



**NOAA
FISHERIES**



Comparing Coral Demographic Surveys From In Situ Observations and Structure-From-Motion Photogrammetry on High Diversity Reefs Shows Low Methodological Bias but Highlights Persistent Areas of Concern

NOAA Technical Memorandum NMFS-PIFSC-153
Pacific Islands Fisheries Science Center
National Oceanographic and Atmospheric Administration
U.S. Department of Commerce
December 2023

Comparing Coral Demographic Surveys From In Situ Observations and Structure-From-Motion Photogrammetry on High Diversity Reefs Shows Low Methodological Bias but Highlights Persistent Areas of Concern

Jonathan A Charendoff^{1,2}, Courtney Couch^{1,2}, Thomas Oliver¹, Mia Lamirand^{1,2}, Corinne Amir^{1,2}, Isabelle Basden^{1,2}, Damaris Torres-Pulliza^{1,2}, Mollie Asbury³, Morgan Winston^{1,2,4}, Brittany Huntington^{1,2}, Frances Lichowski^{1,2}, Nicolas Osborn¹, Jon Ehrenberg³, Joao Garriques^{1,2}, Jennifer Samson¹

¹ Pacific Islands Fisheries Science Center,
National Marine Fisheries Service,
1845 Wasp Blvd
Honolulu, HI, USA

² Cooperative Institute for Marine and Atmospheric Research,
University of Hawai'i at Mānoa,
Honolulu, HI, USA

³ Hawai'i Institute of Marine Biology
University of Hawai'i at Mānoa
Kāne'ohe, Hawaii, USA

⁴ School of Geographical Sciences and Urban Planning
Arizona State University
Tempe, Arizona, USA

NOAA Technical Memorandum NMFS-PIFSC-153

December 2023



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

Charendoff J.A., Couch C, Oliver T, Lamirand M, Amir C, Basden I, Torres-Pulliza D, Asbury M, Winston M, Huntington B, Lichowski F, Osborn N, Ehrenberg J, Garriques J, Samson J. 2023. Comparing coral demographic surveys from in situ observations and structure-from-motion photogrammetry on high diversity reefs shows low methodological bias but highlights persistent areas of concern. U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-PIFSC-153,52 52 p.
doi:10.25923/8r7f-t404

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/>

Contents

List of Tables.....	5
List of Figures.....	6
Executive Summary	8
Introduction	1
Methods	3
In-water collection	3
SfM image collection	5
SfM model generation and data extraction	5
Data analysis	7
Results	10
How does methodological error compare to observer error?	10
Is there a methodological bias in coral demographic metrics at the site-level?.....	11
Adult density of total scleractinians.....	11
Juvenile density of total scleractinians.....	12
Adult colony diameter of total scleractinians.....	13
Average old and recent partial mortality of total scleractinians	14
Condition prevalence of total scleractinians	16
Density of dominant coral genera	18
Diversity of coral genera	20
ESA listed taxa	23
Discussion.....	25
Demographic metrics that show minimal to no bias	25
Demographic metrics that continue to show distinctions across methods	26
Conclusions & Recommendations	29
Acknowledgements	32
Literature Cited.....	33
Appendices	38
Appendix A. Table of statistical results for methodological bias. Bold indicates fixed effect is significant ($p < 0.05$).	38
Appendix B. Lookup table of genus codes used by Pacific NCRMP	39
Appendix C. Cost-Benefit Analysis	40

List of Tables

Table 1. Counts of ESA listed coral species at comparison sites. Other ESA listed coral species were not observed despite their suspected presence in the Mariana Archipelago and American Samoa regions. Bold rows indicate total counts for each species.	23
Table 2. Direction of bias and Midpoint Scaled Mean Absolute Error (MS-MAE) for Couch et al. (2021) and the present study for all transects. A decrease in MS-MAE indicates lower inter-observer variability.....	28

List of Figures

Figure 1. Map of geographic areas (in white) where ESD conducts coral reef monitoring surveys in the Pacific.....	1
Figure 2. Location of 131 survey sites (purple), with 19 repeated survey sites indicated in orange. Black outlines indicate sub-island sectors.	4
Figure 3. Graphic of benthic survey site with in-water visual survey segments (in gray). The SfM swim path is indicated by the black dashed line, covering a 3 m x 20 m area. Scale bar markers, also known as Ground Control Points (GCPs), are represented by the two rectangles that are connected by a thin line.	5
Figure 4. A) Full orthomosaic with all segments (red boxes) and colonies annotated (magenta lines). B) One fully annotated segment in the orthomosaic with a C) raw image containing a Pocillopora sp. colony.	7
Figure 5. Graphic depicting comparisons (A) between observers for the in-water observers (in-water divers), (B) between observers for the SfM imagery, and (C) between the methods of diver-derived data (in-water method) and SfM image-derived data (SfM method).....	8
Figure 7. Plot of the paired site-level SfM adult colony density vs. Diver adult colony density (points). Black line is 1:1 line, red dashed line is linear regression line; (B) Boxplot of adult density by method type; (C) Marginal effects plot of adult density by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (Method or method x depth) from the linear mixed effect models. NS: $p > 0.05$	11
Figure 8. (A) Plot of site-level SfM juvenile colony density vs. Diver juvenile colony density (points). Black line is 1:1 line, red dashed line is linear regression line; (B) Boxplot of juvenile density by method type; (C) Marginal effects plot of juvenile density by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (Method or method x depth) from the linear mixed effect models. NS: $p > 0.05$	12
Figure 9. (A) Plot of site-level SfM adult average maximum colony diameter vs. diver adult average maximum colony diameter (points). Black line is 1:1 line, red dashed line is linear regression line; (B) Boxplot of average maximum diameter by method type; (D) Marginal effects plot of average maximum diameter by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (Method or method x depth) from the linear mixed effect models. NS: $p > 0.05$	13

Figure 10. (A, D) Plot of site-level SfM adult average percent partial mortality vs. Diver adult average percent old partial mortality (points) for old and recent mortality, respectively. Black line is 1:1 line, red dashed line is linear regression line; (B, E) Boxplot of average mortality by method type for old and recent mortality, respectively; (C, F) Marginal effects plot of average old mortality by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models (LMMs). For plots B and C the results of the Likelihood Ratio Tests are included for the different fixed effects (Method, or Method x Depth) from the LMMs. NS: $p > 0.05$. For recent mortality, the difference between methods overall (E) were tested using a nonparametric Wilcoxon test. An outlier in percent recent dead ($>10\%$ recent dead) was removed from the plots to better visualize the majority of data presented. 14

Figure 11. (A) Plot of site-level SfM prevalence vs. Diver prevalence (points) where black line is 1:1 line, and red dashed line is linear regression line. (B) Boxplots of prevalence by method type. Significance ($\alpha = 0.05$) of method was tested using a nonparametric Wilcoxon Test. NS: $p > 0.05$, and (C) Plot of prevalence by maximum depth (m) standard error (shaded areas) for acute disease. (D, E, F) is predation, (G, H, I) is chronic disease, and (J, K, L) is bleaching prevalence. 16

Figure 12. Percent difference of adult colony density for adult coral genera. Only genera with at least 30 colonies observed across all sites for both methods are included. Color corresponds to overall abundance of each genera. Cryptic refers to genera that typically live in cracks or under overhangs and may not be visible in the orthomosaic. Nonparametric Wilcoxon tests were done for genera that had at least 30 colonies observed across both methods. Asterisks (*) indicate significant results ($p < 0.05$). Taxa that had fewer than 30 colonies were not analyzed statistically due to low sample size. See Appendix B for the genus code lookup table. 18

Figure 13. Percent difference of adult colony density for juvenile coral genera. Only genera with at least 8 colonies observed across all sites in-water are included. Color corresponds to overall abundance of each genera. Cryptic refers to genera that typically live in cracks or under overhangs and may not be visible in the orthomosaic. Nonparametric Wilcoxon tests were done for genera that had at least 8 colonies observed across both methods. Asterisks (*) indicate significant results ($p < 0.05$). Taxa that had fewer than 8 colonies were not analyzed statistically due to low sample size. See Appendix B for the genus code lookup table. 19

Figure 14. Adult coral diversity. (A, C, E) Plot of site-level SfM genus diversity vs. Diver genus diversity (points) where black line is 1:1 line, and red dashed line is linear regression line for Shannon-Wiener, Hill, and Richness, respectively. (B, D, F) Boxplots of corresponding diversity metrics. For plots B, D, and F, the result of the likelihood ratio test is included for the fixed effect (Method) from the Linear MMs. NS: $p > 0.05$ 20

Figure 15. Juvenile coral diversity. (A, C, E) Plot of site-level SfM genus diversity vs.

Diver genus diversity (points) where black line is 1:1 line, and red dashed line is linear regression line for Shannon-Wiener, Hill, and Richness, respectively. (B, D, F) Boxplots of corresponding diversity metrics. For plots B, D, and F, the result of the likelihood ratio test is included for the fixed effect (Method) from the linear mixed effect models. NS: $p > 0.05$ 21

Figure 16: Lumping-Splitting Challenge Visualized. Mean genus-level colony size (cm) and density between methods for top quartile of coral genera by mean density. Error-bars are standard errors of site-level means. Gray lines show the direction of lumping-splitting differences between methods. See Appendix B for genus code lookup table.. 22

Figure 17. Examples of ESA listed taxa captured during SfM: A) *Acropora globiceps*, B) *Isopora crateriformis*, and C) *Pavona cf. diffluens*. D) Close up of *Pavona cf. diffluens* colony surface. 24

Executive Summary

The Pacific Islands Fisheries Science Center's Ecosystem Sciences Division (ESD) executes the National Coral Reef Monitoring Program (NCRMP) across 40 primary islands, atolls, and shallow banks in the U.S. Pacific. The goal of NCRMP is to understand how reefs are responding, in both time and space, to local and global disturbances, and what factors enhance reef resilience.

Assessments of benthic cover and coral demography are key components of ESD's Pacific NCRMP and have historically been conducted using in-water visual surveys and photoquadrats by a team of three to four divers. With NOAA ship time becoming more limited and increasing complexities of field missions, Pacific NCRMP has been exploring innovative technologies to more efficiently monitor benthic communities. A primary focus of this effort has been Structure-from-Motion photogrammetry (SfM), which generates 3D reconstructions of a scene from overlapping images. In 2021, we published a study comparing data generated from in-water surveys to SfM-derived metrics for assessing coral demography in the low diversity reefs of the Main Hawaiian Islands (Couch et al., 2021). The goal of that study was to evaluate whether NCRMP can maintain continuity in our long-term NCRMP data sets if the program transitions to SfM survey methods. Couch et al. (2021) presented concerns about the comparability and continuity of estimates of partial mortality and bleaching prevalence between methods as well as high inter-observer error for both methods. As such, ESD modified existing protocols and training which improved consistency between observers in sizing colonies and estimating percent mortality (known as the "lumping-splitting challenge"). In the current study, we reevaluate the comparison on higher diversity reefs in the Mariana Archipelago and American Samoa.

The three objectives of this study:

- 1) Compare error between methods to within-method observer error,
- 2) Test for methodological bias between SfM and in-water visual surveys, and
- 3) Provide recommendations for the transition to SfM for future NCRMP benthic monitoring through a synthesis of Couch et al. (2021) and the present study.

In-water surveys and SfM imagery collection were conducted along the same transects at sites across a range of depths and reef types. Colony density, average colony length, average partial mortality, prevalence of disease, and compromised health states were recorded for each survey and method type.

To compare between-method variability and between-observer variability, multiple divers and SfM annotators surveyed the same 19 transects, covering 32 juvenile and 23 adult segments (i.e. quadrats). For all demographic metrics at this small set of sites, the level of between-method variability we observed (inclusive of between-observer

variability) is comparable to the variability between observers within a single method. This pattern holds true across both methods. This result suggests that the majority of between-method variability is attributable to between-observer variability.

To test for methodological bias, we surveyed the same 131 benthic transects using both diver-based data collection and imaging-based SfM annotation. Overall, our results suggest that a majority of the metrics do not vary significantly between methods nor between methods across depths. However, metrics derived from in-water and SfM methods **differed** in estimates of **partial mortality, juvenile colony density, and adult diversity** (though not adult taxonomic richness). For these three metrics, our results may present concerns for the continuity of SfM-derived surveys with our historical dataset if we do not carefully follow modifications to our existing protocols and training.

SfM provides a variety of advantages over in-water surveys for assessing coral reef communities, but it also presents some disadvantages. For Pacific NCRMP, SfM reduces in-water survey time, enables fine-scale assessments of structural complexity, and provides a permanent visual of the site from which annotators can re-evaluate SfM imagery to control observer error. However, time consuming data extraction, reliance on good image quality and higher performance computing hardware, and the difficulties with imaging highly complex three dimensional reef structures present challenges to operationalizing SfM.

As ESD transitions to SfM for NCRMP benthic monitoring, there are a number of modifications to keep in mind. Firstly, we will continue to emphasize the critical importance of **high image quality** through the use of a **dedicated benthic imaging** team capable of high daily throughput of sites (i.e., > six sites per day). This dedicated team can not only collect the baseline imagery from 1 m above substrate, but can generate higher resolution **fill-in imagery** in the smaller segment areas critical to our survey metrics. The dedicated imaging team could also allow for **hybrid SfM/in-water data collection**. Such an approach would involve extracting metrics—such as density, colony size, partial mortality, and colony conditions from SfM imagery—while potentially recording metrics which are difficult to extract from SfM imagery, such as juvenile density and specific conditions of interest through in-water observations. This can leverage the strengths of both methods, reducing field costs by allowing the hybrid team to complete a high number of surveys per day while maintaining greater continuity with the historical NCRMP data. To streamline post-processing efforts, we will continue to use a well-calibrated, **dedicated annotation team**. Further, we need to maintain our focus on reducing cross-annotator variability by having regular in water and SfM calibration and feedback to maintain **robust annotator training**. Finally, we will continue to improve the **efficiency of data extraction**, by improving both annotator efficiency and automated processing methods to reduce long-term costs associated with manual annotation and model processing.

Introduction

The Pacific Island Fisheries Science Center's Ecosystem Sciences Division (ESD) is funded by the NOAA Coral Reef Conservation Program to monitor the status and trends of coral reefs across 40 U.S. Pacific islands, atolls, and shallow banks as part of the National Coral Reef Monitoring Program (NCRMP). Since 2000, ESD has used a variety of benthic methods including line point intercept (PIFSC, 2020a), benthic photo-quadrat (PIFSC, 2020b), and belt transect surveys (PIFSC, 2015) to quantify coral demography and benthic cover. Each year Pacific NCRMP visits 200–300 benthic sites by using NOAA's seagoing vessels, which are heavily tasked. The continued demand for increasingly limited resources requires innovative solutions for monitoring benthic communities with reduced field teams.



Figure 1. Map of geographic areas (in white) where ESD conducts coral reef monitoring surveys in the Pacific.

To increase the field efficiency of benthic data collection, Pacific-NCRMP began testing an imaging technique called Structure-from-Motion (SfM) photogrammetry. SfM uses overlapping 2D imagery to create a 3D reconstruction of reef areas. This technique is gaining popularity in coral reef research (Burns et al., 2015; Casella et al., 2017; Bayley et al., 2019; Fox et al., 2019; Obura et al., 2019) and allows researchers to study these ecosystems from the coral polyp to the reef-level. To date, the majority of studies on coral reefs utilizing SfM methodology have focused on quantifying structural complexity (Figueira et al., 2015; Burns et al., 2015; Storlazzi et al., 2016; Ferrari et al., 2017; Fukunaga et al., 2020; Torres-Pulliza et al., 2020). Others have used SfM in small-scale

studies to quantify disease and bleaching (Palma, 2016; Fox et al., 2019; Voss et al., 2019; Burns et al., 2020), spatial clustering of corals (Edwards et al., 2017; Pedersen et al., 2019), and coral growth (Lange and Perry, 2020).

In 2019, after a year of methods development and standardizing procedures (Suka et al., 2019), NCRMP conducted a study in the main Hawaiian Islands to compare data generated from traditional in-water surveys to SfM-derived metrics for assessing coral demography (Couch et al. 2021). The goals were to evaluate whether ESD can maintain continuity in our long-term NCRMP data sets if the program transitions to SfM survey methods, as well as identify the strengths and weaknesses, and compare the costs of both methods. Couch et al. (2021) found that a majority of the metrics measured by NCMRP (adult and juvenile colony density, colony diameter, and genus level diversity) do not vary significantly between methodological approaches. They also suggested that improvements in image quality for SfM and continued training of annotators to improve the consistency of challenging metrics such as partial mortality, disease, and bleaching prevalence will likely enhance the comparability of data collected using these two methods. While SfM provides several benefits over traditional in-water methods, its primary disadvantage is the considerable post-processing bottleneck.

The coral reefs of the central and western Pacific vary widely in species diversity, complexity, and geomorphology. In 2022 and 2023, Pacific NCRMP repeated the methods comparison analysis in the Mariana Archipelago (Guam and the Commonwealth of the Northern Mariana Islands) and American Samoa to test the comparability of these two methods in higher diversity reef systems than the Hawaiian Islands.

The **three objectives** of this study were to:

1. Compare error between methods to within-method observer error,
2. Test for methodological bias between SfM and in-water visual surveys, and
3. Provide recommendations for the transition to SfM for future NCRMP benthic monitoring through a synthesis of Couch et al. (2021) and the present study.

Methods

In-water collection

In May 2022 and April 2023, ESD conducted 311 NCRMP benthic surveys across the Mariana Archipelago and American Samoa. Site selection was based on a one-stage stratified random survey design stratified by depth bin (shallow: 0–6 m; mid-depth: 6–18 m; and deep: 18–30 m) and sub-island sectors ([Figure 2](#)). At each site, one 18 m transect line was deployed along the associated depth contour. Using standard NCRMP methods, visual observations were recorded within four segments, each 1 m wide and spaced at 2.5 m intervals along the transect (at 0–2.5 m, 5–7.5 m, 10–12.5 m, and 15–17.5 m). Staggered segments were surveyed to better capture spatial variability across the transect line (details on how data were summarized in the [Data Analysis](#) section). Only three segments were surveyed at deep sites due to bottom time limitations. For each adult coral colony (colony diameter > 5 cm) within the survey area, maximum diameter, coral ID (to the lowest taxonomic level possible), morphology, percent old (denuded skeleton colonized by turf or other organisms), and recent (recently denuded bare skeleton not yet colonized) mortality were recorded. Estimates of percent mortality were recorded in increments of 1 from 0–10 and from 90–100, and in increments of 5 from 10–90. Divers also recorded any signs of disease or abnormal conditions, taking note of the type of disease and the colony surface area affected. Bleaching severity was scored on a scale from 0–3 (0 = none, 1 = mild, 2 = significant pigmentation loss, 3 = full loss of pigmentation) and recorded only for colonies that scored 2 or 3. Juvenile coral colonies (<5 cm max diameter) were recorded within the first 1 × 1 m portion of the first three segments due to limited survey time. For juvenile colonies, genus ID and max diameter were recorded (see Winston et al. 2020 for further details on methods).

Of the 311 benthic sites, 131 sites (102 from the Mariana Archipelago, 29 from American Samoa) were used for the methodological comparison ([Figure 2](#)). This subset of sites was chosen to equally represent each island and depth bins surveyed, as well as a range of habitat complexities, coral cover, and diving conditions. Within these 131 sites, 23 adult and 32 juvenile segments across 19 sites were haphazardly chosen and re-surveyed by a different diver during the same dive to create a replicate in-water observation for the error comparison analysis ([Figure 2](#)). Only segments from the Mariana Archipelago were included in the error comparison analysis as field conditions limited our ability to resurvey segments in American Samoa.

Prior to participating in underwater demographic surveys, divers went through extensive training and calibration on survey choreography, coral identification, and demographic metrics. Divers were given extensive classroom training and required to pass a coral identification and condition test with a minimum score of 90% to participate in NCRMP

surveys. While in the field, divers were encouraged to take pictures of unknown taxa and conditions for further discussion. These improvements in training resulted from the relatively high within-method inter-observer error seen in Couch et al. (2021).

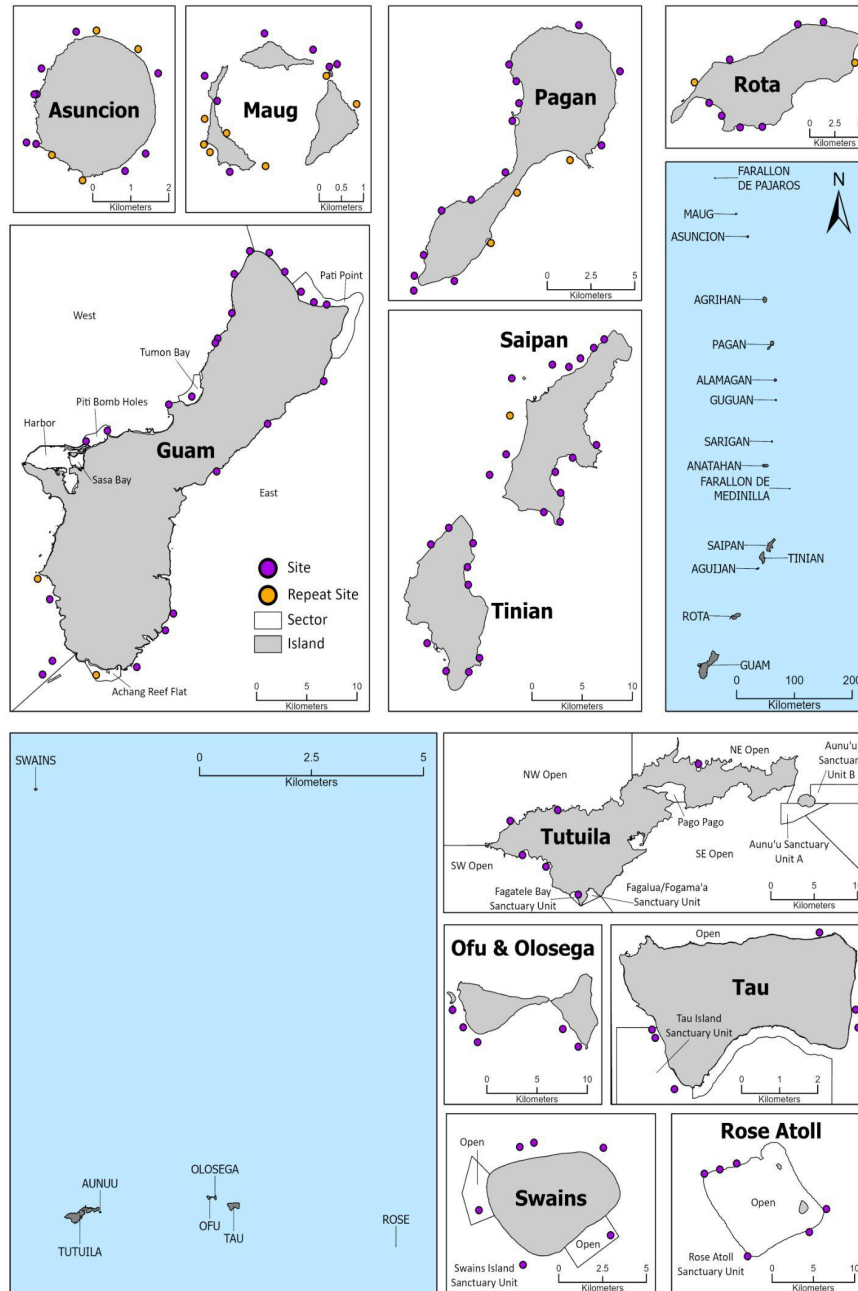


Figure 2. Location of 131 survey sites (purple) with 19 repeated survey sites indicated in orange. Black outlines indicate sub-island sectors where conditions and management contribute to potential differences in benthic communities.

SfM image collection

In conjunction with the standard in-water visual surveys, SfM image collection was also conducted at each site. The SfM survey was conducted over a 3 × 20 m area centered over the transect at depths of 0–18 m (Figure 3), and a 3 × 13 m area at depths >18 m. After color balancing with an 18% gray card, divers took continuous JPEG images along the transect using an entry-level digital SLR camera (Canon EOS Rebel SL2 & SL3, Ikelite underwater housing with six inch dome port) with an 18–55 mm lens set at 18 mm. Divers swam along the transect for six passes (three on each side) spaced 0.5 m apart while maintaining a 1 m height off of the seafloor. This swim pattern allowed for ~60% side overlap and ~80% forward overlap of images. Scale bar markers, also known as Ground Control Points (GCPs), were placed at the beginning of each segment at least 0.5 m away from the transect line. The depth and position along the transect line of each GCP were recorded on a datasheet. Imagery was manually quality controlled (QC'd) to ensure only quality imagery (e.g., no overexposed or blue imagery) was included in the models.

For more details on the process for selecting equipment, image collection methods and survey design refer to Couch et al. (2021) and Torres-Pulliza et al. (2023).

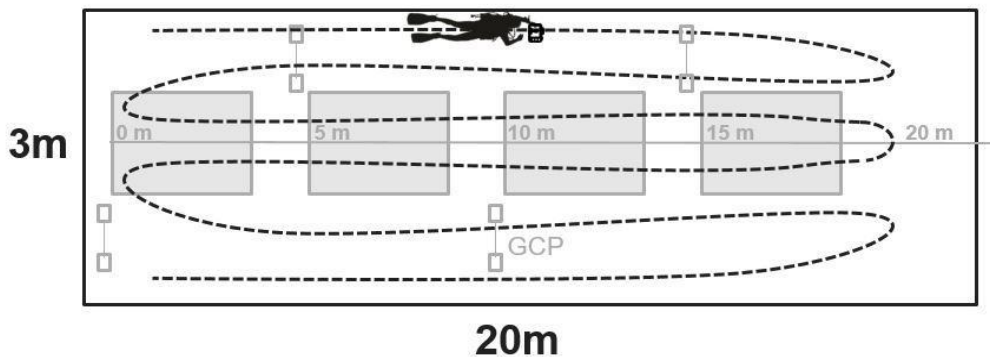


Figure 3. Graphic of benthic survey site with in-water visual survey segments (in gray). The SfM swim path is indicated by the black dashed line, covering a 3 m x 20 m area. Scale bar markers, or GCPs, are represented by the two rectangles that are connected by a thin line.

SfM model generation and data extraction

A 3D model of each site was generated using Agisoft Metashape software (Agisoft Metashape Professional Version 1.8.5). The workflow sequence included aligning images, building and exporting the 3D dense point cloud (DPC), and orthomosaic following parameters described by Burns et al. (2015) and Torres-Pulliza et al. (2023). DPCs were imported into Viscore, a custom 3D model visualization developed at the

Scripps Institution of Oceanography at the University of California, San Diego (Petrovic et al. 2014). Orthomosaics were imported into ArcGIS Pro 2.8.0 for manual colony annotation. See Torres-Pulliza et al. (2023) for detailed information.

In ArcGIS Pro, each site was set up for annotation by manually digitizing the transect and segments as a shapefile ([Figure 4A](#)) and setting up the attribute table in a geodatabase to mirror the in-water visual survey database. To record and extract data from the orthomosaic, coral colonies within each segment were annotated following the in-water visual survey methods. Each colony was measured by digitizing a line across the maximum diameter of the colony ([Figure 4B](#)). Coral ID (to the lowest taxonomic level possible), morphology, disease/condition, and percent estimates of old and recent mortality were recorded. During annotation, the original *JPEG* imagery was viewed alongside the orthomosaic using the Viscore image view feature to see fine-scale colony details and observe colonies from multiple angles ([Figure 4C](#)). Annotators relied on the underlying original imagery to not only identify colonies and conditions, but also locate colonies not visible in the orthomosaic. Using the underlying imagery was crucial to achieving the highest data quality and comparability to in-water surveys.

The SfM observers in this study were part of the in-water diver team and underwent extensive training on coral taxa and condition identification as described above. All SfM observers annotated the same ten segments as part of an inter-observer calibration exercise to correct for potential issues before collecting data from the SfM imagery. Further, observers were encouraged to review annotations together during the annotation process to reduce inter-observer variance. In addition to standard annotation, segments that were repeated in water were also annotated twice by different SfM observers to create replicate data sets for both methods.

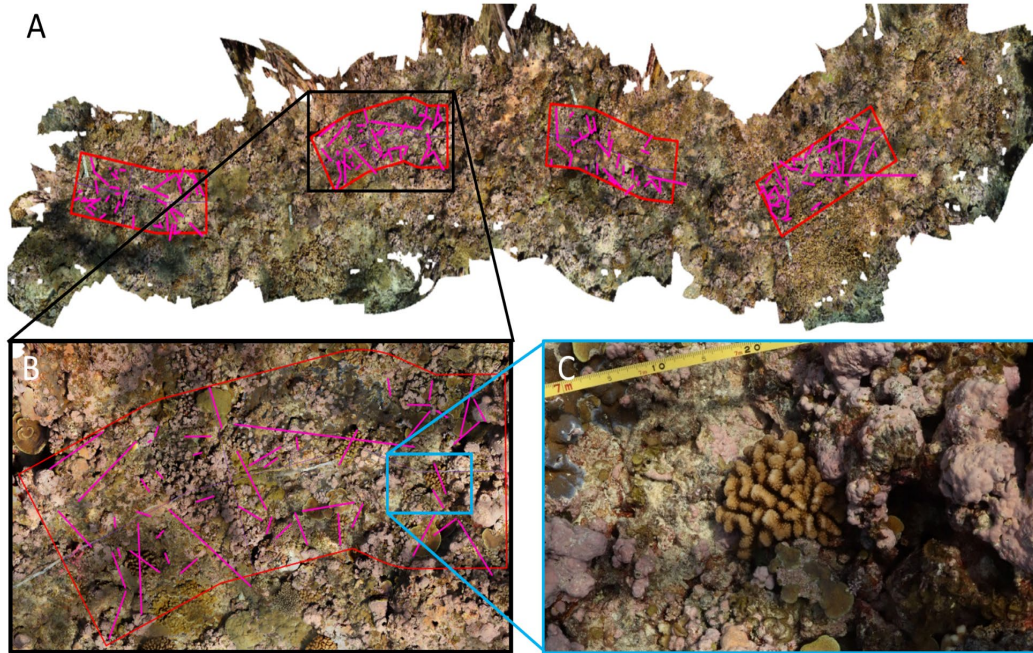


Figure 4. A) Full orthomosaic with all segments (red boxes) and colonies annotated (magenta lines). B) One fully annotated segment in the orthomosaic with a C) raw image containing a *Pocillopora* sp. colony.

Quality control (QC) of the extracted SfM data was carried out in two stages:

1. QC script: R scripts with specific queries were used to identify data entry errors (e.g., misspelled species names, data in the incorrect column) and were corrected in the ArcGIS geodatabase;
2. Spot check: site-level metrics were examined across annotators to identify potential issues for annotators to spot check and correct. If persistent issues were identified for a given annotator (e.g., not identifying specific conditions), the annotator reviewed each site they annotated to correct these issues.

Given the improvements that were made to diver/annotator calibration and enhanced communication between divers/annotators, we found that annotator error in the calibration exercise was sufficiently low, and spot checks of site-level metrics in the full dataset had lower annotator error as well.

Data analysis

All data were analyzed in R v4.1.2 (R Core Team, 2021). Coral demographic metrics were summarized as follows: adult colony density (number of colonies ≥ 5 cm per m^2), juvenile colony density (number of colonies 0.7–4.9 cm per m^2), average maximum adult diameter, average percent old and recent partial mortality, chronic and acute disease prevalence, bleaching prevalence (percent of colonies with a bleaching severity

≥ 2), and adult and juvenile diversity: adult Shannon-Wiener diversity, Hill diversity, and genus richness (number of genera).

To compare error between methods to within-method observer error for the ten demographic metrics, data were summarized at the segment-level for all scleractinians combined. Error was calculated as the absolute difference in values (between methods or observers) divided by the overall mean so that we could compare the relative level of error across metrics (termed “midpoint scaled mean absolute error”: MS-MAE). We calculated error for three different types of comparisons for the 19 paired transects sampled by both SfM and in-water methods. “Diver observer error” represents the difference between divers for a given demographic metric (Figure 5A). “SfM observer error” represents the difference in error for a given metric between SfM annotators (Figure 5B). “Method Error” is the difference between methods for all possible combinations of method x observer divided by the mean of the absolute difference across all method x observer comparisons for a given metric (Figure 5C). We summarize each of these error distributions using the mean and standard error of the mean. We used nonparametric Kruskal-Wallis tests to test for differences between the three errors for metrics that did not meet assumptions of normality and equal variance. We further calculated MS-MAE for all 131 transects to compare the between method error for this study and Couch et al. (2021) to examine if training protocols implemented after 2021 led to better calibration between observers.

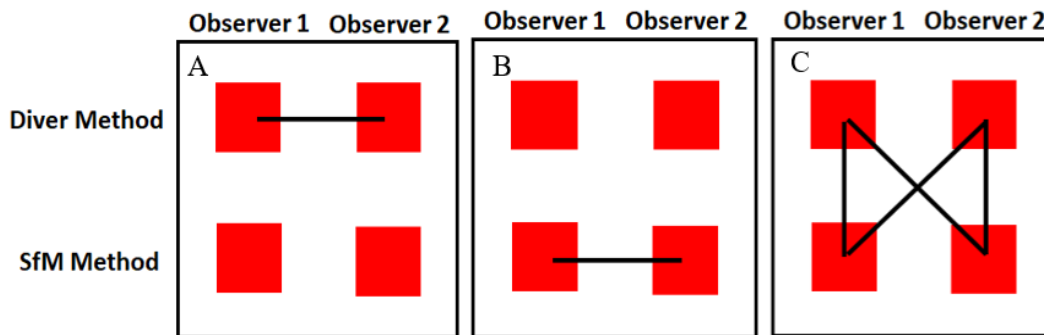


Figure 5. Graphic depicting comparisons (A) between observers for the in-water observers (in-water divers), (B) between observers for the SfM imagery, and (C) between the methods of diver-derived data (in-water method) and SfM image-derived data (SfM method).

To test for differences between methods in the coral demographic metrics, data were summarized at the site-level at 131 sites that were surveyed by one diver and one SfM annotator. Within a site, only segments that were surveyed in both methods were included and then pooled to the site-level (5–10 m² of reef area/site). Data were pooled at the site-level because segments are not considered independent samples, but rather a means to capture spatial variability and this is the lowest spatial resolution that ESD typically summarizes NCRMP data. Results are presented for the eight metrics for total

scleractinian corals combined. Each metric was tested for normality and equal variance. Measures of adult density, juvenile density, and average old partial mortality were square root transformed. Average colony diameter was log transformed. 1:1 plots with linear regressions were used to compare and visualize in-water and SfM-generated estimates for each metric. Root mean squared error (RMSE) was used to evaluate the level of error between methods and was calculated as follows: $RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - x_i)^2}{n}}$; where y_i is the in-water metric value for a given site, x_i is the SfM-generated metric value for a given site and n is the total number of sites.

We then established a series of linear mixed effects models (LMMs) to test two models: method type, and method \times maximum depth. Method and depth were treated as fixed effects and sub-island sector was treated as a random effect. We hypothesize that our metrics may vary across depths given issues of limited bottom time for divers working on deeper reefs, and lower image quality on deeper reefs. To assess the significance of fixed effects, we refit each model using maximum likelihood (ML) estimation and applied likelihood ratio tests (LRTs) (Zuur et al., 2009). Fixed effects that were not significant were sequentially dropped from models. Habitat type was not included in this analysis as it is closely associated with depth, and because LMMs were over-fit when both depth and habitat type were included. The resulting best-fit models were refit using restricted maximum likelihood (REML) to estimate the fixed-effects parameters and associated effect sizes.

Average recent mortality and all prevalence metrics could not be transformed and, therefore, were only tested for overall difference between methods using nonparametric Wilcoxon rank sum tests for each metric. This approach was applied to each coral genera to test for the significance of method type for adult genera that had at least 30 colonies, and for juvenile genera that had at least eight colonies observed across both methods. Different threshold values were chosen for adult and juvenile genera as sample size below these values were uneven across methods. As we were testing for biases in the most common taxa observed, fewer colonies would have limited our ability to detect statistically meaningful results.

Methodological biases in adult and juvenile diversity metrics (Shannon-Weiner diversity, Hill diversity, and richness) were tested using the same LMM and LRT approach as the demographic methods. Diversity was analyzed at the genus level for both adults and juveniles. No diversity metrics were transformed to fit model assumptions.

Results

How does methodological error compare to observer error?

To understand whether the absolute difference between methods was greater or smaller than the difference between observers, we compared the level of between-method error (inclusive of between-observer error, e.g., [Figure 5C](#)) to between-observer error within each method ([Figure 5A, B](#)) in the Mariana Islands ([Figure 6](#)). For all metrics there were no significant differences between the three error types. Overall, this suggests that while there may be variability between methods, the level of error is consistent with that seen between divers and annotators.

It is important to note for all of these comparisons that the identities of observers within each method were randomly assigned to observer “1” or “2” (i.e., Diver 1 vs. Diver 2, SfM 1 vs. SfM 2). Therefore, comparisons between observers within a method highlight general variation among multiple observers but do not reflect the tendencies of a single, human observer.

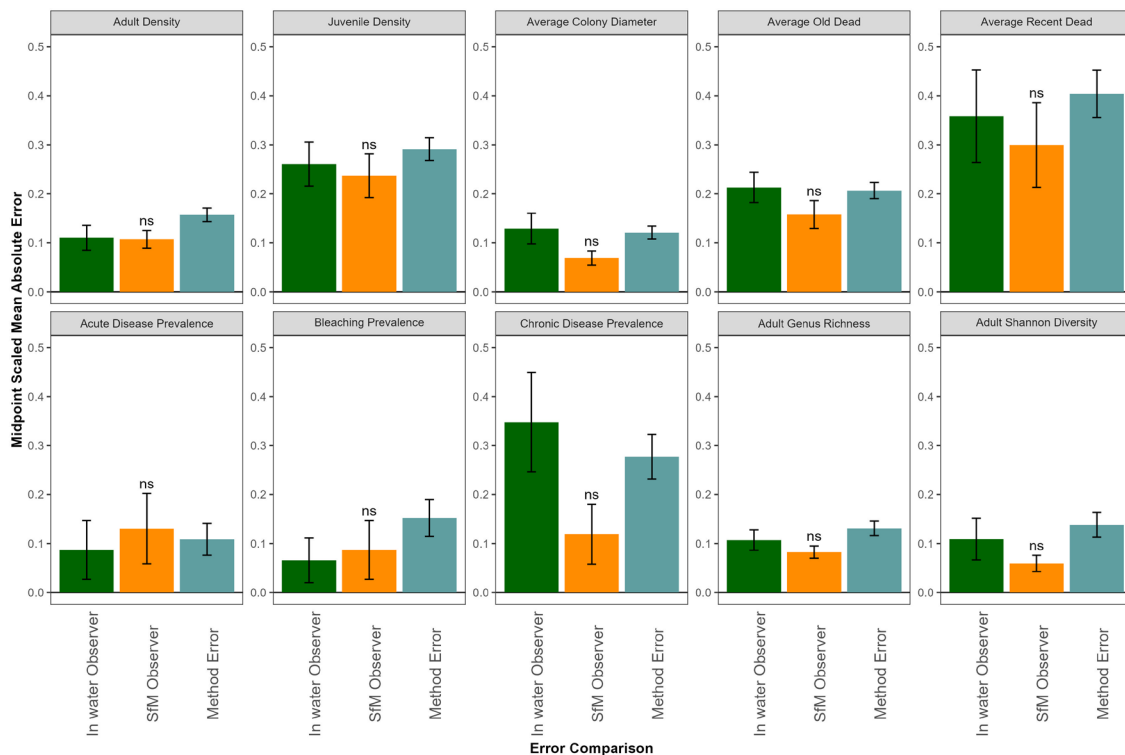


Figure 6. A comparison of the in-water observer error (difference between divers) and SfM observer error (difference between SfM annotators) to method error (difference between methods) for each coral demographic metric. Error is represented as the midpoint scaled mean absolute error (\pm SE) from 23 adult and 32 juvenile segments that were surveyed by multiple divers and SfM annotators. ns = not significant ($p > 0.05$). Data only from the Mariana Archipelago.

Is there a methodological bias in coral demographic metrics at the site-level?

Adult density of total scleractinians

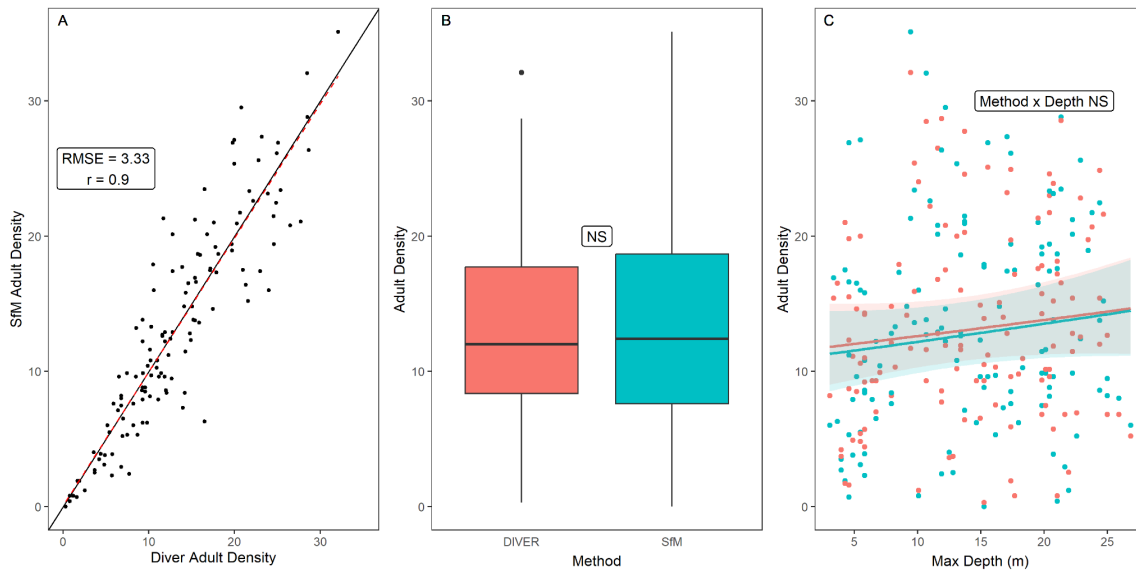


Figure 7. Plot of the paired site-level SfM adult colony density vs. diver adult colony density (points). Black line is 1:1 line, red dashed line is linear regression line; (B) boxplot of adult density by method type; (C) marginal effects plot of adult density by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (method or method x depth) from the linear mixed effect models. NS: $p > 0.05$

At the site-level, total adult colony density showed a strong correlation between methods with a low RMSE and more variability above 10 colonies/m² (Figure 7A). We did not detect a significant methodological difference (Figure 7B, Appendix A) nor an interaction between method and depth (Figure 7C, Appendix A).

Juvenile density of total scleractinians

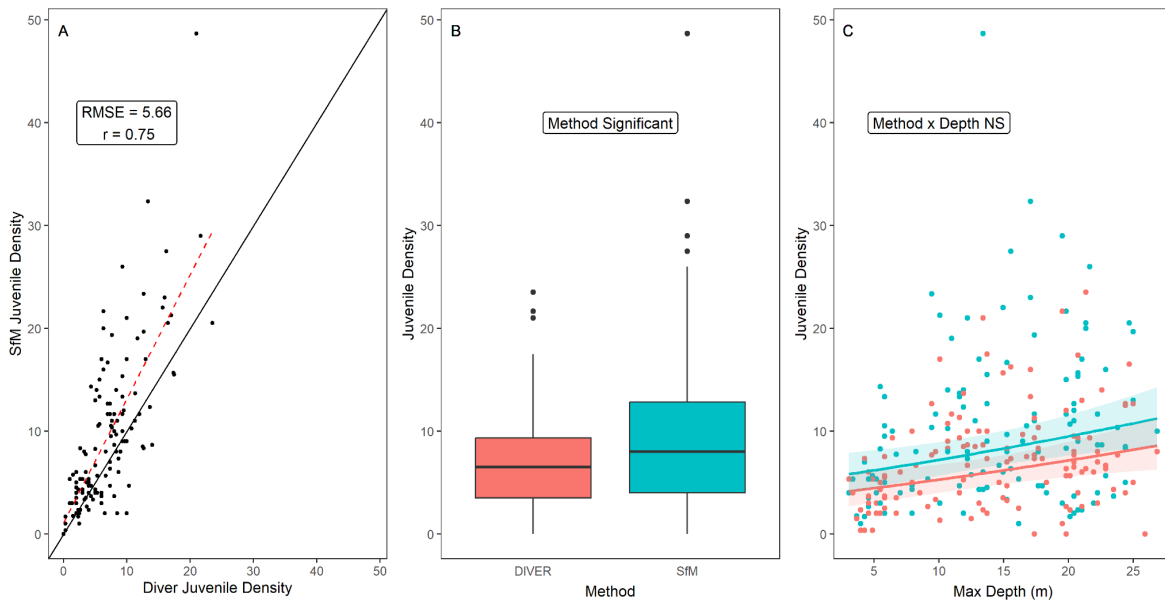


Figure 8. (A) Plot of site-level SfM juvenile colony density vs. diver juvenile colony density (points). Black line is 1:1 line, red dashed line is linear regression line; (B) boxplot of juvenile density by method type ($p < 0.001$); (C) marginal effects plot of juvenile density by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (method or method x depth) from the linear mixed effect models. NS: $p > 0.05$

Total juvenile colony density was strongly correlated between methods with a moderate RMSE (Figure 8A). However, juvenile density was significantly higher in SfM compared to in-water surveys, with SfM returning higher juvenile density by an average of 2.25 colonies/m² (Figures 9A, B; Appendix A). There was no significant interaction between method and depth (Figure 8C, Appendix A). Models were also fit with an outlier (SfM juvenile density close to 50) removed, but this did not change the LRT results.

Adult colony diameter of total scleractinians

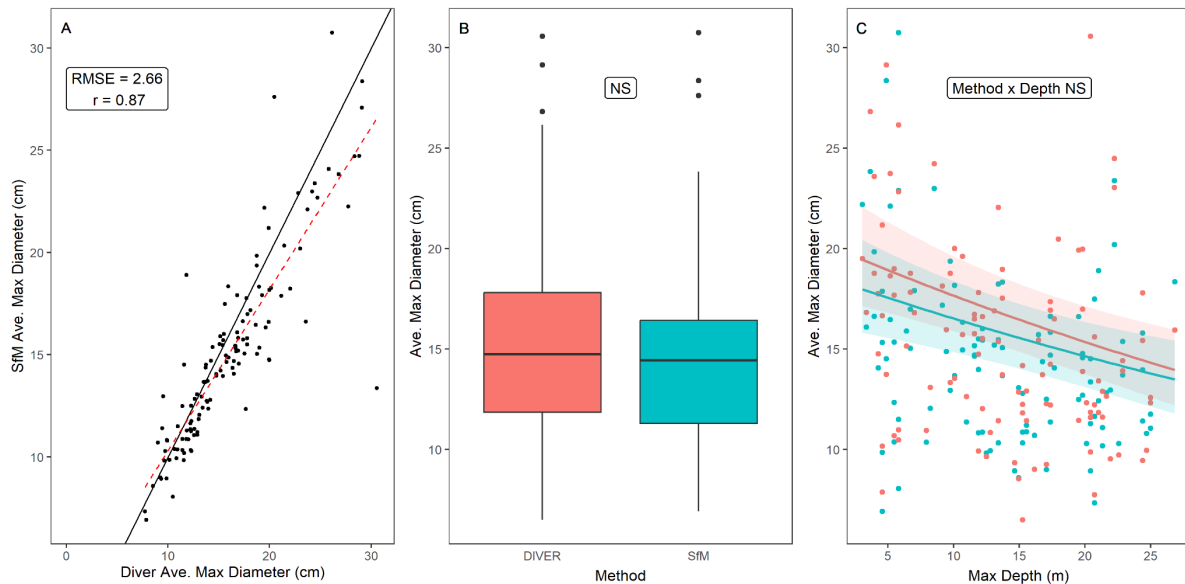


Figure 9. (A) Plot of site-level SfM adult average maximum colony diameter vs. diver adult average maximum colony diameter (points). Black line is 1:1 line, red dashed line is linear regression line; (B) boxplot of average maximum diameter by method type; (C) marginal effects plot of average maximum diameter by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models. For plots B and C, the results of the likelihood ratio tests are included for the different fixed effects (method or method x depth) from the linear mixed effect models. NS: $p > 0.05$

Overall adult average maximum diameter was strongly correlated between methods with a low RMSE (Figure 9A), and we did not detect a significant methodological difference (Figure 9B, Appendix A). There was no significant interaction between method and depth (Figure 9C).

Average old and recent partial mortality of total scleractinians

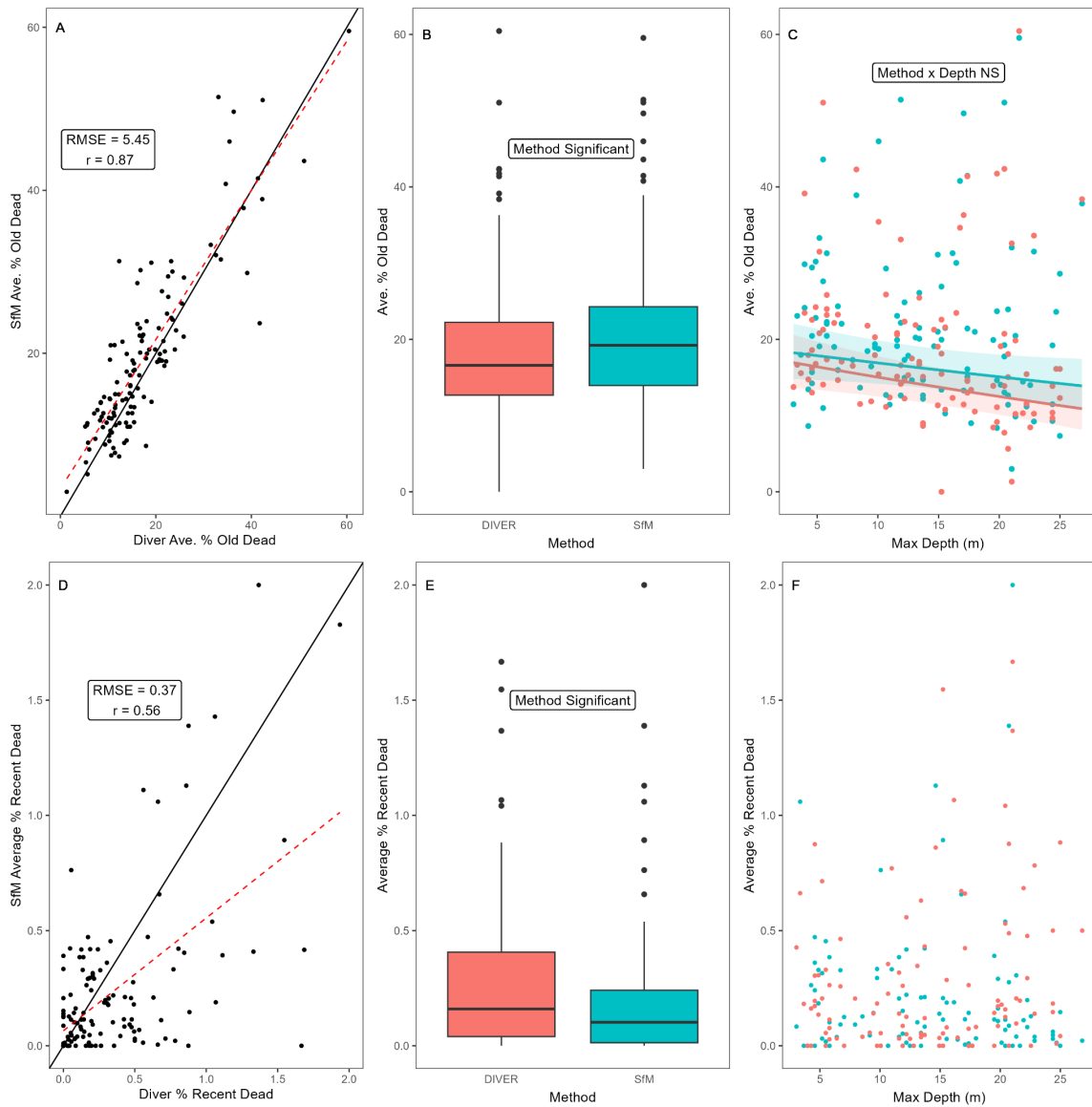


Figure 10. (A, D) Plot of site-level SfM adult average percent partial mortality vs. diver adult average percent old partial mortality (points) for old and recent mortality, respectively. Black line is 1:1 line, red dashed line is linear regression line; (B, E) boxplot of average mortality by method type for old and recent mortality, respectively; (C, F) marginal effects plot of average old mortality by maximum depth (m) including predicted values (lines) and confidence intervals of predictions (shaded areas) from linear mixed effects models (LMMs). For plots B and C the results of the likelihood ratio tests are included for the different fixed effects (method, or method x depth) from the LMMs. NS: $p > 0.05$. For recent mortality, the difference between methods overall (E) were tested using a nonparametric Wilcoxon test. An outlier in percent recent dead ($>10\%$ recent dead) was removed from the plots to better visualize the majority of data presented.

Adult average percent old partial mortality strongly correlated with a moderate RMSE between methods (Figure 10A). SfM percent old partial mortality was significantly higher than in-water mortality, but the mean difference between methods was only 2% (Figure 10C, Appendix A). There was no significant interaction between method and depth (Figure 10C, Appendix A).

Adult average percent recent mortality was moderately correlated (Figure 10D), which is likely due to the high proportion of values less than 1%. Recent mortality collected by divers was significantly higher than in SfM (Figure 10E, Appendix A). Additionally, recent mortality was poorly correlated with depth for both methods, suggesting there is no significant interaction between method and depth (Figure 10F; SfM : Spearman ρ = - 0.10; diver : Spearman ρ = 0.07). At 11.5% of the sites, divers recorded recent mortality that was not recorded by SfM annotators, and at 6.9% of sites, SfM annotators recorded recent mortality that divers did not record.

Condition prevalence of total scleractinians

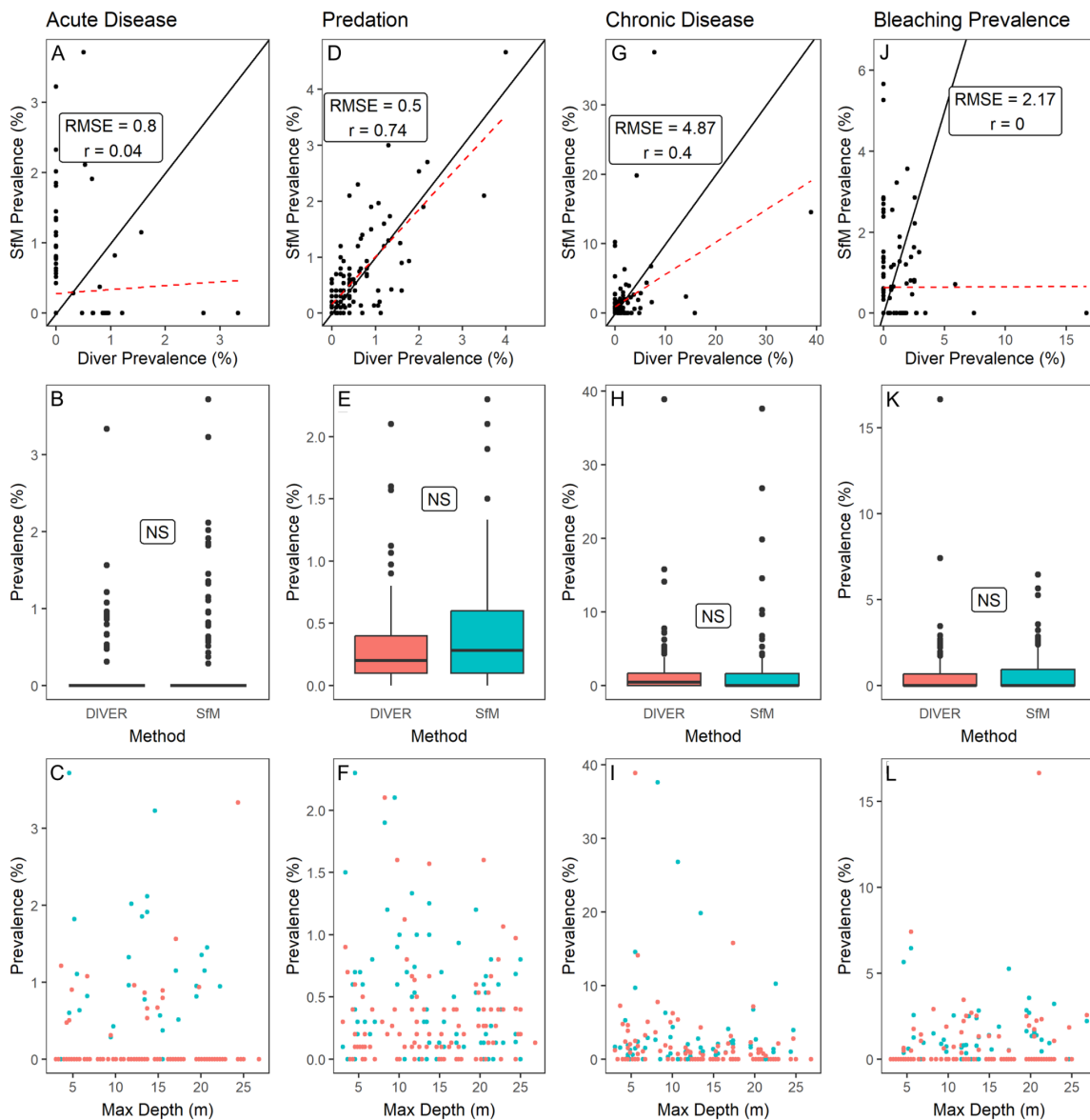


Figure 11. (A) Plot of site-level Sfm prevalence vs. diver prevalence (points) where black line is 1:1 line, and red dashed line is linear regression line; (B) boxplots of prevalence by method type, significance ($\alpha=0.05$) of method was tested using a nonparametric Wilcoxon test, NS: $p>0.05$; and (C) plot of prevalence by maximum depth (m) standard error (shaded areas) for acute disease. (D, E, F) is predation, (G, H, I) is chronic disease, and (J, K, L) is bleaching prevalence.

Acute disease prevalence was not correlated between methods ([Figure 11A](#)). While prevalence was slightly higher for Sfm compared to divers, it did not differ significantly between methods ([Figure 11B](#), [Appendix A](#)). Prevalence was similarly poorly correlated with depth for both methods, suggesting there is not a significant interaction between

method and depth ([Figure 11C](#); SfM: Spearman $\rho = -0.11$; diver : Spearman $\rho = -0.11$). At 9.2% of sites, divers recorded acute disease that was not recorded by SfM annotators, and at 16.8% of sites, SfM annotators recorded acute disease that divers did not record.

Predation prevalence was strongly correlated between methods ([Figure 11D](#)). While prevalence was slightly higher for SfM compared to divers, there was no significant difference between methods ([Figure 11E](#), [Appendix A](#)). Prevalence was similarly poorly correlated with depth for both methods, suggesting there is no significant interaction of method and depth ([Figure 11F](#); SfM: Spearman $\rho = -0.04$; diver: Spearman $\rho = 0.08$). At 12.2% of sites, divers recorded predation that was not recorded by SfM annotators, and at 10.7% of sites, SfM annotators recorded predation that divers did not record.

Chronic disease prevalence was weakly correlated between methods with a moderate RMSE ([Figure 11G](#)). Chronic disease prevalence was not significantly different between methods ([Figure 11H](#), [Appendix A](#)). Additionally, prevalence was poorly correlated with depth for both methods, suggesting there is no significant interaction between method and depth ([Figure 11I](#); SfM: Spearman $\rho = -0.19$; diver: Spearman $\rho = -0.29$). At 17.6% of sites, divers recorded chronic disease that was not recorded by SfM annotators, and 13.7% of sites, SfM annotators recorded chronic disease that divers did not record.

Bleaching prevalence was not correlated between methods with a higher RMSE ([Figure 11J](#)). Bleaching prevalence was not significantly different between methods ([Figure 11K](#), [Appendix A](#)). Prevalence was similarly poorly correlated with depth for both methods, suggesting there is no significant interaction between method and depth ([Figure 11L](#); SfM: Spearman $\rho = 0.07$; diver: Spearman $\rho = 0.02$). At 16.7% of sites, divers recorded bleaching that was not recorded by SfM annotators, and 18.3% of sites, SfM annotators recorded bleaching that divers did not record.

Density of dominant coral genera

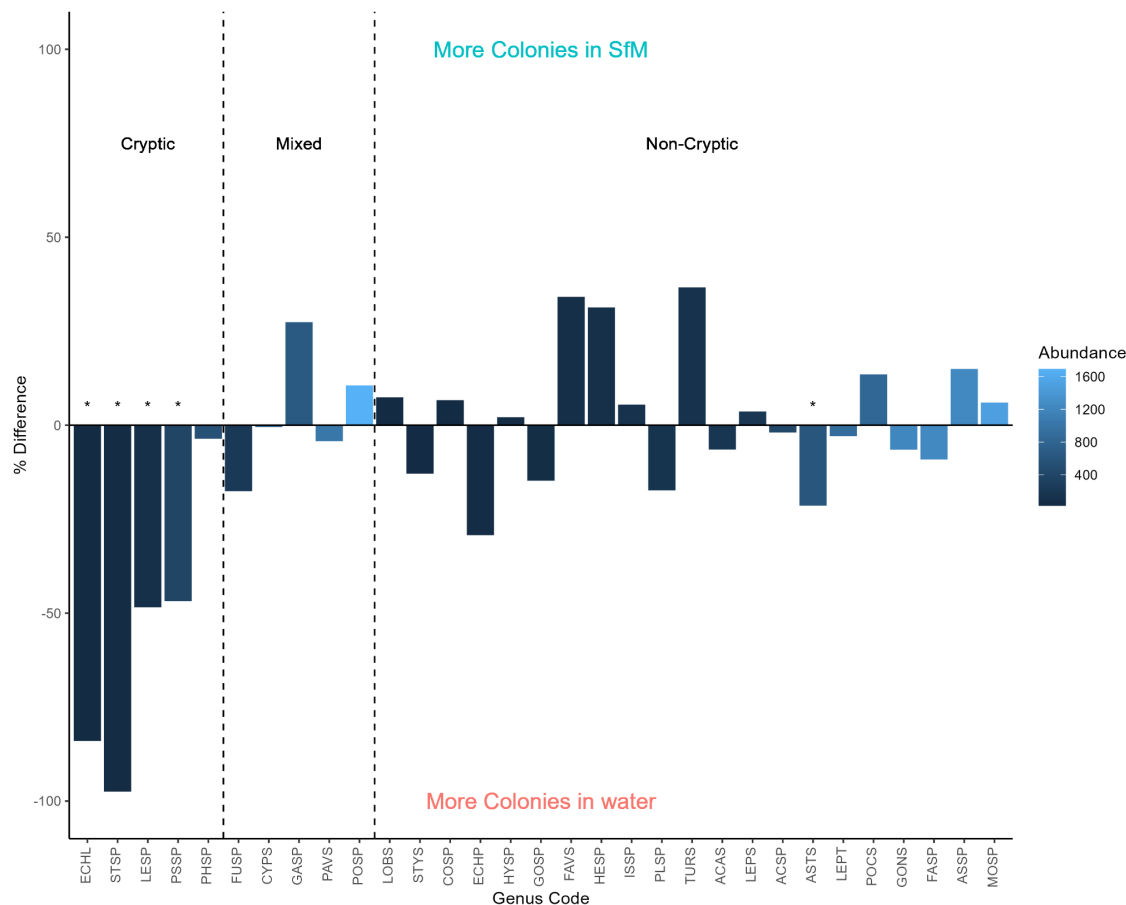


Figure 12. Percent difference of adult colony density for different coral genera. Only genera with at least 30 colonies observed across all sites for both methods are included. Color corresponds to overall abundance of each genera. Cryptic refers to genera that typically live in cracks or under overhangs and may not be visible in the orthomosaic. Nonparametric Wilcoxon tests were done for genera that had at least 30 colonies observed across both methods. Asterisks (*) indicate significant results ($p < 0.05$). Taxa that had fewer than 30 colonies were not analyzed statistically due to low sample size. See [Appendix B](#) for the genus code lookup table.

For most coral genera, colony density was moderately to strongly correlated (50–95%) between methods (Figure 12). Cryptic taxa were generally observed more often in-water than in SfM. For example, observers counted 41 colonies of *Stylocoeniella* sp. (STSP) in-water, while only one colony was counted in SfM. Of the 49 genera observed, colony density was significantly higher in-water for *Astrea* spp. (ASTS), *Psammocora* spp. (PSSP), *Leptoseris* spp. (LESP), *Echinophyllia* spp. (ECHL), and STSP.

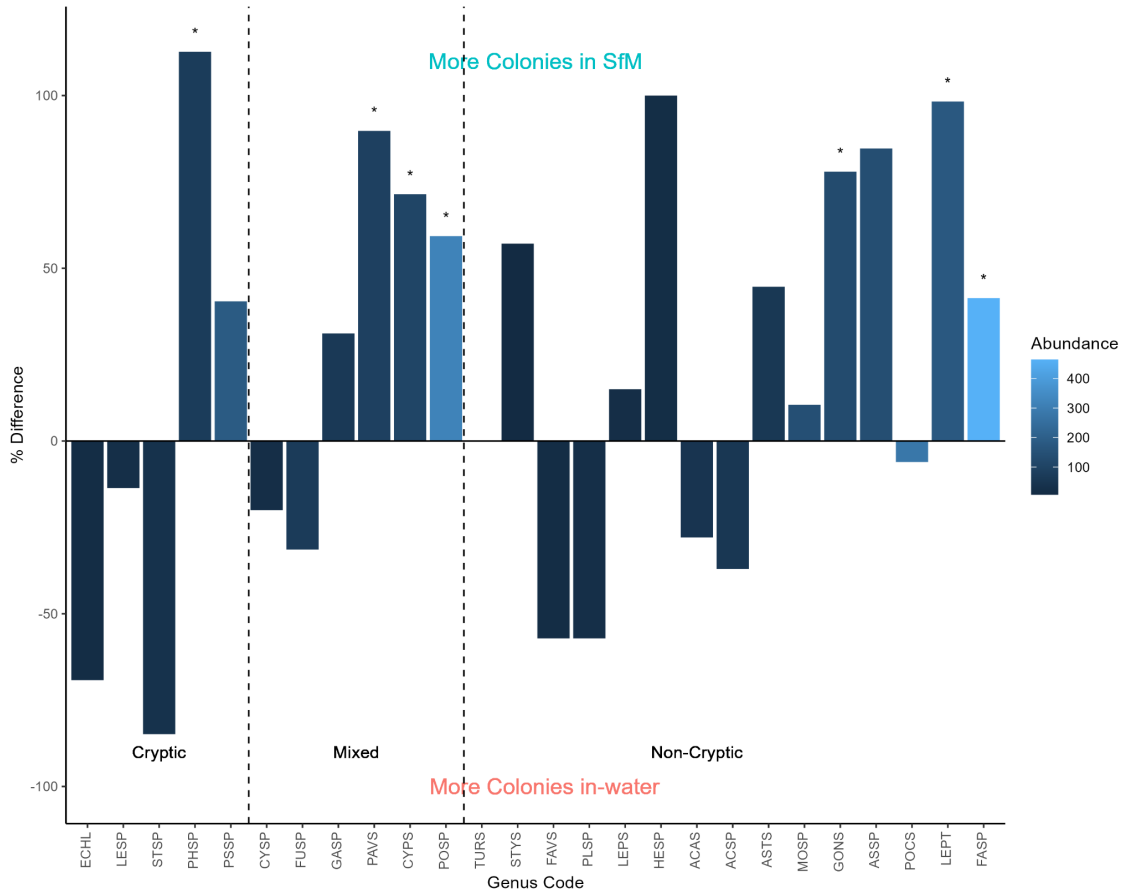


Figure 13. Percent difference of juvenile colony density for different coral genera. Only genera with at least eight colonies observed across all sites in-water are included. Color corresponds to overall abundance of each genera. Cryptic refers to genera that typically live in cracks or under overhangs and may not be visible in the orthomosaic. Nonparametric Wilcoxon tests were done for genera that had at least eight colonies observed across both methods. Asterisks (*) indicate significant results ($p < 0.05$). Taxa that had fewer than eight colonies were not analyzed statistically due to low sample size. See [Appendix B](#) for the genus code lookup table.

Juvenile colony density observed in-water was weakly to strongly correlated (5–90%) with colony density observed in SfM ([Figure 13](#)). Juvenile colony density was significantly higher in SfM for seven taxa: *Favia sp.* (FASP), *Goniastrea sp.* (GONS), *Porites sp.* (POSP), *Leptastrea sp.* (LEPT), *Cyphastrea sp.* (CYPS), *Pavona sp.* (PAVS), and *Phymastrea sp.* (PHSP). Juvenile STSP colony density was significantly higher in-water. There appears to be no consistent bias in cryptic taxa across methods. For example, PHSP and PSSP were observed more in SfM, whereas LESP and STSP were observed more in-water.

Diversity of coral genera

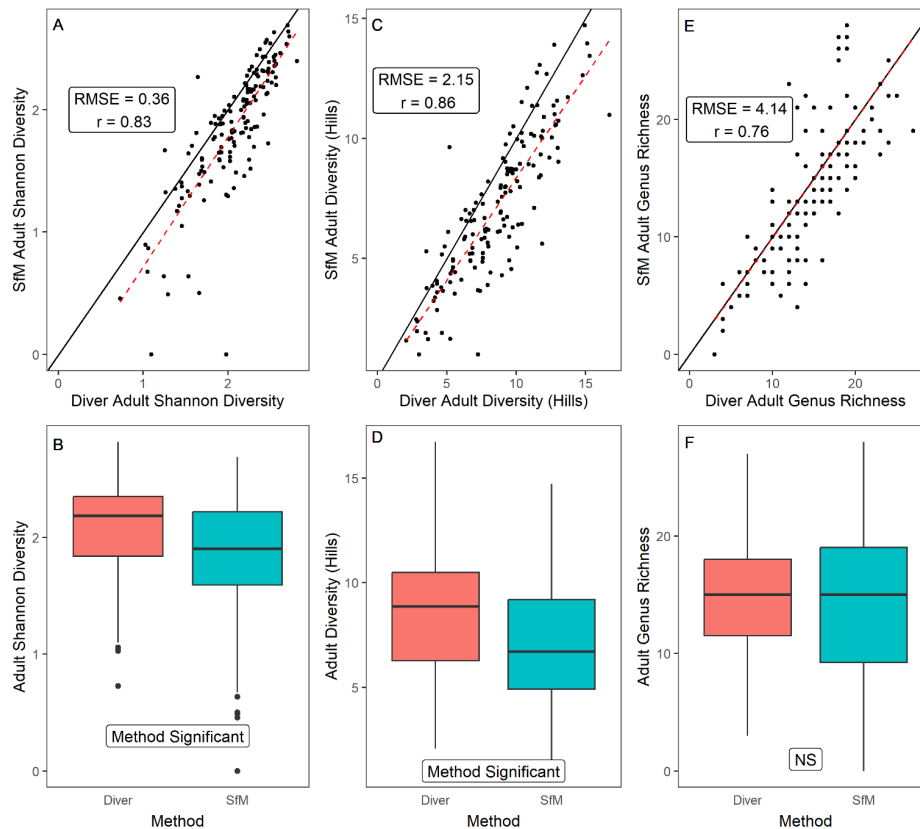


Figure 14. Adult coral diversity. (A, C, E) Plot of site-level SfM taxa diversity vs. diver taxa diversity (points) where black line is 1:1 line, and red dashed line is linear regression line for Shannon-Wiener, Hill, and Richness, respectively. (B, D, F) Boxplots of corresponding diversity metrics (Shannon-Wiener: $p < 0.001$, Hill: $p < 0.001$). For plots B, D, and F, the result of the likelihood ratio test is included for the fixed effect (method) from the linear mixed effect models. NS: $p > 0.05$

Adult genus diversity was strongly correlated between methods for all three diversity metrics with relatively low RMSE (Figure 14A, C, E). Both Shannon and Hill diversity were significantly higher in-water than in SfM (Figure 14B, D; Appendix A), while genus richness did not vary significantly between methods (Figure 14F; Appendix A).

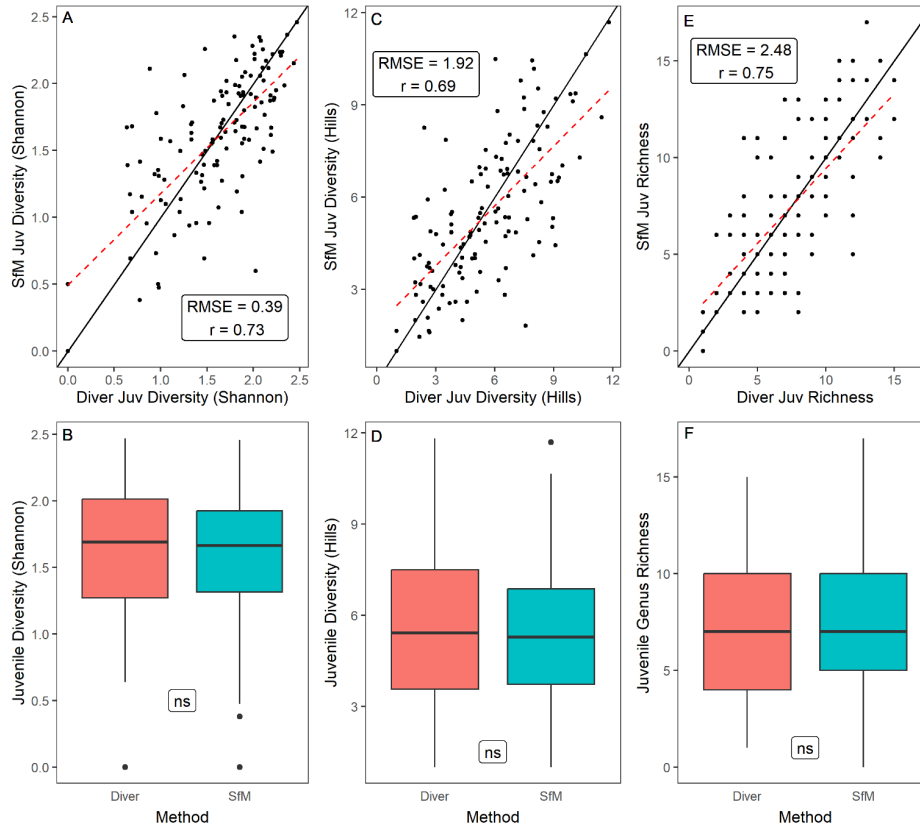


Figure 15. Juvenile coral diversity. (A, C, E) Plot of site-level SfM genus diversity vs. diver genus diversity (points) where black line is 1:1 line, and red dashed line is linear regression line for Shannon-Wiener, Hill, and Richness, respectively. (B, D, F) Boxplots of corresponding diversity metrics. For plots B, D, and F, the result of the likelihood ratio test is included for the fixed effect (method) from the linear mixed effect models. NS: $p > 0.05$

Juvenile genus diversity was moderately to strongly correlated between methods for all three diversity metrics with low RMSE ([Figure 15A, C, E](#)). However, there was no significant difference between methods for any juvenile diversity metric ([Figure 15B, D, E](#); [Appendix A](#)).

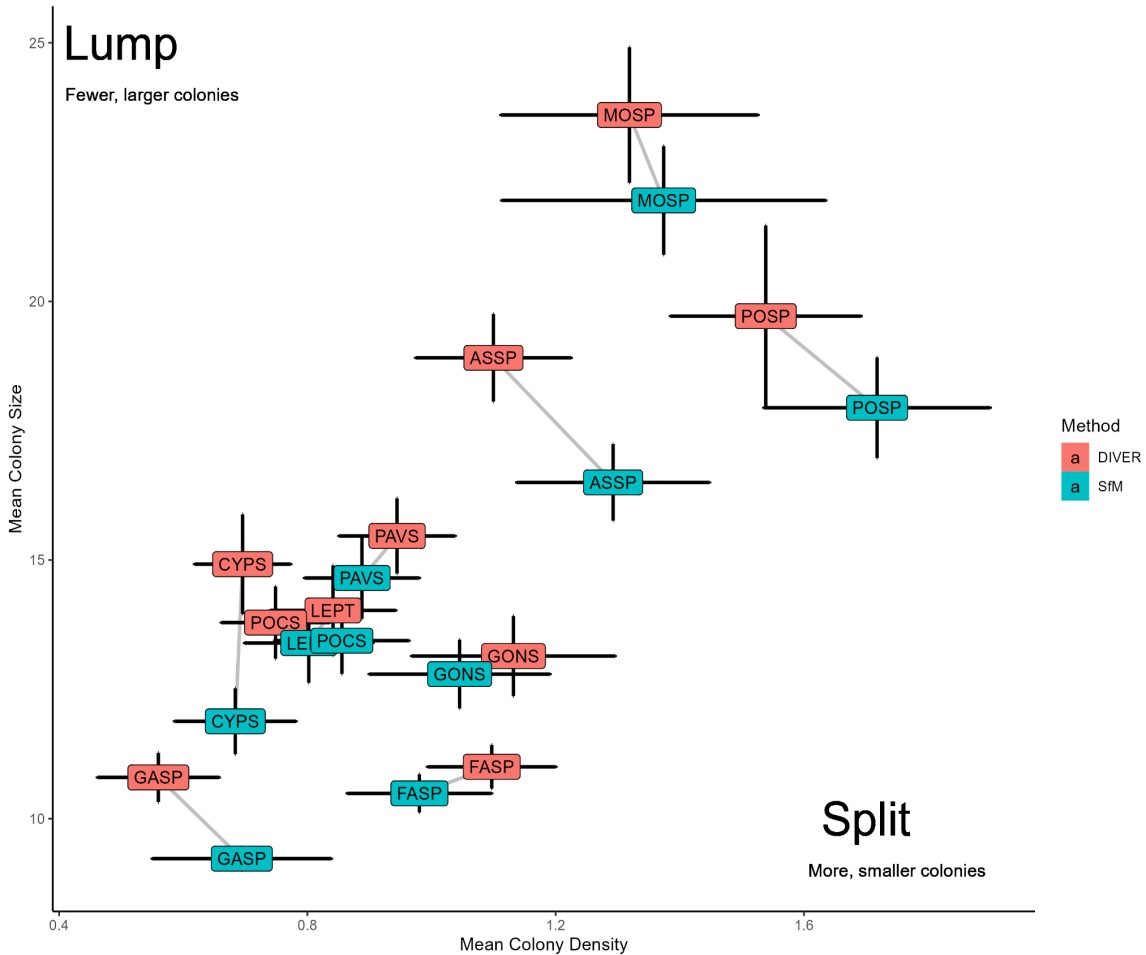


Figure 16. Lumping-splitting challenge visualized. Mean genus-level colony size (cm) and density between methods for top quartile of coral genera by mean density. Error-bars are standard errors of site-level means. Gray lines show the direction of lumping-splitting differences between methods. See [Appendix B](#) for genus code lookup table.

In this higher diversity dataset, relative to divers, our SfM annotators tend to split up colonies of common reef taxa, resulting in higher densities and smaller colonies for the three most common taxa (POSP, *Astreopora* sp. (ASSP), and *Montipora* sp. (MOSP); [Figure 16](#)). However, the differences in colony density were not statistically significant for these taxa ([Figure 12](#)).

ESA listed taxa

Table 1. Counts of ESA listed coral species at comparison sites. Other ESA listed coral species were not observed despite their suspected presence in the Mariana Archipelago and American Samoa regions. Bold rows indicate total counts for each species.

Site	Species	In-water Count	SfM Count
ASC-00623	<i>Acropora globiceps</i>	1	0
MAU-01248	<i>Acropora globiceps</i>	1	0
OFU-01322	<i>Acropora globiceps</i>	1	2
OFU-01340	<i>Acropora globiceps</i>	1	1
ROT-00830	<i>Acropora globiceps</i>	1	1
SAI-01909	<i>Acropora globiceps</i>	1	0
TIN-00868	<i>Acropora globiceps</i>	1	0
ALL	<i>Acropora globiceps</i>	7	4
OFU-01342	<i>Isopora crateriformis</i>	5	9
TAU-01306	<i>Isopora crateriformis</i>	27	27
TUT-05690	<i>Isopora crateriformis</i>	3	1
TUT-05795	<i>Isopora crateriformis</i>	18	12
TUT-05876	<i>Isopora crateriformis</i>	69	76
ALL	<i>Isopora crateriformis</i>	117	125

Three ESA listed coral species (*Acropora globiceps*, and *Isopora crateriformis*) were observed at our survey sites ([Figure 17](#)). Colony counts were similar across methods, with the largest percent difference in counts in *Acropora globiceps* where the presence of this taxa was not observed in SfM when observed in-water ([Table 2](#)). These species were not recorded in SfM when absent at the same site in-water. *Isopora crateriformis* was found more often in SfM than in-water at two sites, OFU-01342 and TUT-05876, and *Acropora globiceps* was found more often in SfM than in-water at one site, OFU-01322.

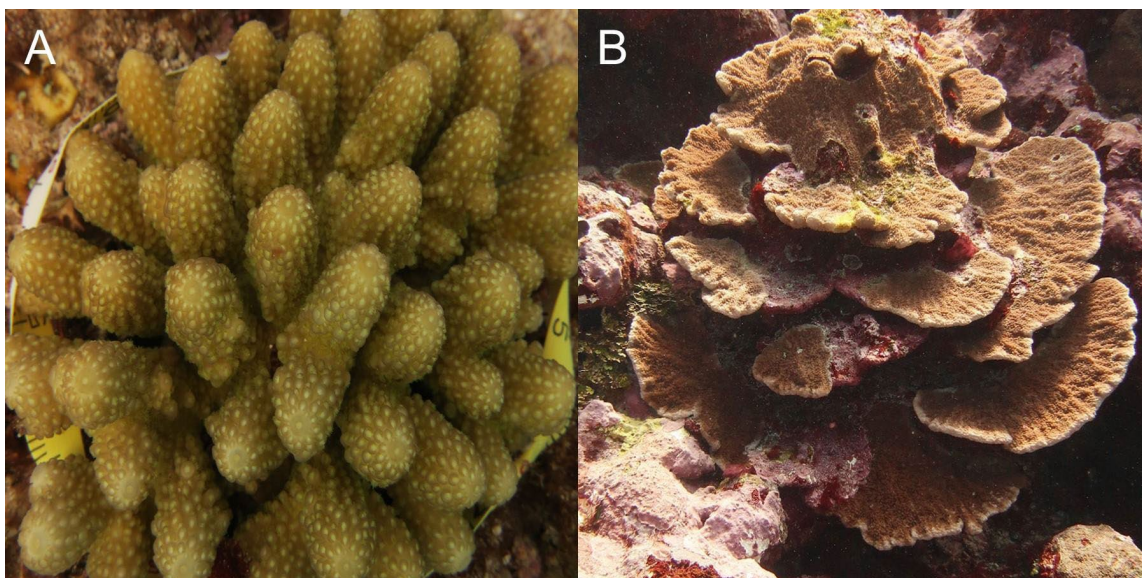


Figure 17. Examples of ESA listed taxa captured during SfM: A) *Acropora globiceps*, and B) *Isopora crateriformis*.

Discussion

Demographic metrics that show minimal to no bias

On the higher diversity reef systems of the Mariana Archipelago and American Samoa, the total colony density and the dominant coral genera, as well as adult colony diameter, showed the strongest correlation between methods and no significant bias between methods (Figures [7](#), [9](#), [12](#)). At the genus level, adult colony density was significantly higher in-water for 5 of the 49 observed adult genera: *Stylocoeniella* sp., *Echinophyllia* sp., *Psammocora* sp., *Leptoseris* sp., and *Astrea* sp. (comprising 6% of all adult colonies counted). These taxa with (which were almost completely missed in SfM) are often found in cryptic environments — within holes or crevices and beneath overhangs — that can be hard for SfM imagery and top-down orthomosaics to capture resulting in *Stylocoeniella* sp. and *Echinophyllia* sp. being almost missed completely in SfM. These results are consistent with our previous study in the main Hawaiian Islands (Couch et al., 2021) and suggest that transitioning to SfM would allow NCRMP to maintain continuity with our historical NCRMP data for adult density and maximum diameter for a majority of taxa.

Old partial mortality was significantly higher using SfM compared to in-water methods and there was no significant interaction between method and depth. However, similar to the main Hawaiian Islands study (Couch et al., 2021), the difference in old partial mortality estimates between methods was small ([Figure 10](#)). Interestingly, we documented a considerable improvement in the correlation and reduced variability between methods from the previous main Hawaiian Islands study. The improved alignment between methods suggests that the more comprehensive training and calibration (see [Methods](#) section *In Situ Collection*) deployed in this study meaningfully improved estimates of old partial mortality. Still, estimating old partial mortality can be challenging due to the binned nature of this metric (recorded in 5% increments) and the need for live tissue and partial mortality to sum to 100%. This latter estimation of old colony boundaries is particularly challenging in communities where corals are highly fragmented and/or branching. We recommend that enhanced training and calibration both in-water and behind a computer continue as a fundamental part of preparation for each monitoring season.

Identifying the boundaries of colonies is a fundamental challenge for these types of coral demographic surveys, regardless of whether surveys are conducted underwater or behind a computer. As colonial organisms, coral colonies often fragment into discrete tissue patches. NCRMP methods dictate that observers identify colonies by lumping together tissue fragments of a similar color and morphology on the original skeletal structure into one colony. Enumerating and sizing colonies can be challenging when

partial mortality is not recent and colonies are densely aggregated. The role of some level of observer error in these patterns is also supported by the fact that we reported that observer variability is comparable to differences between methods ([Figure 6](#)). We have demonstrated that SfM provides an opportunity to reduce observer error compared to divers by allowing annotators to discuss during annotation, revisit sites, and correct errors, which is not possible for in-water observers.

The results of both Couch et al. (2021) and the present study indicate no significant bias between methods in disease prevalence, but rather that assessing colony health consistently is challenging regardless of method. Previous studies have tested the comparability of in-water methods to imagery-based methods to extract coral health information. Contrary to our study, Page et al. (2016) found that disease prevalence recorded by in-water divers was six times higher compared to imagery at 12 coral reef sites in Australia. However, unlike our studies, Page et al. (2016) used photoquadrat images that were taken from one angle perpendicular to the colony surface, which likely resulted in missed lesions that would have been better captured by SfM given its 3D nature. Burns et al. (2020) used SfM to test the comparability of in-water and SfM methods on Hawai'i Island. At the colony-scale, they found that in-water assessments had a better ability than SfM to identify diseases and conditions correctly but a lower ability to detect colonies without a lesion (i.e., specificity). They conclude that given the modest magnitude of the difference between methods, in-water surveys should not be considered the gold standard method given its lower specificity, and SfM should be considered as a viable replacement for in-water observations. As suggested by Burns et al. (2020), inter-observer variability and the subjective nature of coral health assessments likely explain the overall weak correlation between the two methods for recent mortality and prevalence metrics. The weak relationship between the methods may also be related to suboptimal image quality at some sites, which impairs disease and lesion identification. Due to the generally uncommon nature of diseases, assessing patterns at the site-level for partial mortality and prevalence is often challenging, with many sites having a very low or zero prevalence of lesions.

Predation observations showed no significant bias between methods and, unlike acute disease prevalence, there was a strong correlation between methods ([Figure 11](#)). Burns et al. (2020) found that SfM assessments had the lowest ability to detect coral predation as lesions often form around the base and sides of colonies, which can be hard to see in imagery. However, coral predators often leave distinctive scars (Rotjan and Lewis, 2008) that can be consistently identified with the level of training required to participate in NCRMP surveys. Additionally, the use of *Viscore* in our SfM workflow allows for easy access to raw imagery from multiple angles, allowing observers to see parts of colonies not always visible from a single view.

Bleaching prevalence was also consistent across methods, which is contrary to our previous study that found significantly higher bleaching prevalence in SfM surveys across the main Hawaiian Islands. Unlike the previous study, we encountered low bleaching severity in the Mariana Archipelago and American Samoa. Our results agree with previous methodological comparison studies that found no difference in reported bleaching between in-water and imagery-based methods (Page et al., 2016; Burns et al., 2020). However, it appears that while the overall prevalence of bleaching was consistent between methods ([Figure 11K](#)), we did not record the same instances of bleaching as there was no correlation between methods ([Figure 11J](#)), pointing to the difficulty in assessing this metric in either survey.

Demographic metrics that continue to show distinctions across methods

Quantifying juvenile corals in imagery is challenging and often leads to underestimates relative to in-water surveys (Edmunds et al., 1998; Burgess et al., 2010). These patterns are likely a result of image quality combined with the challenge of quantifying juveniles given their small size and cryptic nature, sometimes preferring crevices and vertical surfaces to exposed substrates (Babcock and Mundy, 1996; Edmunds et al., 2004). While image quality has greatly improved in recent years, we still hypothesized that SfM may underestimate juvenile density. However, our present results indicate that juvenile density was significantly higher in SfM estimates than in-water. At the genus level, juvenile colony density was significantly higher for seven of the 39 observed juvenile genera (comprising 55% of all juvenile colonies counted), including the most common genera (*Porites* sp., *Favia* sp., *Leptastrea* sp., *Goniastrea* sp.; [Figure 13](#)).

While Couch et al. (2021) did not conduct an analysis of percent difference by genera, the previous study did find similar variability and methodological bias in some of the dominant juvenile genera. The higher density in SfM and higher observer error overall may be a result of difficulty in determining colony boundaries of juveniles, especially in taxa that are prone to high levels of fragmentation and partial mortality. Studies using SfM to examine coral juvenile demographics have excluded certain encrusting taxa such as *Montipora* sp. and *Pavona* sp., as they could not be sure if colonies were true recruits or resulting from fragmentation or partial mortality (Pedersen et al., 2019; Sarribouette et al., 2022). The inherent ability of SfM to more precisely measure colony size may also have contributed to the differences. In-water observers appeared to record juvenile colonies >4.5 cm as adults more frequently than SfM annotators, since we had higher densities of juvenile corals greater than 4.5 cm in SfM surveys than in in-water surveys. In light of the methodological bias in juvenile populations, we are concerned about maintaining continuity with historical data and discuss several strategies for capturing juvenile communities in the [Recommendations](#) section below.

Adult genus level diversity was significantly higher in-water compared to SfM surveys. Conversely, juvenile genus level diversity was not significantly different between methods for any diversity metric. Juvenile corals are generally hard to identify to a high taxonomic resolution both underwater and in SfM, which may explain the lack of bias towards one method. This methodological bias seen in adult diversity is in contrast to results from the lower diversity reefs in the main Hawaiian Islands, where Couch et al. (2021) found no significant bias in diversity between methods. Further, we did not detect a significant bias in adult genus richness between methods on these higher diversity reefs suggesting that colony evenness is not similar across methods and is the source of the diversity distinctions.

Evenness was primarily affected by the three most abundant taxa, *Porites* sp., *Astreopora* sp., and *Montipora* sp. and when these taxa were removed, we no longer detected methodological bias in Shannon-Weiner diversity (LRT: 2.11, p -value: 0.147). On the other hand, when rare and cryptic taxa were removed, including up to half of the rarest taxa, the methodological difference persisted. This cross-method distinction appears to derive from higher colony counts of common taxa in SfM data, and we further hypothesize that this bias is linked to differences in how observers identified colony boundaries ([Figure 16](#)).

The “lumper/splitter” challenge (whether to include distinct patches of live coral as a single, recorded colony) is a persistent issue in coral reef colony-focused surveys. In NCRMP benthic rapid ecological assessment (REA) surveys, we visit a spatially random site a single time and cannot rely on temporal change to clarify whether neighboring tissue patches are fragmented clones of a single colony. However, given that the surveys are designed to assess partial mortality, observers are required to measure colony size including living and dead areas. Faced with an aggregation of coral patches, reasonable experts can justifiably disagree on the “correct” answer to lumping/splitting of colony boundaries. This pattern could be responsible for the relative over-count of colonies in these common genera and could reasonably generate the distinctions in diversity ([Figure 14](#)). Therefore, as a program, we train for cross-annotator consistency by applying distinct rules based on coral life history knowledge and repeatedly cross-validating our annotators.

When comparing inter-observer error for all transects across regions, there was a decrease (based on 95% CI overlap) in MS-MAE in previous surveys in Hawai'i (Couch et al. 2021) and this study for adult colony density, average colony diameter, average old and recent partial mortality, and adult genus Shannon diversity. MS-MAE did not change between regions for juvenile colony density prevalence metrics (acute disease, chronic disease, and bleaching) when comparing surveys from the Hawaiian Islands to the higher diversity Mariana Archipelago and American Samoa regions ([Table 1](#)).

Table 2. Direction of bias and MS-MAE for Couch et al. (2021) and the present study for all transects. A decrease in MS-MAE indicates lower inter-observer variability.

Response Variable	Couch et al. (2021)		Present Study		Change in MS-MAE
	Bias	MS-MAE	Bias	MS-MAE	
Adult Colony Density	None	0.233	None	0.186	Decrease
Juvenile Colony Density	None	0.428	Higher in SfM	0.455	None
Average Colony Diameter	None	0.171	None	0.112	Decrease
Average Old Mortality	Higher in SfM	0.365	Higher in SfM	0.221	Decrease
Average Recent Mortality	None	1.253	Higher in-water	0.686	Decrease
Acute Disease Prevalence	None	1.424	None	1.699	None
Chronic Disease Prevalence	None	1.017	None	1.177	None
Bleaching Prevalence	Higher in SfM	0.904	None	1.403	None
Adult Richness	None	0.224	None	0.176	None
Adult Shannon Diversity	None	0.225	Higher in-water	0.108	Decrease

As such, a focus on training and cross-validating between observers, especially on colony boundary definitions, has helped reduce some of this bias. In particular, we have already seen improvements in reduced variability among observers in lumping/splitting distinctions when we focus training on consistent lumping/splitting rules and cross-annotator feedback (i.e., where each annotator stands on lumping/splitting relative to their peers). To address this issue, we plan to make this training and feedback more frequent and consistent to maintain consistency across annotators and compatibility with our existing time series.

Conclusions & Recommendations

The application of imaging technology for coral reef monitoring is growing rapidly, but the question remains: can monitoring programs maintain time series historically collected with in-water methods? Our method comparison studies across low and high diversity reef systems indicate that demographic metrics such as total adult density, adult size, and partial mortality show very low to no bias between methods overall or when interacting with depth and habitat type (Couch et al., 2021; current study). Disease and bleaching prevalence, although weakly correlated between methods due to continued issues with inter-observer variability, also show low to no bias. Furthermore, the level of variability between methods for these metrics is comparable to the variability we normally see between divers. However, on the more diverse reefs of the Mariana Archipelago and American Samoa, total and genus-level juvenile colony density, as well as adult diversity, show significant bias. This bias may prevent us from maintaining continuity with historical NCRMP data for these metrics if NCRMP chooses to switch to SfM.

Both in-water and SfM survey methods have a variety of strengths and weaknesses, and which method is chosen depends on the research question and logistical constraints at hand (details provided in Couch et al., 2021). Monitoring coral demography with SfM provides a number of advantages over in-water methods, such as reduced in-water survey effort, finer-scale and more accurate measurements of structural complexity, and the ability to re-evaluate SfM imagery to control for observer error. However, SfM requires a considerable amount of time to extract data relative to in-water methods. It carries the risk of taking poor-quality imagery, resulting in poor data quality, and SfM's inability to capture all the crevices and overhangs that can be observed underwater may limit operationalizing this method.

The feasibility of replacing standard in-water coral demographic surveys with SfM surveys is dependent on being able to maintain continuity with legacy NCRMP data. Enabling continuity includes maintaining high standards for training to minimize inter-annotator variability, exploring a hybrid approach to benthic monitoring, and balancing field and post-processing costs, which are expanded on below:

Dedicated SfM imaging team. To more effectively collect high-quality imagery, we propose that ESD dedicates one boat to capturing SfM imagery, whenever our field allocations permit. This would allow divers to focus on collecting the highest quality image data without the time and air constraints of conducting other surveys concurrently with SfM surveys. The added value of a dedicated imaging team is the opportunity to record natural history information on aspects such as difficult taxonomic, morphological, or coral health information that will be invaluable to the annotation team. We estimate that each SfM-only dive would be 20 minutes long and would allow a three-person SfM

team to visit six sites per day. As efficiency requirements will likely dictate that we maintain a joint Fish-SfM survey effort in some contexts, we also will focus on training and any methodological improvements to ensure that fish divers can collect the highest quality imagery possible.

Well-calibrated annotation team. We also recommend that NCRMP continue efforts to incorporate enhanced field and in-water observer training to ensure consistency between annotators. Improvements will continue to focus on not only high-quality taxonomic identification, recent mortality, and disease training but also reaching cross-annotator consistency with lumping and splitting across colony boundaries. This includes data collection necessary to measure cross-annotator consistency, which is compiled and reported to annotators throughout the annotation process.

Supplemental higher quality imaging within segments. In addition to a more specialized imaging team, we also suggest that the divers capture not only full transect imaging at 1 m above substrate (~0.2 mm ground resolution), but also in the adult and juvenile segment areas, collect supplemental imagery that is higher quality, closer to substrate (~0.5m), with more representative angles. This “fill-in” imagery will help resolve difficult taxonomic differences, better capture complex 3D structures, and potentially resolve some of the issues present in juvenile surveys.

Hybrid approaches. One method that would both prioritize metric quality and balance in-field and annotation costs is applying the SfM approach with supplemental in-situ benthic monitoring. This would involve extracting metrics—such as density, colony size, partial mortality, and incidences of disease, bleaching, and other compromised health states from SfM—but recording juvenile colony density using in-water observations. This would allow ESD to use SfM to minimize field time while using in-water surveys to maintain continuity in assessments of juvenile coral density. We estimate that divers would still be able to survey six sites per day so field costs would not change from what is proposed in [Appendix C](#). However, we estimate that it would save SfM annotators one hour of data extraction time for each site surveyed and improve continuity with our historical data set. While the hybrid method has not been tested and personnel costs could vary, we estimate that this would save an hour of manual annotation per site as proposed in [Appendix C](#). An additional limitation of this approach is that it still requires at least one highly trained benthic specialist in the field.

Strive for high image quality. It is paramount that divers continue collecting quality imagery using guidelines listed in the SOP (Torres-Pulliza et al. 2023) and Appendices. In addition to properly color-balanced and sharp images, our annotations would also benefit from shooting from more than one angle. This would allow divers to capture colonies on vertical or concave surfaces and reduce gaps in the model. When time permits, we recommend that plots with high rugosity be photographed from a minimum

of two angles. Taking extra care when photographing these sites will ensure all surfaces are captured. If the hybrid approach discussed above is not possible, then close-up imagery (0.5 m off the benthos) of coral segments drastically improves the ability of observers to capture true juvenile colonies.

Improved efficiency of data extraction. One of the largest hurdles to overcome with SfM is the significant annotation and post-processing time and effort necessary to extract data. In the short-term, we recommend developing a tool to allow annotators to directly enter SfM data into the established benthic database from ArcPro rather than entering data into a geodatabase. This will allow NCRMP to leverage the efficient data entry tool already used for in-water surveys and drastically reduce time spent correcting database QC errors. In the long-term, we recommend working with the computer science community to develop ‘human-in-the-loop’ annotation tools to reduce manual annotation time (see section *Increasing efficiency through innovation and technology* in Couch et al. 2021 for in-depth discussion).

Acknowledgements

We would like to thank the crew of the NOAA Ship *Rainier* for providing field support during the Mariana Archipelago and American Samoa RICHARD cruises. We would also like to thank Michael Akridge and Lori Luers for their assistance with development and maintenance of the data management pipeline. We thank Stuart Sandin, John Burns, Nicole Pederson, Vid Petrovic, and Clinton Edwards for their support with field testing, method development, and implementation of SfM. This work is funded by the NOAA Coral Reef Conservation Program as part of the Pacific National Coral Reef Monitoring Program (project # 31360 and 743).

Literature Cited

Babcock R, Mundy C. (1996). Coral recruitment: Consequences of settlement choice for early growth and survivorship in two scleractinians. *Journal of Experimental Marine Biology and Ecology*, 206(1), 179–201. doi:10.1016/S0022-0981(96)02622-6.

Bayley D.T.I., Mogg A, Koldewey H, Purvis A. (2019). Capturing complexity: field-testing the use of ‘structure from motion’ derived virtual models to replicate standard measures of reef physical depth and structure. *PeerJ*. 7:e6540. doi:10.7717/peerj.6540.

Burgess S.C., Osborne K, Sfiligoj B, Sweatman H. Can juvenile corals be surveyed effectively using digital photography?: implications for rapid assessment techniques. *Environmental Monitoring and Assessment*, 171(1-4), 345-51. doi: 10.1007/s10661-009-1282-1.

Burns J.H.R., Delparte D, Gates R, Takabayashi M. (2015). Integrating structure-from-motion photogrammetry with geospatial software as a novel technique for quantifying 3D ecological characteristics of coral reefs. *PeerJ*.:3:e1077. doi:10.7717/peerj.1077.

Burns J.H.R., Weyenberg G, Mandel T, Ferreira S.B., Gotshalk D, Kinoshita C.K., Marshall M.J., Del Moral N.A., Murphy S.J., Pascoe K.H., et al. (2020). A Comparison of the Diagnostic Accuracy of in-situ and Digital Image-Based Assessments of Coral Health and Disease. *Front Mar Sci*. 7:304. doi:10.3389/fmars.2020.00304.

Casella E, Collin A, Harris D, Ferse S, Bejarano S, Parravicini V, Hench J.L., Rovere A. (2017). Mapping coral reefs using consumer-grade drones and structure from motion photogrammetry techniques. *Coral Reefs*, 36(1), 269–275. doi:10.1007/s00338-016-1522-0.

Couch C.S., Oliver T.A., Suka R, Lamirand M, Asbury M, Amir C, Vargas-Ángel B, Winston M, Huntington B.E., Huntington B, Lichowski F, Halperin A, Gray A.K., Garriques J, Samson J. (2021). Comparing Coral Colony Surveys From In-Water Observations and Structure-From-Motion Imagery Shows Low Methodological Bias. *Frontiers in Marine Science*. DOI: 10.3389/fmars.2021.647943.

Edmunds P, Aronson R, Swanson D, Levitan D, Precht W. (1998). Photographic versus visual census techniques for the quantification of juvenile corals. *Bulletin Marine Science*, 62, 937–946.

Edmunds P.J., Bruno J, Carlon D. (2004). Effects of depth and microhabitat on growth and survivorship of juvenile corals in the Florida Keys. *Marine Ecological Progress Series*, 278, 115–124. doi:10.3354/meps278115.

Edwards C.B., Eynaud Y, Williams G.J., Pedersen N.E., Zgliczynski B.J., Gleason A.C.R., Smith J.E., Sandin S.A. (2017). Large-area imaging reveals biologically driven non-random spatial patterns of corals at a remote reef. *Coral Reefs*, 36(4), 1291–1305. doi:10.1007/s00338-017-1624-3.

Ferrari R, Figueira W.F., Pratchett M.S., Boube T, Adam A, Kobelkowsky-Vidrio T, Doo

S.S., Atwood T.B., Byrne M. (2017). 3D photogrammetry quantifies growth and external erosion of individual coral colonies and skeletons. *Science Report*, 7(1), 16737. doi:10.1038/s41598-017-16408-z.

Figueira W, Ferrari R, Weatherby E, Porter A, Hawes S, Byrne M. (2015). Accuracy and Precision of Habitat Structural Complexity Metrics Derived from Underwater Photogrammetry. *Remote Sensing*, 7, 16883–16900. doi:10.3390/rs71215859.

Fox M.D., Carter A.L., Edwards C.B., Takeshita Y, Johnson M.D., Petrovic V, Amir C.G., Sala E, Sandin S.A., Smith J.E. (2019). Limited coral mortality following acute thermal stress and widespread bleaching on Palmyra Atoll, central Pacific. *Coral Reefs*, 38(4), 701–712. doi:10.1007/s00338-019-01796-7.

Fukunaga A., Burns J.H.R., Pascoe K.H., Kosaki R.K. (2020). Associations between Benthic Cover and Habitat Complexity Metrics Obtained from 3D Reconstruction of Coral Reefs at Different Resolutions. *Remote Sensing*, 12, 1011. doi:10.3390/rs12061011.

Lange I.D., Perry C.T. (2020). A quick, easy and non-invasive method to quantify coral growth rates using photogrammetry and 3D model comparisons. *Methods Ecological Evolution*, 11(6), 714–726. doi:10.1111/2041-210X.13388.

[PIFSC] Pacific Islands Fisheries Science Center. (2015). Pacific Reef Assessment and Monitoring Program: Belt Transect Surveys for Demographics and Condition of Corals at U.S. Pacific Reefs from 2007 to 2012. <https://inport.nmfs.noaa.gov/inport/item/25260>.

[PIFSC] Pacific Islands Fisheries Science Center ESD. (2020a). Pacific Reef Monitoring and Assessment Program: Line-Point-Intercept (LPI) Surveys for Benthic Cover at Coral Reef Sites across the Pacific Ocean from 2005 to 2012. <https://www.fisheries.noaa.gov/inport/item/25261>.

[PIFSC] Pacific Islands Fisheries Science Center ESD. (2020b). National Coral Reef Monitoring Program: Benthic Images Collected from Stratified Random Sites (StRS) across the Hawaiian Archipelago from 2013 to 2019. <https://www.fisheries.noaa.gov/inport/item/25247>.

Obura D, Aeby G, Amornthammarong N, Appeltans W, Bax N, Joe B, Brainard R.E., Chan S, Pamela F, Gordon T.A.C., et al. (2019). Coral Reef Monitoring, Reef Assessment Technologies, and Ecosystem-Based Management. *Frontiers Marine Science*, 6, 580. doi:10.3389/fmars.2019.00580.

Page C.A., Field S.N., Pollock F.J., Lamb J.B., Shedrawi G, Wilson S.K. (2016). Assessing coral health and disease from digital photographs and in situ surveys. *Environmental Monitoring Assessment*, 189(1), 18. doi:10.1007/s10661-016-5743-z.

Palma N. (2016). 3D Modeling of Coral Reefs Using Structure-from-Motion Photogrammetry as an Inventory and Monitoring Tool. https://scholarship.miami.edu/discovery/fulldisplay/alma991031447804602976/01UOML_INST:ResearchRepository.

Pedersen N.E., Edwards C.B., Eynaud Y, Gleason A.C.R., Smith J.E., Sandin S.A. (2019). The influence of habitat and adults on the spatial distribution of juvenile corals. *Ecography (Cop)*, 42(10), 1703–1713. doi:10.1111/ecog.04520.

Sarribouette L, Pedersen N.E., Edwards C.B., Sandin S.A. (2022). Post-settlement demographics of reef building corals suggest prolonged recruitment bottlenecks. *Oecologia*. 199(2), 387-396. doi: 10.1007/s00442-022-05196-7.

Storlazzi C.D., Dartnell P, Hatcher G.A., Gibbs A.E. (2016). End of the chain? Rugosity and fine-scale bathymetry from existing underwater digital imagery using structure-from-motion (SfM) technology. *Coral Reefs*, 35(3). doi:10.1007/s00338-016-1462-8.

Suka R, Asbury M, Gray A.E., Winston M, Oliver T, Couch C.S. (2019). Processing Photomosaic Imagery of Coral Reefs Using Structure-from-Motion Standard Operating Procedures. <https://doi.org/10.25923/h2q8-jv47>.

Torres-Pulliza D, Dornelas M.A., Pizarro O, Bewley M, Blowes S.A., Boutros N, Brambilla V, Chase T.J., Frank G, Friedman A, et al. (2020). A geometric basis for surface habitat complexity and biodiversity. *Nat Ecol Evol* 4, 1495–1501 <https://doi.org/10.1038/s41559-020-1281-8>.

Voss J, Shilling E, Combs I. (2019). Intervention and fate tracking for corals Florida, affected by stony coral tissue loss disease in the northern Florida Reef Tract. [https://floridadep.gov/sites/default/files/Voss SEFL Disease Intervention 2019_FDEP FINAL Report.Fully508compliant.pdf](https://floridadep.gov/sites/default/files/Voss%20SEFL%20Disease%20Intervention%202019_FDEP_FINAL%20Report.Fully508compliant.pdf).

Zuur A.F., Ieno E.N., Walker N.J., Anatoly A. Saveliev, Smith G.M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. Springer Science & Business Media.

Appendices

Appendix A. Table of statistical results for methodological bias. Bold indicates fixed effect is significant ($p < 0.05$).

Response Variable	Test	Fixed effect	Likelihood Ratio	p
Adult Colony Density	LMM	Method × Depth	0.016	0.899
		Method	3.569	0.059
Juvenile Colony Density	LMM	Method × Depth	0.013	0.908
		Method	11.276	0.000
Average Colony Diameter	LMM	Method × Depth	0.163	0.687
		Method	3.449	0.063
Average Old Mortality	LMM	Method × Depth	0.479	0.489
		Method	6.426	0.011
Adult Shannon Diversity	LMM	Method × Depth	0.295	0.587
		Method	19.888	0.000
Adult Hill Diversity	LMM	Method × Depth	0.115	0.735
		Method	20.774	0.000
Adult Richness	LMM	Method × Depth	0.543	0.461
		Method	0.054	0.817
Juvenile Shannon Diversity	LMM	Method × Depth	0.137	0.711
		Method	0.006	0.936
Juvenile Hill Diversity	LMM	Method × Depth	0.032	0.857
		Method	0.076	0.783
Juvenile Richness	LMM	Method × Depth	0.154	0.695
		Method	0.103	0.745
Average Recent Mortality	Wilcoxon	Method	-	0.009
Acute Disease Prevalence	Wilcoxon	Method	-	0.083
Predation Prevalence	Wilcoxon	Method	-	0.296
Chronic Disease Prevalence	Wilcoxon	Method	-	0.592
Bleaching Prevalence	Wilcoxon	Method	-	0.962

Appendix B. Lookup table of genus codes used by Pacific NCRMP.

Genus Code	Genus Name	Genus Code	Genus Name	Genus Code	Genus Name
ACAS	<i>Acanthastrea sp</i>	GASP	<i>Galaxea sp</i>	PACS	<i>Pachyseris sp</i>
ACSP	<i>Acropora sp</i>	GONS	<i>Goniastrea sp</i>	PAVS	<i>Pavona sp</i>
ALSP	<i>Alveopora sp</i>	GOSP	<i>Goniopora sp</i>	PESP	<i>Pectinia sp</i>
ASSP	<i>Astreopora sp</i>	HASP	<i>Halomitra sp</i>	PHSP	<i>Phymastrea sp</i>
ASTS	<i>Astrea sp</i>	HERS	<i>Herpolitha sp</i>	PLER	<i>Plerogyra sp</i>
CASP	<i>Caulastrea sp</i>	HESP	<i>Heliopora sp</i>	PLES	<i>Plesiastrea sp</i>
COSP	<i>Coscinaraea sp</i>	HYSP	<i>Hydnophora sp</i>	PLSP	<i>Platygyra sp</i>
CTSP	<i>Ctenactis sp</i>	ISSP	<i>Isopora sp</i>	POLY	<i>Polyphilia sp</i>
CYPS	<i>Cyphastrea sp</i>	LEPS	<i>Leptoria sp</i>	POSP	<i>Porites sp</i>
CYSP	<i>Cycloseris sp</i>	LEPT	<i>Leptastrea sp</i>	PSSP	<i>Psammocora sp</i>
DISP	<i>Diploastrea sp</i>	LESP	<i>Leptoseris sp</i>	SCAS	<i>Scapophyllia sp</i>
ECHL	<i>Echinophyllia sp</i>	LOBS	<i>Lobophyllia sp</i>	SESP	<i>Seriatopora sp</i>
ECHP	<i>Echinopora sp</i>	MESP	<i>Merulina sp</i>	STSP	<i>Stylocoeniella sp</i>
EUSP	<i>Euphyllia sp</i>	MISP	<i>Millepora sp</i>	STYP	<i>Stylaraea sp</i>
FASP	<i>Favia sp</i>	MOSP	<i>Montipora sp</i>	STYS	<i>Stylophora sp</i>
FAVS	<i>Favites sp</i>	MYSP	<i>Mycedium sp</i>	SYSP	<i>Symphyllia sp</i>
FUSP	<i>Fungia sp</i>	OUSP	<i>Oulophyllia sp</i>	TURS	<i>Turbinaria sp</i>
GARS	<i>Gardineroseris</i>	OXSP	<i>Oxypora sp</i>	TUSP	<i>Tubastraea sp</i>

Appendix C. Cost-Benefit Analysis

In this section, we provide a breakdown and comparison of the estimated time and hardware costs of generating the standard suite of coral demographic and benthic metrics each year for in-water surveys and SfM surveys (Tables [1](#), [2](#)). These costs have been tailored to meet NCRMP programmatic requirements. Please contact the authors if you are an external partner interested in identifying costs specific to your program. For the purposes of this analysis, we assume that data will be collected from an average of 240 sites per year to meet the average NCRMP benthic allocations. The time estimates provided in [Table 1](#) reflect improvements in efficiency and data extraction from the previous study (Couch et al., 2021) and involve a mixture of well-trained and new staff. We assume that we will have the same mixture of well-trained and new staff for both methods.

Personnel Cost

Training

Training is an important component of benthic surveys. It takes time to develop a team of benthic analysts that is consistent and accurate. The amount of time to train staff is equivalent between the two methods. Weeks of classroom and in-water training and calibration are required for both methods prior to each cruise. SfM image collection training can be conducted in just a few hours and involves staff not trained in benthic monitoring. However, SfM annotation and in-water surveys require the same in-depth training in taxonomic identification and demographic methods both in-water and in the classroom. In 2022 and 2023, we updated our training materials and instituted more rigorous calibration procedures, which reduced the inter-diver error, leading to reduced QC time and more timely data extraction from SfM.

Field Time

In-water surveys require on average 45 minutes with three divers to complete, compared to an average of 25 minutes with two divers to complete a SfM survey. In-water surveys require a three-diver team with one backup diver that will need a total of 69 days to complete 240 sites and a total of 2,208 personnel hours. This total time estimate does not include non-dive days associated with a cruise. SfM requires a 2-diver team with 1 backup diver and would require 40 days to complete and 960 personnel hours. Note, with improvements to training, we are able to survey more sites per day compared to the previous study. Overall when comparing field time, transitioning to SfM results in a 56% reduction in field time over in-water surveys.

Processing Time

Processing time is substantial for SfM and minimal for in-water methods. In 2022, we deployed a Python script in Agisoft Metashape to automate the generation and scaling of dense point clouds and digital elevation models, which has decreased hands-on time by 35%. The hands-on processing time includes one personnel hour per site to run the Python script, engage in several human-in-the-loop checkpoints, and set up the survey area in ArcPro. A majority of the processing time per site (6.5 hours) is required to manually delineate colonies in ArcPro and QC delineations. With the processing infrastructure we have developed within the Pacific Islands Fisheries Science Center, we are able to generate eight models simultaneously in approximately 17 hours (hands-off time). It will take approximately three weeks of time to generate the 3D and 2D products from 240 sites on our most powerful servers (assuming continuous generation of models) and 1800 personnel hours to annotate these sites using 100% manual annotation.

Data Summary and Archiving Time

Assessing data quality and generating standard data summaries is conducted in R and requires the same amount of time regardless of method. Archiving benthic data is largely the same for both methods, with the exception of archival of SfM imagery in NCEI, which requires an additional 80 hours of personnel time.

Summary of Personnel Costs

In summary, tasks such as training, generation of data summaries, and data archiving are very similar between the two methods. The primary difference between these two methods is seen in the field and processing time. Field costs are reduced by 56% by switching to SfM under the proposed model. However, this method requires an estimated 2,040 personnel hours of processing time, which largely does not exist for in-water surveys. Overall, this would result in a 17% increase in personnel hours by switching from in-water to SfM, which is down from the 35% increase published in Couch et al. (2021). These efficiency gains were the result of the implementation of the automated model processing script and improved benthic training. Given the rapidly advancing field of SfM and artificial intelligence (AI) and our desire to reduce the human annotation burden, NCRMP scientists are continuing to partner with computer scientists to improve the efficiency of the SfM processing pipeline (see Section below on *Increasing efficiency through innovation and technology*).

Table C1. A comparison of the average hands-on time, per year, for field data collection and data processing between in-water and SfM using estimates. Annual time estimates include time to generate data for 240 benthic sites. These estimates do not include time for training and calibration, which is consistent across methods.

Field time/site & year	In-water	SfM
Bottom time (h)	0:45	0:25
# divers needed/site*	4	3
# of sites/day	3.5	6
# of dive days/year	69	40
Total field hours/year**	2,208	960
Processing time/site & year	In-water	SfM
Hands-on hours/site to generate 3D and 2D products (h)	0	1:00
Demographic data extraction & QC hands-on/site (h)	0:30	6:30
Total data archival hours/year	160	240
Total hands-on personnel hours/year	280	2,040
Overall personnel hours for field collection and data extraction	2,488	3,000

*Includes rotating topside diver

** Total field hours/year = (240 sites/ # sites per day) x # of divers x 8 hours

Equipment Cost

Equipment costs are associated with both methods, but the equipment costs for SfM is approximately 19 times higher than the in-water method. [Table 2](#) provides a comparison of costs for each method divided by general category and initial, annual, and every five-year purchases. For camera purchases, in-water surveys require an initial purchase of point-and-shoot cameras and housings to photograph colonies that divers have questions about and require an annual purchase of one new camera or housing. For SfM, four new Canon Rebel SL2/3 cameras and housings are initially required. We estimate that half of the camera shutters will need to be rebuilt each year, and all cameras and housings will need replacement every fifth year. Field gear SfM costs are higher for one-time purchases of markers not associated with in-water surveys. The most expensive components of equipment costs for SfM are computers and servers. Computer costs for in-water surveys are minimal, with two backup hard drives purchased every five years. Computer costs for SfM are substantially higher with an initial purchase of a processing server with GPU acceleration, a Synology server for cruise data storage, an in-house storage server, one high-powered workstation for

annotation, backup hard drives, monitors, and six Agisoft licenses. The annual cost includes additional hard drive storage and replacement parts for computers. These items will continue to improve in performance and drop in price over time. Because the actual lifespan of these items can be variable, a complete replacement for these systems is included in the five-year cost estimate, but it is possible that the equipment may last longer than five years.

In addition, a larger proportion of this work and these costs will be performed in the NMFS cloud system, likely lowering costs overall and pushing the expenses to broader institutional budgets. That said, we are still in the process of running cloud cost models, and the actual savings are as of yet unclear.

Table C2. List of equipment and cost for in-water and SfM data collection and processing.

Item	In-water Cost (\$)	SfM Cost (\$)
Cameras and accessories		
Initial	4760.00	10,492.00
Each year	644.00	1,124.00
Every 5 years*	0	7,036.00
Field Gear		
Each year	235.00	240.00
Every 5 years*	445.00	435.00
Computers		
Initial	200.00**	41,306.0
Each year	0	7,500.00
Every 5 years*	200.00	52,718.00
<i>Total</i>	6,284.00	120,851.00

* 5-year costs include annual costs.

**Initial computer costs for in-water surveys do not include laptops that all ESD staff are issued and are needed to summarize data.