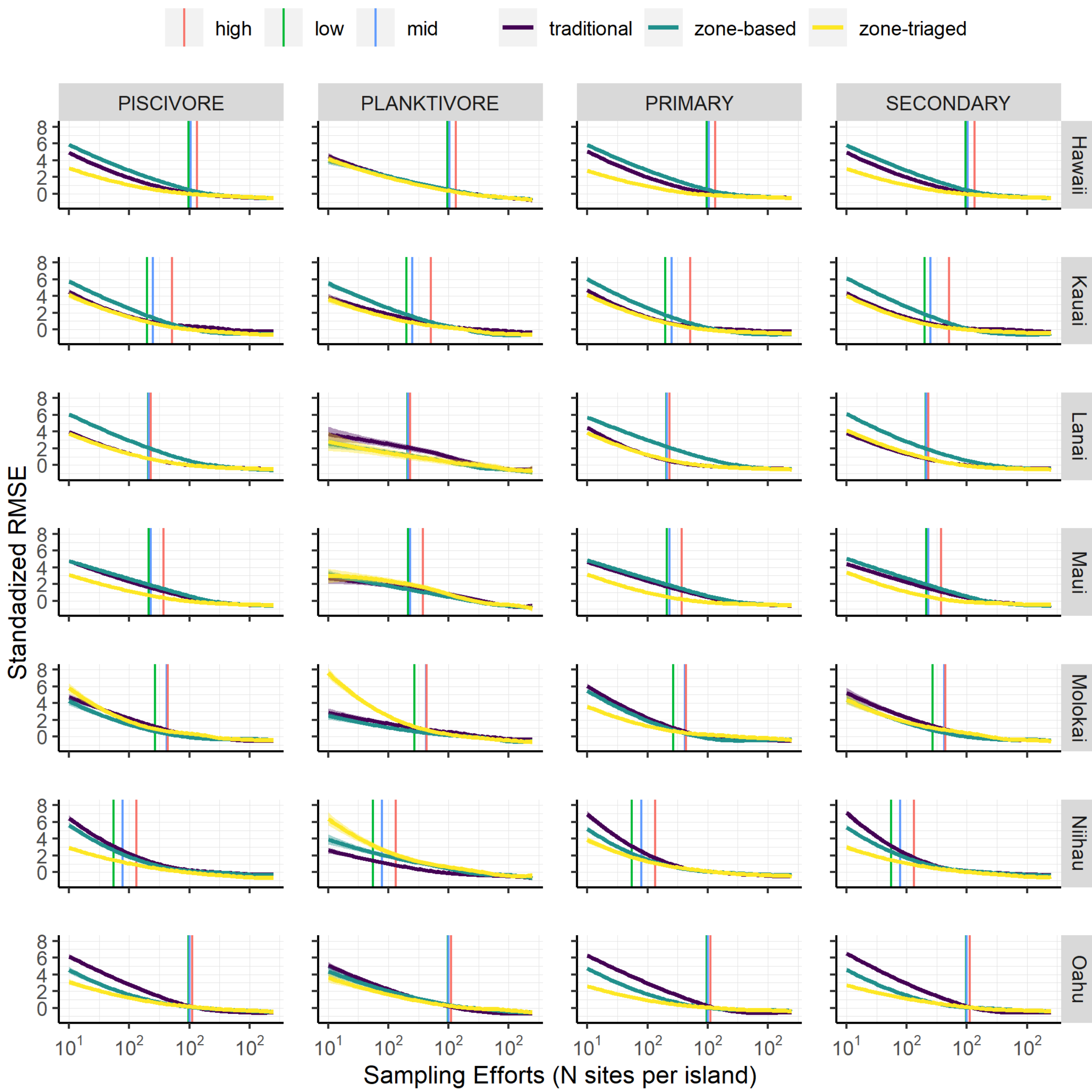

Figure 8. Changes in the precision and accuracy (root-mean-squared error; RMSE) standardized over an array of survey designs across seven islands and four functional groups. Note the x-axis has been log10 transformed. Vertical lines represent actual levels of historical sampling, ranked from "low" to "median" to "high."

Discussion

The NOAA's National Coral Reef Monitoring Program (NCRMP) survey maintains long-term monitoring efforts with a geographically comprehensive sampling scheme for detecting temporal and spatial changes in living marine resource (LMR) status from either natural or anthropogenic causes (Heenan et al. 2017). We developed a simulation approach using historic NCRMP data to evaluate the precision and accuracy of multiple fishery-independent survey designs to meet NCRMP objectives. We compared the performances of the sampling designs with different stratification schemes for the spatiotemporal distribution of reef fish biomass among functional groups, explicitly testing traditional NCRMP scenarios against more ecologically homogeneous survey sectors. Though somewhat narrow in scope, the mixed-effects spatiotemporal modeling of reef functional populations incorporated sufficient biogeographical realities to evaluate the efficacy of these distinct sampling strategies.

Overall, our results suggest that, assuming selected zones are broadly representative of management needs, targeting fewer ecologically homogeneous and meaningful zones (zone-triaged design) can be a viable alternative when survey resources are scarce (e.g., reduced number of days allocated for sampling each island). At low levels of sampling, zone-triaged methods substantially out-performed the traditional sector stratification (Fig. 6). Among the survey designs targeted at piscivores (the functional group with the lowest RMSE overall), the zone-triaged survey design emerged as the best performing survey strategy in 5 out of the 7 islands (Fig. 7, Fig. 8). However, the relative strength of either the zone-triaged or zone-based designs tends to disappear with greater sampling effort, and the performances of the different designs are hard to distinguish when sampling efforts reach the levels traditionally required to meet NCRMP historical survey standards (Fig. 6; Table 1).

While the objective of the NCRMP survey is to apply a “wide-but-thin” strategy to cover shallow benthic and reef fish habitats in the most geographically comprehensive way possible, the enormous scale of NCRMP’s survey domain provides many logistical challenges to the maintenance of adequate, random, and well-balanced representation across space, time, and stratified variables within a sector. For example, this “wide-but-thin” survey strategy requires practitioners to cover a relatively large number of strata (i.e., reporting sectors), and is vulnerable to unexpected effort reductions due to funding shortfalls, vessel unavailability, weather, and other complications that require immediate or strategic revisions. However, many survey practitioners and resource managers are skeptical of hypothesized changes to existing survey designs until the potential benefits of alternative designs have been empirically demonstrated through a simulation from existing data. Our results imply that 1) ecological inference from the actual survey can be potentially improved by altering the survey designs and 2) valuable survey power can be potentially maintained by cutting back on the number of strata. Both of these findings may improve the relevance of NCRMP data products and support resilience-based management in an era of uncertain field access. In other words, in times of limited opportunities for survey effort, adopting a zone-triaged design appears to provide a better estimate than traditional methods. However, at historic levels of sampling, the distinctions among designs are small enough that it would be hard to justify a redesign around the zone-triaged strategy.

Our study demonstrates the benefit of evaluating alternate stratification schemes for complex surveys with multiple objectives. In theory, survey practitioners are interested in designing a survey that can effectively address multiple objectives such as describing spatiotemporal dynamics of wildlife populations, predicting abundances or impact assessment at unsampled locations, or accurately estimating autocorrelation model parameters (Bijleveld et al., 2012). In reality, each objective can have a different optimal survey design and sampling protocols, as no single design would fulfill all criteria. Any successful design must reasonably compromise among these objectives. Fisheries-independent LMR surveys often incorporate multiple sampling schemes, and marine populations often show structured spatiotemporal dynamics (Regular et al., 2020). A single survey design does not easily capture these complex and dynamic features, making it difficult to determine optimal sampling and subsampling schemes (Bijleveld et al. 2012). In such situations, it is common for practitioners to resort to simulations to identify a survey design that minimizes the number of sampling units and maximizes the accuracy of the estimates (Cao et al. 2017; Li et al. 2015; Regular et al. 2020).

Statistical power analyses that can provide quantitative comparisons of alternate survey strategies and performances are critical in designing effective LMR monitoring programs. For a given survey, the ability to accurately detect ecological trends depends on careful planning and well-articulated and realistic objectives (Nichols & Williams, 2006; Yoccoz et al., 2001). Ideally, survey practitioners should routinely conduct benchmark survey evaluations to ensure reasonable monitoring objectives and evaluate trade-offs in sampling effort (Field et al. 2005; Rhodes et al. 2006). Comparing multiple survey designs can provide an opportunity to evaluate the prioritization of existing monitoring tasks by exploring alternative sampling protocols for potential gains in survey efficiencies. For example, survey practitioners can develop comprehensive simulation studies on cost-benefit analysis by incorporating the various subcomponents inherent in survey design and metric calculations (e.g., minimum and maximum survey effort reduction related to reducing the number of biological samples, days at sea, survey areas, or survey frequency). The simulated impact of certain changes in survey protocols can be used to assess total survey uncertainty and provide insight into the appropriate weighting in subsequent likelihood-based assessments of survey products (e.g., spatial availability, density dependence, diver effects, timing, and environmental conditions). Future power analysis platforms should incorporate environmental data in model-based LMR abundance estimation and include functions to evaluate multispecies/multi-objective trade-offs. Our findings contribute to existing NCRMP survey efforts and provide an opportunity to develop quantitative applications to determine the impacts of different monitoring strategies in terms of inputs (efforts) and outputs (accuracy) that can be used for similar scientific survey evaluation frameworks.

Acknowledgments

The authors would like to acknowledge funding from NOAA's Coral Reef Conservation Program, and the useful contributions of Hannah Barkley, Tye Kindinger, and Courtney Couch.

Literature Cited

Amante C, Eakins BW. 2009. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24, National Geophysial Data Center, Marine Geology and Geophysics Division, Boulder, CO, USA. 19 pp.

Anderson SC, Ward E J, English PA, Barnett LAK. 2022. sdmTMB : an R package for fast , flexible , and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. BioRxiv. 1–17.

Arab A. 2015. Spatial and spatio-temporal models for modeling epidemiological data with excess zeros. Int J Environ Res Public Health. 12(9):10536–10548. https://doi.org/10.3390/ijerph120910536

Barnett LAK, Ward EJ., Anderson SC. 2021. Improving estimates of species distribution change by incorporating local trends. Ecography. 44(3): 427–439. https://doi.org/10.1111/ecog.05176

Bijleveld AI, van Gils JA, van der Meer J, Dekinga A, Kraan C, van der Veer HW, Piersma T. 2012. Designing a benthic monitoring programme with multiple conflicting objectives. Methods Ecol Evol. 3(3):526–536.

Cao J, Thorson JT, Richards RA, Chen Y. 2017. Spatio-temporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. CanJ Fish Aquat Sci. 74(11):1781–1793. https://doi.org/10.1139/cjfas-2016-0137

Cressie N, Wikle CK. 2015. Statistics for spatio-temporal data. Hoboken, NJ: John Wiley & Sons.

Friedlander AM, Donovan MK, Stamoulis KA, Williams ID, Brown EK, Conklin EJ, DeMartini EE, Rodgers KS, Sparks RT, Walsh WJ. 2018. Human-induced gradients of reef fish declines in the Hawaiian Archipelago viewed through the lens of traditional management boundaries. Aquat Conserv Mar FreshwEcosyst.28(1):146–157. https://doi.org/10.1002/aqc.2832

Gunderson DR. 1993. Surveys of fisheries resources. New York, NY: John Wiley & Sons.

Heenan A, Williams ID, Acoba T, DesRochers A, Kosaki RK, Kanemura T, Nadon MO, Brainard RE. 2017. Data Descriptor: Long-term monitoring of coral reef fish assemblages in the Western central pacific. Sci Data. 4:1–13. https://doi.org/10.1038/sdata.2017.176

International Council for the Exploration of the Sea (ICES). (2020). ICES Workshop on unavoidable survey effort reduction (WKUSER). ICES Sci Rep. 2(72):72.

Katsanevakis S, Weber A, Pipitone C, Leopold M, Cronin M, Scheidat M, Doyle TK, Buhl-Mortensen L, Buhl-Mortensen P, Anna GD. 2012. Monitoring marine populations and communities: methods dealing with imperfect detectability. Aquat Biol. 16(1):31–52.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. (2015). TMB: automatic differentiation and Laplace approximation. *ArXiv Preprint ArXiv:1509.00660*.

Li B, Cao J, Chang, J, Wilson C, and Chen Y. 2015. Evaluation of Effectiveness of Fixed-Station Sampling for Monitoring American Lobster Settlement. N Am J Fish Manag. 35,: 942–957. https://doi.org/10.1080/02755947.2015.1074961

McCoy K, Asher J, Ayotte P, Gray A, Kindinger T, Williams I. 2019. Pacific Reef Assessment and Monitoring Program Data Report - Ecological monitoring 2019 - Reef fishes and benthic habitats of the main Hawaiian Islands: PIFSC data report DR-19-039. https://doi.org/10.25923/he4m-6n68

Murphy HM, Jenkins G P. 2010. Observational methods used in marine spatial monitoring of fishes and associated habitats: a review. MarFreshwRes. 61(2):236–252.

Nichols JD, Williams BK. 2006. Monitoring for conservation. Trends EcolEvol. 21(12): 668–673.

Oliver TA, Barkley H, Couch C, Williams I, Kindinger T. 2020. Downscaling Ecological Trends from the Spatially Randomized Datasets of the National Coral Reef Monitoring Program. July. https://doi.org/10.25923/2fef-8r42

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.r-project.org/

Regular PM, Robertson GJ, Lewis KP, Babyn J, Healey B, Mowbray F. 2020. SimSurvey: An R package for comparing the design and analysis of surveys by simulating spatially-correlated populations. PLoS ONE, 15(5), 1–28. https://doi.org/10.1371/journal.pone.0232822

Suarez B, Grabowski TB. 2021. Estimating detection and occupancy coefficients for the Pacific Islands coral reef fish species. Hawaii Cooperative Studies Unit Technical Report. USGS Publications Warehouse. http://pubs.er.usgs.gov/publication/70229376

Tanaka KR, Chang J-H, Xue Y, Li Z, Jacobson L, Chen Y. 2018. Mesoscale climatic impacts on the distribution of Homarus americanus in the US inshore Gulf of Maine. Can J Fish Aquat Sci..18(July):1–58.

Thorson JT. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. Fish Res. 210:143–161. https://doi.org/10.1016/j.fishres.2018.10.013

Tommasi D, Stock C, Hobday A, Methot R, Kaplan I, Eveson P, Holsman K, Miller T, Gaichas

S, Gehlen M, et al. 2017. Managing living marine resources in a dynamic environment: the role of seasonal to decadal climate forecasts. Prog Ocean.152:15–49. https://doi.org/10.1016/j.pocean.2016.12.011

Towle EK, Donovan EC, Kelsey H, Allen ME, Barkley H, Blondeau J, Brainard R E, Carew A, Couch CS, Dillard MK. et al. 2022. A National Status Report on United States Coral Reefs Based on 2012–2018 Dat From National Oceanic and Atmospheric Administration's National Coral Reef Monitoring Program. Front Mar Sci. 8(February). https://doi.org/10.3389/fmars.2021.812216

Tweedie MCK. 1984. An index which distinguishes between some important exponential families. Statistics: Applications and New Directions. In J. K. Ghosh & J. Roy (Eds.), Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (pp. 579–604). Indian Statistical Institute.

Tyre AJ, Tenhumberg B, Field SA, Niejalke D, Parris K, Possingham HP. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecol Appl. 13(6):1790–1801.

Yoccoz NG, Nichols JD, Boulinier T. 2001. Monitoring of biological diversity in space and time. Trends Ecol Evol. 16(8): 446–453.

Zimmermann F, Enberg K. 2017. Can less be more? Effects of reduced frequency of surveys and stock assessments. ICES J MarSci. 74(1):56–68.

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM. 2009. Mixed Effects Models and Extension in Ecology with R (M. Gail, K. Krickeberg, J. Samet, A. Tsiatis, & W. Wong (eds.)). Springer Science & Business Media.

Supplemental Figures

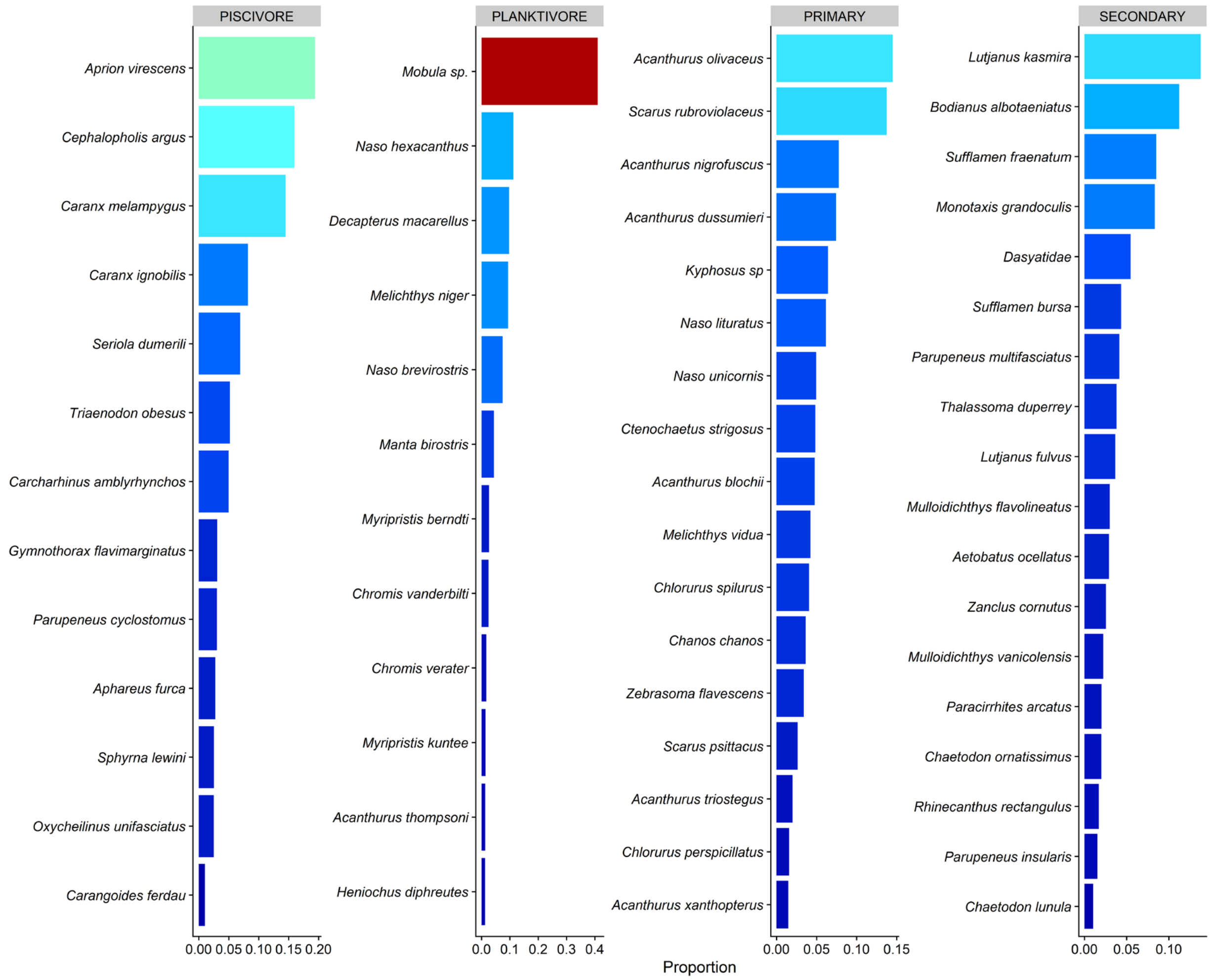

Figure S 1. Proportion of commonly detected species in functional fish biomass survey data between 2010 and 2019. Only species with proportion larger than 1% are included. Data were drawn from the National Coral Reef Monitoring Program (www.coris.noaa.gov/monitoring/).

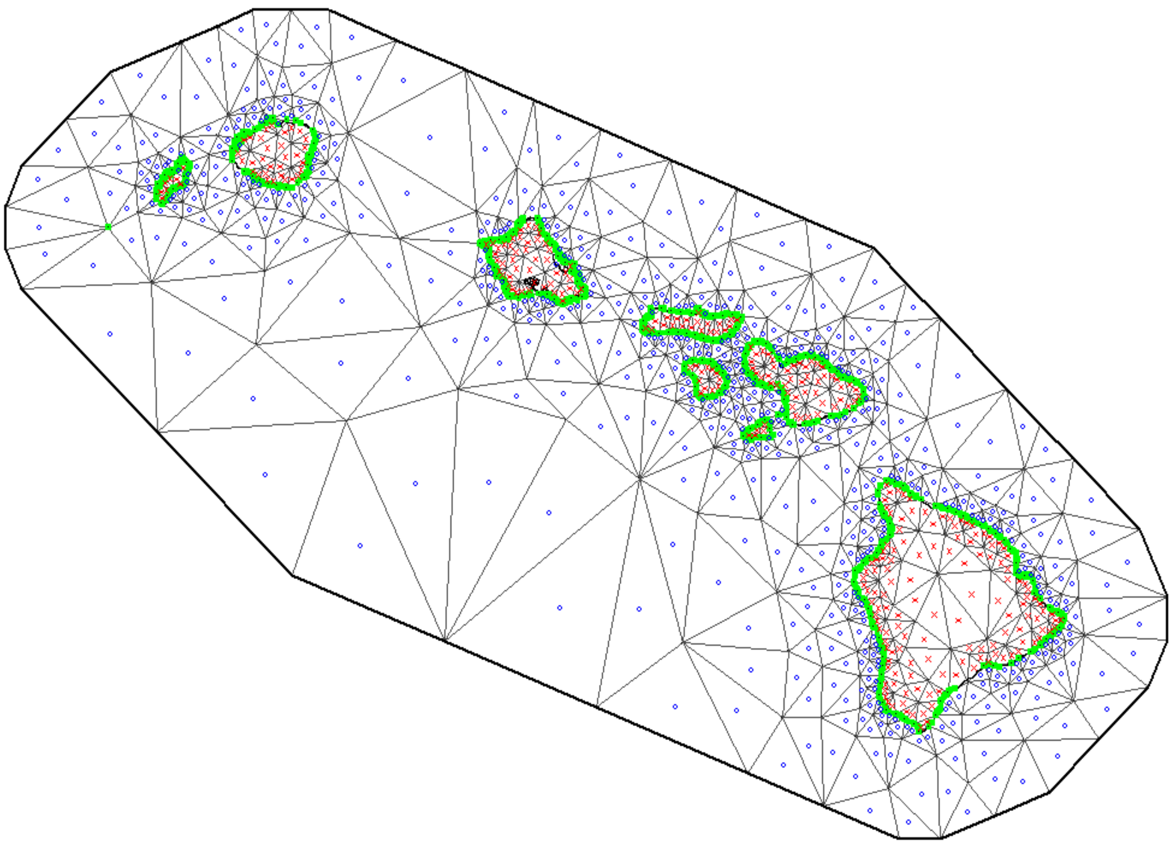

Figure S 2. Triangulated mesh covering the main Hawaiian Islands region prepared with vertices at 500 knots using the R INLA package. Polygons with blue points represent spatial domains considered for spatial autocorrelations in the spatiotemporal generalized linear mixed model calibration process. Polygons with red dots represent land masses and were not included in the spatial autocorrelations in the spatiotemporal generalized linear mixed model calibration process.

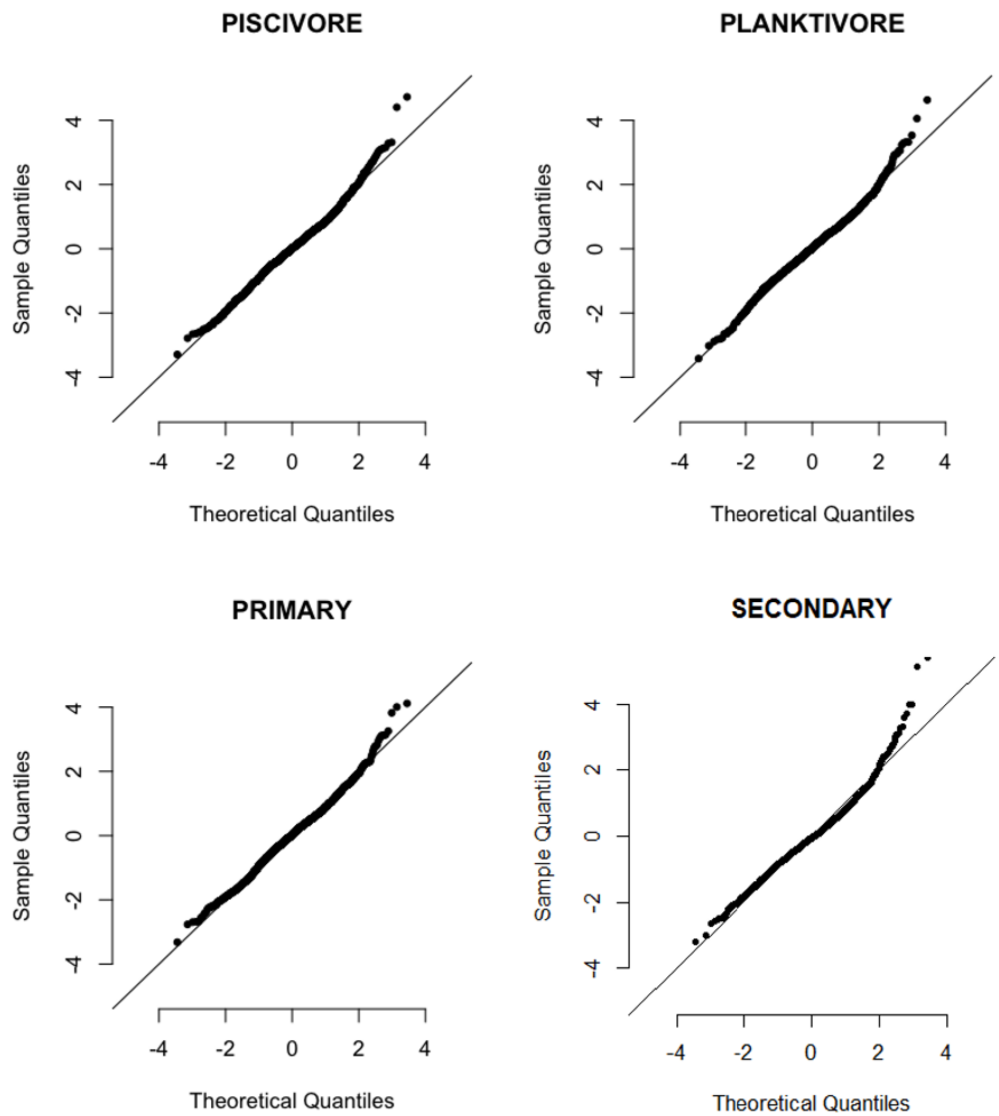

Figure S 3. Quantile–quantile plots of the deviance residuals of the spatiotemporal models fitting four National Coral Reef Monitoring Program reef fish functional group biomass data using Tweedie error distribution. secondary: secondary consumers (omnivores and invertivores) primary: primary consumers (herbivores).