UNIVERSITY OF CALIFORNIA SAN DIEGO

Experiments in optical data collection, processing, and analysis for ocean science

A Dissertation submitted in partial satisfaction of the requirements

for the degree Doctor of Philosophy

in

Oceanography

by

Hugh Runyan

Committee in charge:

Professor Falko Kuester, Co-Chair
Professor Stuart Sandin, Co-Chair
Professor Dariusz Stramski, Co-Chair
Professor David Kriegman
Professor Hao Su

2023

The Dissertation of Hugh Runyan is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

To my family, partner, and friends for their unwavering support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisors, Stuart Sandin and Falko Kuester, for years of support and generosity with their time and expertise, along with the entirety of their labs, who have been friendly and knowledgeable colleagues. I would particularly like to thank Clinton Edwards, Nicole Pedersen, and Vid Petrovic for their regular collaboration.

I would also like to thank Dariusz Stramski for admitting me and advising me during my first two years at Scripps. I would also like to thank the rest of his lab, particularly Rick Reynolds. Dariusz and Rick taught me how to make precise lab measurements in the field, how to analyze those measurements, and how to prepare top quality scientific manuscripts. My first chapter, and first publication, is based heavily on measurements they have collected on numerous cruises over the past few decades.

Additionally, I would like to thank the other members of my committee—David Kriegman and Hao Su—for their guidance, assistance, and support. David Kriegman is responsible for the first AI coral imagery annotation tool (CoralNet), which was an inspiration for much of my work, and was the basis for my fourth dissertation chapter. Hao Su has done some incredible work in the domain of 3D neural networks, which was a primary source and inspiration for my own work with 3D neural networks.

Chapter 1, in full, is a reprint of material as it appears in *Journal of Geophysical Research: Oceans*, Runyan, H.; Reynolds, R. A.; Stramski, D.; 2020. The dissertation author was the primary investigator and author of this paper.

Chapter 2 contains material being prepared for publication, coauthored with Petrovic, V.; Edwards, C. B.; Pedersen, N.; Alcantar, E.; Kuester, F.; Sandin, S. A. The dissertation author was the primary researcher and author of this material.

Chapter 3, in full, is a reprint of material as it appears in *Frontiers in Robotics and AI*, Runyan, H.; Petrovic, V.; Edwards, C. B.; Pedersen, N.; Alcantar, E.; Kuester, F.; Sandin, S. A; 2022. The dissertation author was the primary investigator and author of this paper.

Chapter 4 contains material being prepared for publication coauthored with Petrovic, V.; Edwards, C. B.; Pedersen, N.; Lubarsky, K.; Holloway, N.; Kuester, F.; Sandin, S. A. The dissertation author was the primary researcher and author of this material.

VITA

2011 Bachelor of Science in Mechanical Engineering, Duke University

2023 Doctor of Philosophy in Oceanography, University of California San Diego

PUBLICATIONS AND CONFERENCE PRESENTATIONS

Runyan, H., Petrovic, V., Edwards, C. B., Pedersen, N. E., Alcantar, E., Kuester, F., Sandin, S. A. 2022. Automated 2D, 2.5D, and 3D segmentation of coral reef pointclouds and orthoprojections. *Frontiers in Robotics and AI.* 9:884317. doi: 10.3389/frobt.2022.884317

Runyan, H., Petrovic, V., Edwards, C. B., Pedersen, N. E., Alcantar, E., Kuester, F., Sandin, S. A. 2021 and 2022. Experiments in automation to expedite high-precision 2D/3D coral segmentation. 15th International Coral Reef Symposium (Bremen, Germany).

Runyan, H., Petrovic, V., Edwards, C. B., Pedersen, N. E., Kuester, F., Sandin, S. A. 2021. Automation-assisted segmentation to expedite 3D coral mapping. 2nd NOAA AI Workshop (virtual).

Runyan, H., Petrovic, V., Edwards, C. B., Pedersen, N. E., Kuester, F., Sandin, S. A. 2020. Investigation of 3D, 4D, and hybrid automated methods to expedite high-precision coral segmentation. Ocean Sciences Meeting (San Diego, California).

Runyan, H., R. A. Reynolds, D. Stramski. 2020. Evaluation of Particle Size Distribution Metrics to Estimate the Relative Contributions of Different Size Fractions based on Measurements in Arctic Waters. *J. Geophys. Res. Oceans* 125, e2020JC016218. doi: 10.1029/2020JC016218

ABSTRACT OF THE DISSERTATION

Experiments in optical data collection, processing, and analysis for ocean science

by

Hugh Runyan

Doctor of Philosophy in Oceanography

University of California San Diego, 2023

Professor Falko Kuester, Co-Chair,

Professor Stuart Sandin, Co-Chair

Professor Dariusz Stramski, Co-Chair

Improving oceanographic data collection involves two components: improving instrumentation, and improving the processing and analysis of resulting measurements. While advancing technology has improved and expanded data collection, processing these data has become a significant problem for under-resourced academic labs, rendering much of these data underutilized. The experiments described in this dissertation are contributions to the effort of

improving data processing in oceanography. Chapter 1 examines parametrization of particle size distribution measurements in the Arctic, collected during decades of field expeditions by Dariusz Stramski, Rick Reynolds, and others. Experiments suggest that the commonly-used Junge-type power law model for parametrizing particle size distributions is insufficient, and cumulative distribution functions may offer a superior alternative. Particle size directly physically affects light, so the particle size distribution affects the signals of optical instruments. These results will increase the utility of satellite imagery, both by assisting the measurement of particle size from satellites, and by improving understanding of the impact of different seawater characteristics on optical signals. Chapter 2 discusses cutting-edge, technology-enabled survey image and 3D model techniques for studying coral reefs. The report focuses on lessons learned by the Sandin/Kuester labs from a decade of experience. Improving and standardizing data collection in the field allows research groups to pool datasets and compare results. Pooled databases are valuable for developing processing and analytical tools like neural networks, make those tools useful to more research groups, and enable ecological analysis at larger scales. Chapters 3 and 4 evaluate neural-network-assisted tools that can expedite taxonomic labeling of survey image products such as orthoprojections and pointclouds. Chapter 3 contains the first published investigation of 3D coral pointcloud segmentation with 3D neural networks. This new technology shows great promise, as evidenced by the time savings (36%) and high prediction accuracy in some contexts (~70-90%) recorded by our experiments, but requires more comprehensive, standardized datasets than what is currently available to fully develop or confidently evaluate. These tools are already capable of increasing the utility of survey imagery by expediting their annotation, which is necessary for many types of analysis, and further development will further improve their performance.

INTRODUCTION TO THE DISSERTATION

I will start with a cliché because it is one of the true ones: Earth's oceans are very, very big. Ships are expensive and slow, and anything that goes under the water is even slower. It takes a research ship in the region of 2-4 weeks to cross the full width of the Pacific Ocean, and that is without stopping to do any science. Any measurements from the same day, let alone the same morning or tide or etc, must be taken within a few miles – contemporaneous large-area datasets are nearly impossible to create from a ship. Human divers are additionally limited by available underwater instrumentation and human physiology. For these reasons, ocean science has traditionally been limited by researchers' ability to acquire measurements. However, modern technology is increasingly enabling scientists to transcend these historic impediments. With these novel sensing capabilities has come a new problem: it is difficult for under-resourced academic research groups to process and analyze measurements at the rate modern sensors can collect them.

Optical instruments are a cornerstone of expanding sensing capabilities. Underwater, cameras enable researchers to capture expansive environments at millimeter resolution, where they were formerly restricted (by oxygen tank capacity and human physiology) to just a few minutes with handheld tools (Weinberg, 1981; Gracias and Santos-Victor, 2000). Commercial cameras with relatively simple three-channel RGB sensors are generally used for this application, though some researchers are experimenting with more advanced multi- and hyperspectral instruments. Similar optical sensors mounted on unmanned aerial vehicles (UAVs), airplanes, or satellites can cover vastly more area than manned or unmanned underwater platforms, but they are much further away, and so suffer a loss in resolution. They are also limited to imaging the shallow surface layers of the ocean, and must account for its complex reflective surface, as well as the atmosphere

between the sensor and the ocean. Nevertheless, optical oceanographers have had great success, both with in-water and in-lab instrumentation, as well as building algorithms to interpret optical measurements made from satellites.

Progress in processing and analytical techniques for these new sensing capabilities is quite varied, so I cannot summarize it in its entirety in this venue. There is constant incremental progress in measurement interpretation techniques; some utilize novel machine learning and neural network capabilities, while others rely on e.g. mechanistic or statistical models. Ever-increasing computational processing speeds contribute to this ongoing development. However, many processing and analytical limitations remain, resulting in underutilized sensing technology.

SATELLITE-BASED OPTICAL SENSORS

Satellite-mounted optical instruments are successful and widely-used tools in the ocean sciences. Depending on their orbit, they are able to take measurements of all the world's oceans on a time scale of days to weeks. A single satellite image with a pixel size of one kilometer square represents thousands of independent measurements in a precise gridded pattern over a large area – replicating those measurements from a ship would take weeks, and they would not be contemporaneous.

Optical satellite oceanography attempts to solve a simple problem: given a satellite sensor measurement of light intensity at some number of wavelengths, determine as many properties of the water and its contents as possible (e.g. temperature, salinity, chlorophyll). Since the launch of the Coastal Zone Color Scanner (CZCS) in 1978, a variety of optical oceanographic instruments, ranging in target applications, country of origin, age, and more, have been launched into orbit.

Some algorithms for estimating oceanographic properties from satellite optical signals are well-established, while many others are in development, aided by the deployment of newer, more advanced sensors (e.g. the hyperspectral NASA Plankton, Aerosol, cloud, ocean Ecosystem (PACE) mission) (Gorman et al., 2019). In order to create algorithms for interpreting satellite optical signals, researchers must first A) optically characterize a volume of seawater, and then B) evaluate its characteristics and suspended particulate contents. The measurements of seawater properties/contents and associated optical properties can then be used to investigate causal links between properties of the seawater and resulting optical characteristics. When algorithms are derived to deduce water properties from optical signals, the satellite-measured optical signals can be used to infer seawater characteristics.

SCIENTIFIC VALUE OF PARTICLE SIZE DISTRIBUTION MEASUREMENTS

One desirable target for optical satellite algorithms is characterizing the particle size distribution (PSD) in the ocean. Knowledge of the PSD contributes to scientific understanding of numerous physical, chemical, and biological processes. Examples of such size-dependent phenomena include:

1) Rates of particle aggregation and sinking (Jackson, 1995; Stemmann et al., 2004; Burd, 2013)
2) Particle colonization and remineralization rates (Kiørboe, 2000; Ploug and Grossart, 2000; Kiørboe et al., 2004), and
3) Planktonic metabolic processes and trophic interactions (Chisholm, 1992; Hansen et al., 1997; Gillooly et al., 2001; Jennings and Warr, 2003; Brown et al., 2004; Woodward et al., 2005)

Particle size also plays a significant role in determining the light scattering and absorption properties of seawater, and thus the extent of penetration of different wavelengths of light within the ocean (Morel and Bricaud, 1986; Stramski and Kiefer, 1991; Stramski et al., 2001; Stemmann and Boss, 2012; Agagliate et al., 2018). The PSD directly influences the behavior of light in the ocean, suggesting that it may be possible to infer the PSD from optical measurements.

CHARACTERIZATION AND MEASUREMENT OF PSD

The size distribution of marine particle assemblages varies continually in time and space, as many overlapping processes add, remove, or alter particles in a given volume of seawater. Observations suggest the PSD often exhibits a complex shape (Sheldon, 1972; Jonasz and Fournier, 2007; Reynolds et al., 2010), so simplifying descriptors of the PSD struggle to capture the full complexity of these dynamic systems. However, it has long been recognized that different particle size ranges often have different physiological, ecological, or biogeochemical roles (Quéré et al., 2005; Stemmann and Boss, 2012; Ward et al., 2012). The PSD is thus often parametrized by size classes approximating different planktonic "functional types" based on size (IOCCG, 2014; Mouw et al., 2017). The most common size grouping of plankton used in oceanographic studies is based on particle diameter (Sieburth et al., 1978); picoplankton (diameter range 0.2–2 μm), nanoplankton (2–20 μm), and microplankton (20–200 μm).

In addition to size ranges, single metrics such as the average or median particle diameter (by particle number, area, or volume concentration) are another means to simplify the complete PSD function (Bernard et al., 2007; Woźniak et al., 2010; Briggs et al., 2013; Slade and Boss, 2015). Alternatively, parametrizations which model the shape of the PSD in seawater have been proposed, such as power law models (Bader, 1970), Gaussian or log-normal distributions (Jonasz,

1983, 1987), and the gamma function (Risović, 1993). Of these, the most commonly used is the power law model (often referred to as the Junge distribution); research has offered justification for its applicability to marine biological and biogeochemical systems (Platt and Denman, 1978; Kiefer and Berwald, 1992; Buonassissi and Dierssen, 2010).

Recent advances in terrestrial optically-based instrument capabilities, such as improved imaging and evaluation of light scattering, have increased capabilities to measure the PSD (Jackson et al., 1997; Agrawal and Pottsmith, 2000; Moore et al., 2009; Graham and Nimmo Smith, 2010; Picheral et al., 2010). However, due to the inherent limitations of ship, shore, stationary buoy, or drifter/glider sensor platforms, they alone likely cannot provide the spatial or temporal resolution necessary to fully characterize complex global oceanic ecosystems. For this reason, efforts to develop space-based remote-sensing approaches for estimation of the PSD from imaging sensors (ocean color) are a subject of ongoing research interest (Bowers et al., 2007; Kostadinov et al., 2009; Shi and Wang, 2019).

Chapter 1 of this dissertation addresses the first step in optical algorithm development: parametrizations of the PSD are investigated, with the goal of evaluating potential targets of optical signal interpretation models. The PSD is a complex and continuously varying function, so it is unlikely that its entirety can be determined from optical measurements; however, if a relatively simple parametrization or non-parametric model can be developed that captures much of the full PSD, it may be possible to determine causal relationships between the variables of such a parametrization and optical measurements that can be made from satellites.

DIGITAL IMAGING UNDERWATER FOR MAPPING CORAL REEFS

While satellite-based optical instrumentation can inform many investigations of scientific value, *in situ* measurements are more suitable for other research goals. Diver-based digital imagery and derived products (e.g., 3D pointclouds, 2D orthographic projections) enable coral researchers to capture high-resolution maps of the benthos, allowing investigation of e.g. the growth or death of coral colonies and their spatial interrelationships through time. This is valuable for a variety of biological investigations.

BIOLOGICAL MOTIVATION

The description of coral colonial demography—rates of recruitment, survivorship, growth, bleaching, and tissue loss—requires data collection focused on individual coral organisms, as do other interesting research goals, such as evaluation of inter- and intra-genera/species competition. However, to date, most studies of coral communities focusing on site (Goreau, 1959; Kenyon et al., 2010), island (Newman et al., 2006; Sandin et al., 2008), or region (Smith et al., 2016), have relied on percent cover data. Percent cover is a metric often used in coral research that estimates the percentage of the benthos occupied by different taxonomic categories. To determine accurate context-dependent demographic rates, it is necessary to track many individual coral colonies though time and obtain accurate measurements of colony size and location. A handful of studies have accomplished this with laborious in-water monitoring protocols (Lewis, 1970; Stimson, 1974; Bradbury and Young, 1981; Bak et al., 1982; Hughes, 1984; Carlon and Olson, 1993; Bak and Nieuwland, 1995; Connell et al., 1997; Fong and Glynn, 1998; Hughes and Tanner, 2000; Jolles et al., 2002; Karlson et al., 2007; Zvuloni et al., 2009; Vardi et al., 2012; Deignan and Pawlik, 2015; Doropoulos et al., 2015; Edmunds, 2015; Dana, 1976), but their comprehensiveness was necessarily limited by the demanding methodology. The painstaking nature of reliable and

repeated measurements of coral sizes and fates has been a bottleneck for expanded community-level analyses; a complete understanding of population-level demography as well as community-level succession requires long time series of spatially-referenced data.

Analogous studies in terrestrial ecology have generally utilized large-area sampling of individual organisms in the context of long-term study sites, including Barro Colorado Island (Hubbell and Foster, 1992), the Hubbard Brook experimental forest, and others (Condit et al., 2000). Among many other applications, these data have been used to identify or evaluate fundamental ecological dynamics, such as dispersion patterns (Hubbell, 1979; Lieberman et al., 1985; Condit et al., 2000), structuring mechanisms (e.g., recruitment patterns, habitat preference and availability, dispersal probabilities, resource limitation; (Hubbell, 1979; Connell, 1985; Turner, 1989; Condit et al., 2000; Rietkerk and van de Koppel, 2008), and space use (Harms et al., 2000; Marhaver et al., 2013). Chapters 2, 3, and 4 of this dissertation aim to contribute to the development of survey/processing procedures for underwater ecosystems that support large-area ecological investigation in the context of coral reefs.

UNDERWATER IMAGE SURVEYS

Using structure from motion (SFM) photogrammetric algorithms, 3D models of entire reef tracts can be derived from survey imagery (Pizarro et al., 2009; Smith et al., 2016; Edwards et al., 2017; Ferrari et al., 2017; Kodera et al., 2020; Sandin et al., 2020). A variety of metrics can be extracted from these 3D models or 2D orthographic projections (orthoprojections) thereof, including growth, percent cover, species composition, or disease or bleaching incidence.

To provide SFM software with imagery sufficient to create high-resolution, geometrically-accurate 3D models, the field survey itself must capture many high-resolution images. SFM algorithms require extensive overlap (> 90%) between images—if insufficient overlap is captured, the software will have difficulty solving the relative 3D locations of features in the images and might create a 3D model that is inaccurate or incomplete, or fail to compute one altogether. Additionally, model quality/resolution is limited by detail captured in the original images. Divers properly using high-quality cameras while swimming slowly and close to the bottom will capture substantial biologically-interesting detail. Similar surveys done by divers using low-quality cameras with lower-resolution sensors, or improperly using cameras and their associated underwater housings, while swimming quicker and further from the reef will collect less detail.

My second chapter focuses on this field procedure—specifically, the lessons the Kuester and Sandin labs have learned from a decade of doing precise above- and below-water image surveys to construct 3D models of maximal scientific utility.

DATA ANALYSIS

Modern sensing platforms enable collection of vastly more imagery and associated data than was previously possible. However, these new sensing capabilities have introduced the much-discussed Big Data Problem (BDP): measurements are captured at a much higher rate than even well-funded research labs have the human and computational resources to analyze. Our research teams — the Sandin and Smith coral labs, 100 Island Challenge partners, and Falko Kuester's Cultural Heritage Engineering Initiative (CHEI) — have managed to amass a library of thousands of reef sites, of which the majority remain partially or entirely uninvestigated.

NNS FOR IMAGE/POINTCLOUD ANALYSIS

Machine learning tools such as neural networks (NNs) and other processing/analytical strategies offer hope that such backlogs can be cleared. The story of recent advances in neural network technology has been told widely and repeatedly, so I will keep my review brief.

First, a formal definition: "An Artificial Neural Network is an information or signal processing system composed of a large number of simple processing elements which are interconnected by direct links and which cooperate to perform parallel distributed processing in order to solve a desired computational task" (Macukow, 2016: 3). In layman's terms, NNs are computational tools that learn to interpret data, such as the aforementioned large-area 2D/3D image-derived products.

The fundamental unit of a NN is modeled after the fundamental unit of our own brains: a neuron. The function of a biological neuron is to receive multiple inputs from other neurons and then, if a threshold is reached, pass an electrical signal to downstream neurons (Vodrahalli and Bhowmik, 2017). That simple definition does not capture the full physiological complexity of biological neurons, but it is a useful conceptual model. Vodrahalli and Bhowmik explain the assumptions used to translate that conceptual model into computational neurons:

> There are a few key abstractions we can make to simplify the complex biology involved in a neuron. First, a single neuron receives inputs from many other neurons and transmits its action potential to many other neurons; there is a high connectivity between neurons. Second, the action potential is triggered at a threshold value; on a given input, only a portion of the neurons will activate indicating that the arrangement and relation of neurons is able to capture information about the structure of the input. Lastly, neurons are highly nonlinear; the transmitting of the signal from one neuron to the next is governed by the complex biochemistry at the synapses of the neurons (Vodrahalli and Bhowmik, 2017: 677).

Of note, the point concerning threshold values is (for the most part) not true in modern NNs: computational neurons can pass on continuously-varying values, instead of a simple binary

0 or 1. However, the now-commonly-used rectified linear activation function (ReL) still has a threshold at 0: any input signal below 0 is output as 0 (Hahnloser et al., 2000). The third point is critical: non-linear activation functions allow NNs to model realistic, complex non-linear systems.

Early simple neural networks based on this conceptual framework were introduced in the 1940s (McCulloch and Pitts, 1943). NN research developed over the proceeding decades, with peaks and valleys in funding and research effort due to cycles of over-optimism and the resulting backlash (Macukow, 2016). These NNs were much smaller than modern iterations, but were still limited by computational capabilities of the time, as well as algorithmic/mathematical design and implementation weaknesses that made optimization unstable and reduced their ability to map complex functions. A series of design developments, including backpropagation (Werbos, 1988), the aforementioned ReL activation function, normalization and regularization strategies, and optimization strategies, enabled better and more stable NNs. With immense modern computational speed, these developments enabled modern NNs composed of millions of parameters arranged in many layers.

While applications of NNs vary widely, from here on I will focus on their use in computer vision—the computational interpretation of imagery. In 2012, the NN AlexNet achieved a 15.3% error rate compared to 26.2% for the second placed submission on the image classification component of the annual ImageNet Large Scale Visual Recognition Challenge. Since that dramatic success, neural networks have generally been the highest-performing method for computer vision tasks such as classification, segmentation, and object detection (Krizhevsky et al., 2012a; Russakovsky et al., 2015).

It is prudent to pause here and define pointclouds (sometimes written point clouds). Pointclouds are 3D data products analogous to images, except that they are composed of points in

3D space, each with corresponding XYZ values. Both image pixels and pointcloud points generally have RGB color values, but any information can be stored at these spatial locations (e.g., range/depth, simple monochromatic tone, or anything else of interest). Unlike images, pointclouds are irregular—pixels occupy a defined spatial grid, but points can have any XYZ value. Pointclouds are also unordered—because there is no grid, it is not trivial to determine which points are nearest each other or the distance between them. These properties make pointclouds relatively geometrically complicated to work with, but also more flexible.

Chapter 3 of this dissertation concerns segmentation (sometimes referred to as semantic segmentation, but I will use the shorter version), which is defined as assigning a label to every pixel an image or point in a pointcloud. Tracing is segmentation—each organism is outlined and assigned a taxonomic label, which is also applied to pixels/points inside the outline. By this means, every pixel/point is assigned to a taxonomic category or a background or other category. At the end, every single pixel/point has a label.

Chapter 4 concerns classification. The typical definition of classification is assigning a single label to an entire scene/image/pointcloud. However, images and derived data products in this dissertation contain many organisms, so pixels/points were randomly sampled and the corresponding organism was taxonomically identified. I refer to this process as point sampling. Generally speaking, full segmentations contain more information than a subset of sampled points, but they are also much more laborious to produce.

NNs for computer vision function on the fundamental principle of an encoder: the objective is to translate complex 2D or 3D visual information contained in images/pointclouds/etc into a many-dimensional feature space, with scores in each dimension representing the presence or absence of certain patterns and combinations of patterns in the input (referred to as features). In

feature space, similar inputs are expected to produce similar feature space representations, allowing the input to be identified by its similarity to other known samples.

While some recent NNs are moving beyond convolutions (Vaswani et al., 2017), convolutions were a critical component during the development of computer vision NNs, and all NN architectures in this dissertation utilize them; I will use the abbreviation CNN to refer to convolutional neural networks. Convolutions are windowed pattern-detection functions: if the pattern in the input matches the pattern in the convolutional window (which generally looks at only a small patch of the total input at any one time), a high value will be returned for that location in the input space; otherwise a low value will be returned.

The fundamental encoding structure of CNNs is simple: a variety of convolutional filters (which were used in computer vision well before NNs) examine the input image in small patches, testing for the presence of simple features like colors, edges, or gradients. In the earliest NN strategies, these pattern-detecting filters were designed by hand, drawing from successful pre-NN digital image processing techniques (e.g., Sobel or Canny edge detection), while in more modern designs the filter patterns are determined automatically through training. Typically, the first CNN layer encodes the input image as a series of presence/absence/magnitude-of-presence scores for simple features, while later layers apply convolutional pattern detectors to the feature scores of the proceeding layer. In progressive layers, the windowed nature of convolutions typically causes the spatial dimensions to decrease relative to the input as encoded features become more complex. To better understand this complicated idea, consider an input image of 500 x 500 pixels, with non-overlapping convolution windows of size 5x5. A single score (for each feature) will be generated for the first 5x5 set of pixels, and again for the next, etc. Along a single 500 pixel spatial dimension, this results in feature scores for 100 patches. The spatial dimension for the output of this layer will

be 100 x 100, where each spatial grid cell contains a vector of pattern detection (feature) scores for a number of patterns. This 100 x 100 set of feature scores is passed to the next layer, which by the same process reduces it to 20 x 20. At this point, each of the 20 x 20 spatial locations in this layer is drawn from a 25 x 25 patch of the original input image.

Simple patterns are generally detected first, and then combined with nearby patterns to form higher-level, more complex encodings of the patterns present in the image. This process mimics a simplified understanding of how our own vision system works. It is important to note that individual convolutional filters are typically applied to all patches in the input—the network is not learning different filters for different spatial locations. Spatial understanding is developed by examining the presence or absence of the same patterns in different locations. A simplified description of the training process of a standard supervised CNN is as follows:

1) An image, RGBD depth map, fully 3D pointcloud, or other visual media is input.
2) The image is passed through the multi-scale encoder, generating either a single 1x1 set of feature scores representing the entire image (classification), or the encoder result is upsampled back to the spatial dimensions of the input with feature scores for each pixel (segmentation).
3) The softmax activation function (Bridle, 1989, 1990) or a similar function is used to generate probabilities for each possible class, either for each pixel or the entire image. This is typically done with a final NN layer that is “fully connected”, meaning every input feature score is visible to the final layer, not just those in a spatial window. This is slightly different in a segmentation network, where the final layer prediction is generally based only on feature scores at each pixel location.

4) A loss function—typically cross entropy (Shannon, 1948), but a wide variety exist—is used to generate an error value that represents the difference between the supplied ground truth (i.e., the correct class label as determined by a human expert) and the NN prediction.
5) A backpropagation algorithm is applied, which computes the gradient of the loss function with respect to individual weights connecting layers.
6) An optimization function, such as simple gradient descent or more complicated functions like Adam (Kingma and Ba, 2015), is used to update the network weights based on the loss score (technically, on the partial derivative of the cost function with respect to each network parameter). If functioning properly, this will improve the performance of the network when presented with similar inputs in the future.
7) The process is repeated with more training samples.
8) Generally, training is stopped when the loss function value no longer decreases with additional training samples.

NNs FOR CORAL

A number of studies have evaluated automated methods, both NN-based and otherwise, for classifying or segmenting 2D images of coral (Beijbom et al., 2012; King et al., 2018; Nguyen et al., 2021; Burns et al., 2022), and 2D orthoprojections of 3D reef models (Alonso et al., 2017, 2019; Alonso and Murillo, 2018; Yuval et al., 2021). The NN tool most widely used in the coral science community is CoralNet, hosted at coralnet.ucsd.edu (Beijbom et al., 2012). CoralNet is free to use and cloud-based, so researchers from around the world can upload their own images, define label sets, create sample points, use the built-in user interface to manually classify the points, and train NNs to provide recommendations or automatically classify points.

Use of segmentation NNs in the coral community is more limited, and the available options are more fractured. There are several available methodologies—determining the best option depends primarily on the type and amount of available input data. Creating a ground truth mask by fully manually segmenting a 2D image or orthoprojection or 3D pointcloud is extremely labor-intensive. Researchers that can marshal the resources to accomplish this task on hundreds or more samples can train and evaluate standard fully-supervised segmentation NNs (Runyan et al., 2022). Alternative approaches include:

1) Training segmentation algorithms directly from image-level classifications (Alonso et al., 2017). CoralNet also has segmentation functionality, but it is only available by API, not graphical user interface.
2) Using experts or an image patch classification NN like CoralNet to label a subsample of pixels in an image, and then calculate segmentation error only at those points (Alonso et al., 2019), or expand those points to partially or entirely fill the ground truth mask, either by simple geometric methods or superpixels (Alonso et al., 2019; Yuval et al., 2021).

Performance of these alternative methods to achieve automated segmentation predictions, which do not require experts to laboriously create full ground truth masks, sometimes approaches that of fully-supervised methods but does not exceed it. This means to minimize the effort required to produce training data would be quite promising if accuracy of fully-supervised NNs was considered adequate, but it is not.

Beyond coral applications, NNs for top-down photographic/remote sensing applications such as airplane/UAV/satellite imagery for e.g. scientific, planning, or agricultural applications have been evaluated in 2D, 2.5D (RGBD), 3D, and hybrid formats, with promising results

(Mohammadi et al., 2019; Bachhofner et al., 2020; Saralioglu and Gungor, 2020; Song and Choi, 2020). More generically, many studies have examined 2D, 3D, and hybrid methods of segmentation, with two of the most common applications being 1) scene analysis to inform automated decision-making for driverless cars, grasping arms, and other robotic applications, and 2) interpretation of diagnostic medical imagery.

COMPARISON OF 2D AND 3D IMAGERY PRODUCTS

Simply point sampling survey imagery is useful, but generally only provides area cover percentages—this approach does not identify or track individual organisms. Additionally, survey imagery is usually heavily and variably redundant, so the same organism will appear many times in different images, further complicating use of this kind of data. Human divers cannot maintain an exactly constant speed, orientation of camera, or height from the benthos, so the number of images capturing any given organism will not be exactly consistent. In Chapter 4 I discuss NN-automated point sampling of 2D images. Segmenting survey imagery instead of point sampling enables identification of individual organisms, but does not solve the statistical problem of inconsistently-overlapping survey imagery.

Some of these shortcomings can be overcome by using SFM software (e.g., Agisoft Metashape) to derive 3D models from the images. Unlike any individual image, a 3D model contains all the organisms at a survey site and their spatial relationships. The 3D model construction process requires calculating the location/orientation from which each image was captured, and the location on the 3D reef surface that corresponds to each image pixel. This information can be used to account for the overlapping nature of the images, or to transfer point

sample classification/segmentation annotations from 2D images to 3D models or vice versa (Hopkinson et al., 2020).

3D models themselves can also be directly point sampled or segmented, as well as 2D projections thereof. Annotations can be included in this projection process (Runyan et al., 2022), or the projection can be annotated after it is generated. Annotations can be projected from orthoprojections back to 3D models, but complications arise in areas of the 3D model that are at oblique angles to the projection angle or are occluded by other surfaces.

The work discussed in Chapter 3 is, to the best of my knowledge, the first published examination of 3D neural networks, trained with 3D pointclouds segmented by experts, that apply predictions directly to 3D pointclouds. Others have evaluated transferring predictions made on 2D images to pointclouds (King et al., 2019; Mizuno et al., 2020).

NN PERFORMANCE ON CORAL DATA

Studies in Chapters 3 and 4 show varying levels of performance, but they do not exceed ~80-95% accuracy (though performance metrics and difficulty of the automated task vary from case to case). However, Chapter 3 and investigations by others show neural networks approaching or matching the performance of human experts (Yuval et al., 2021; Runyan et al., 2022). This distinction is very important. Coral classification or segmentation is quite difficult: some species/genera look quite similar, taxonomic organization shifts over time, images/pointclouds are not always adequate for accurate taxonomic identification, and even experts make mistakes. Studies comparing classification or segmentation annotations generated by different experts show agreement rates of ~75-95%, depending on e.g. the specificity of taxonomic classifications, image/pointcloud quality, level of biological expertise or geographic familiarity, and more

(Beijbom et al., 2015; Yuval et al., 2021; Runyan et al., 2022). In the context of segmentation, these disagreements most commonly occur at organisms' boundaries (Yuval et al., 2021; Runyan et al., 2022). Locating these boundaries is difficult and often subjective due to occlusions, shadows, and insufficient imagery resolution, so even experts find it difficult to identify the exact pixel/point at which an organism begins or ends.

This uncertainty in the ground truth data introduces uncertainty into the training data, and also into the evaluation data used to generate performance metrics. A NN prediction that is 100% correct evaluated on an expert-produced dataset that is 90% correct will be reported as 90% correct due to the deficiencies in the ground truth. It is therefore difficult to determine exactly where deficiency in NN performance ends and human uncertainty/fallibility begins. Much larger labeled datasets, and many more investigations of expert disagreement, are required to map out this error v. uncertainty relationship.

In the current paradigm, coral researchers often consider NN segmentation predictions insufficient for investigating growth of individual organisms because they are not 100% accurate. However, this is based on the obviously and well-documented-to-be false assumption that expert-produced data in the past was 100% accurate, and expert-provided training and testing ground truth data for the NNs is 100% accurate. If we are to understand how NN performance compares to past and present expert performance, we must first exhaustively study the performance of the experts.

NN-driven point sample classification methods like CoralNet, have, however, gained increasing acceptance for use in area cover studies of broad categories like hard coral or for common genera. Evidence shows that CoralNet-derived percent cover metrics correlate well with those produced manually (Williams et al., 2019).

REFERENCES FOR THE INTRODUCTION

Agagliate, J., Lefering, I., and McKee, D. (2018). Forward modeling of inherent optical properties from flow cytometry estimates of particle size and refractive index. *Applied Optics* 57, 1777–1788. doi: 10.1364/AO.57.001777.

Agrawal, Y. C., and Pottsmith, H. C. (2000). Instruments for particle size and settling velocity observations in sediment transport. *Marine Geology* 168, 89–114. doi: 10.1016/S0025-3227(00)00044-X.

Alonso, I., Cambra, A., Munoz, A., Treibitz, T., and Murillo, A. C. (2017). Coral-Segmentation: Training Dense Labeling Models with Sparse Ground Truth. in *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)* (Venice, Italy: IEEE), 2874–2882. doi: 10.1109/ICCVW.2017.339.

Alonso, I., and Murillo, A. C. (2018). Semantic Segmentation from Sparse Labeling Using Multi-Level Superpixels. in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Madrid: IEEE), 5785–5792. doi: 10.1109/IROS.2018.8594185.

Alonso, I., Yuval, M., Eyal, G., Treibitz, T., and Murillo, A. C. (2019). CoralSeg: Learning coral segmentation from sparse annotations. *J Field Robotics* 36, 1456–1477. doi: 10.1002/rob.21915.

Bachhofner, S., Loghin, A.-M., Otepka, J., Pfeifer, N., Hornacek, M., Siposova, A., Schmidinger, N., Hornik, K., Schiller, N., Kahler, O., Hochreiter, R. (2020). Generalized Sparse Convolutional Neural Networks for Semantic Segmentation of Point Clouds Derived from Tri-Stereo Satellite Imagery. *Remote Sensing* 12, 1289. doi: 10.3390/rs12081289.

Bader, H. (1970). The hyperbolic distribution of particle sizes. *Journal of Geophysical Research* 75, 2822–2830. doi: 10.1029/JC075i015p02822.

Bak, R. P. M., and Nieuwland, G. (1995). Long-Term Change in Coral Communities Along Depth Gradients Over Leeward Reefs in the Netherlands Antilles. *Bulletin of Marine Science* 56, 609–619.

Bak, R. P. M., Termaat, R. M., and Dekker, R. (1982). Complexity of coral interactions: Influence of time, location of interaction and epifauna. *Marine Biology.* 69, 215–222. doi: 10.1007/BF00396901.

Beijbom, O., Edmunds, P. J., Kline, D. I., Mitchell, B. G., and Kriegman, D. (2012). Automated annotation of coral reef survey images. in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 1170–1177. doi: 10.1109/CVPR.2012.6247798.

Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P., Dunlap, V. M., Fan, T., Tan, C., Chan, S., Treibitz, T., Gamst, A., Mitchell, G., Kriegman, D. (2015). Towards Automated Annotation of Benthic Survey Images: Variability of Human Experts

and Operational Modes of Automation. *PLOS ONE* 10, e0130312. doi: 10.1371/journal.pone.0130312.

Bernard, S., Shillington, F. A., and Probyn, T. A. (2007). The use of equivalent size distributions of natural phytoplankton assemblages for optical modeling. *Opt. Express, OE* 15, 1995–2007. doi: 10.1364/OE.15.001995.

Bowers, D. G., Binding, C. E., and Ellis, K. M. (2007). Satellite remote sensing of the geographical distribution of suspended particle size in an energetic shelf sea. *Estuarine, Coastal and Shelf Science* 73, 457–466. doi: 10.1016/j.ecss.2007.02.005.

Bradbury, R., and Young, P. (1981). The Effects of a Major Forcing Function, Wave Energy, on a Coral Reef Ecosystem. *Marine Ecology-progress Series* 5, 229–241. doi: 10.3354/meps005229.

Bridle, J. (1989). Training Stochastic Model Recognition Algorithms as Networks can Lead to Maximum Mutual Information Estimation of Parameters. in *Advances in Neural Information Processing Systems* (Morgan-Kaufmann). Available at: https://papers.nips.cc/paper/1989/hash/0336dcbab05b9d5ad24f4333c7658a0e-Abstract.html [Accessed October 23, 2022].

Bridle, J. S. (1990). Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. in *Neurocomputing* NATO ASI Series., eds. F. F. Soulié and J. Hérault (Berlin, Heidelberg: Springer), 227–236. doi: 10.1007/978-3-642-76153-9_28.

Briggs, N. T., Slade, W. H., Boss, E., and Perry, M. J. (2013). Method for estimating mean particle size from high-frequency fluctuations in beam attenuation or scattering measurements. *Applied Optics* 52, 6710–6725. doi: 10.1364/AO.52.006710.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a Metabolic Theory of Ecology. *Ecology* 85, 1771–1789. doi: 10.1890/03-9000.

Buonassissi, C. J., and Dierssen, H. M. (2010). A regional comparison of particle size distributions and the power law approximation in oceanic and estuarine surface waters. *Journal of Geophysical Research. Oceans; Washington* 115. doi: http://dx.doi.org/10.1029/2010JC006256.

Burd, A. B. (2013). Modeling particle aggregation using size class and size spectrum approaches. *Journal of Geophysical Research: Oceans* 118, 3431–3443. doi: 10.1002/jgrc.20255.

Burns, C., Bollard, B., and Narayanan, A. (2022). Machine-Learning for Mapping and Monitoring Shallow Coral Reef Habitats. *Remote Sensing* 14, 2666. doi: 10.3390/rs14112666.

Carlon, D. B., and Olson, R. R. (1993). Larval dispersal distance as an explanation for adult spatial pattern in two Caribbean reef corals. *Journal of Experimental Marine Biology and Ecology* 173, 247–263. doi: 10.1016/0022-0981(93)90056-T.

Chisholm, S. W. (1992). "Phytoplankton Size," in *Primary Productivity and Biogeochemical Cycles in the Sea* Environmental Science Research. (Springer, Boston, MA), 213–237. doi: 10.1007/978-1-4899-0762-2_12.

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbel, S. P., Foster, R.B., Itoh, A., Lafrankie, J. V., Lee, H. S., Losos, E., Manokaran, N., Sukumar, R., Yamakura, T. (2000). Spatial Patterns in the Distribution of Tropical Tree Species. *Science* 288, 1414–1418. doi: 10.1126/science.288.5470.1414.

Connell, J. H. (1985). The consequences of variation in initial settlement vs. post-settlement mortality in rocky intertidal communities. *Journal of Experimental Marine Biology and Ecology* 93, 11–45. doi: 10.1016/0022-0981(85)90146-7.

Connell, J. H., Hughes, T. P., and Wallace, C. C. (1997). A 30-Year Study of Coral Abundance, Recruitment, and Disturbance at Several Scales in Space and Time. *Ecological Monographs* 67, 461–488. doi: 10.1890/0012-9615(1997)067[0461:AYSOCA]2.0.CO;2.

Dana, T. F. Reef-Coral Dispersion Patterns and Environmental Variables on a Caribbean Coral Reef (1976). *Bulletin of Marine Science* 26(1), 1-13.

Deignan, L. K., and Pawlik, J. R. (2015). Perilous proximity: Does the Janzen–Connell hypothesis explain the distribution of giant barrel sponges on a Florida coral reef? *Coral Reefs* 34, 561–567. doi: 10.1007/s00338-014-1255-x.

Doropoulos, C., Ward, S., Roff, G., González-Rivero, M., and Mumby, P. J. (2015). Linking Demographic Processes of Juvenile Corals to Benthic Recovery Trajectories in Two Common Reef Habitats. *PLOS ONE* 10, e0128535. doi: 10.1371/journal.pone.0128535.

Edmunds, P. J. (2015). A quarter-century demographic analysis of the Caribbean coral, Orbicella annularis, and projections of population size over the next century. *Limnology and Oceanography* 60, 840–855. doi: 10.1002/lno.10075.

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, 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., Atwood, T. B., Byrne, M. (2017). 3D photogrammetry quantifies growth and external erosion of individual coral colonies and skeletons. *Sci Rep* 7, 1–9. doi: 10.1038/s41598-017-16408-z.

Fong, P., and Glynn, P. W. (1998). A dynamic size-structured population model: does disturbance control size structure of a population of the massive coral Gardineroseris planulata in the Eastern Pacific? *Marine Biology* 130, 663–674. doi: 10.1007/s002270050288.

Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M., and Charnov, E. L. (2001). Effects of Size and Temperature on Metabolic Rate. *Science* 293, 2248–2251. doi: 10.1126/science.1061967.

Goreau, T. F. (1959). The Ecology of Jamaican Coral Reefs I. Species Composition and Zonation. *Ecology* 40, 67–90. doi: 10.2307/1929924.

Gorman, E. T., Kubalak, D. A., Patel, D., Dress, A., Mott, D. B., Meister, G., Werdell, J. P. (2019). The NASA Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission: an emerging era of global, hyperspectral Earth system remote sensing. *Sensors, Systems, and Next-Generation Satellites XXIII* (SPIE), 78–84. doi: 10.1117/12.2537146.

Gracias, N. R., and Santos-Victor, J. (2000). Underwater Video Mosaics as Visual Navigation Maps. *Computer Vision and Image Understanding* 79, 66–91. doi: 10.1006/cviu.2000.0848.

Graham, G. W., and Nimmo Smith, W. A. M. (2010). The application of holography to the analysis of size and settling velocity of suspended cohesive sediments. *Limnology and Oceanography: Methods* 8, 1–15. doi: 10.4319/lom.2010.8.1.

Hahnloser, R. H. R., Sarpeshkar, R., Mahowald, M. A., Douglas, R. J., and Seung, H. S. (2000). Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405, 947–951. doi: 10.1038/35016072.

Hansen, P. J., Bjørnsen, P. K., and Hansen, B. W. (1997). Zooplankton grazing and growth: Scaling within the 2-2,-µm body size range. *Limnology and Oceanography* 42, 687–704. doi: 10.4319/lo.1997.42.4.0687.

Harms, K. E., Wright, S. J., Calderón, O., Hernández, A., and Herre, E. A. (2000). Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* 404, 493–495. doi: 10.1038/35006630.

Hopkinson, B. M., King, A. C., Owen, D. P., Johnson-Roberson, M., Long, M. H., and Bhandarkar, S. M. (2020). Automated classification of three-dimensional reconstructions of coral reefs using convolutional neural networks. *PLOS ONE* 15, e0230671. doi: 10.1371/journal.pone.0230671.

Hubbell, S. P. (1979). Tree Dispersion, Abundance, and Diversity in a Tropical Dry Forest. *Science* 203, 1299–1309. doi: 10.1126/science.203.4387.1299.

Hubbell, S. P., and Foster, R. B. (1992). Short-Term Dynamics of a Neotropical Forest: Why Ecological Research Matters to Tropical Conservation and Management. *Oikos* 63, 48–61. doi: 10.2307/3545515.

Hughes, T. P. (1984). Population Dynamics Based on Individual Size Rather than Age: A General Model with a Reef Coral Example. *The American Naturalist* 123, 778–795. doi: 10.1086/284239.

Hughes, T. P., and Tanner, J. E. (2000). Recruitment Failure, Life Histories, and Long-Term Decline of Caribbean Corals. *Ecology* 81, 2250–2263. doi: 10.2307/177112.

IOCCG (2014). Phytoplankton Functional Types from Space. International Ocean Colour Coordinating Group (IOCCG). doi: 10.25607/OBP-106.

Jackson, G. A. (1995). Comparing observed changes in particle size spectra with those predicted using coagulation theory. *Deep Sea Research Part II: Topical Studies in Oceanography* 42, 159–184. doi: 10.1016/0967-0645(95)00010-N.

Jackson, G. A., Maffione, R., Costello, D. K., Alldredge, A. L., Logan, B. E., and Dam, H. G. (1997). Particle size spectra between 1 μm and 1 cm at Monterey Bay determined using multiple instruments. *Deep Sea Research Part I: Oceanographic Research Papers* 44, 1739–1767. doi: 10.1016/S0967-0637(97)00029-0.

Jennings, S., and Warr, K. J. (2003). Smaller predator-prey body size ratios in longer food chains. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, 1413–1417. doi: 10.1098/rspb.2003.2392.

Jolles, A. E., Sullivan, P., Alker, A. P., and Harvell, C. D. (2002). Disease Transmission of Aspergillosis in Sea Fans: Inferring Process from Spatial Pattern. *Ecology* 83, 2373–2378. doi: 10.1890/0012-9658(2002)083[2373:DTOAIS]2.0.CO;2.

Jonasz, M. (1983). Particle-size distributions in the Baltic. *Tellus B: Chemical and Physical Meteorology* 35, 346–358. doi: 10.3402/tellusb.v35i5.14624.

Jonasz, M. (1987). Nonspherical sediment particles: Comparison of size and volume distributions obtained with an optical and a resistive particle counter. *Marine Geology* 78, 137–142. doi: 10.1016/0025-3227(87)90072-7.

Jonasz, M., and Fournier, G. (2007). *Light Scattering by Particles in Water: Theoretical and Experimental Foundations - 1st Edition*. Available at: https://www.elsevier.com/books/light-scattering-by-particles-in-water-theoretical-and-experimental-foundations/jonasz/978-0-12-388751-1 [Accessed June 17, 2018].

Karlson, R., Cornell, H., and Hughes, T. (2007). Aggregation influences coral species richness at multiple spatial scales. *Ecology* 88, 170–7. doi: 10.1890/0012-9658(2007)88[170:AICSRA]2.0.CO;2.

Kenyon, J. C., Maragos, J. E., and Cooper, S. (2010). Characterization of Coral Communities at Rose Atoll, American Samoa. Available at: http://repository.si.edu//handle/10088/11728 [Accessed November 12, 2019].

Kiefer, D. A., and Berwald, J. (1992). A random encounter model for the microbial planktonic community. *Limnology and Oceanography* 37, 457–467. doi: 10.4319/lo.1992.37.3.0457.

King, A., Bhandarkar, S. M., and Hopkinson, B. M. (2018). A Comparison of Deep Learning Methods for Semantic Segmentation of Coral Reef Survey Images. 2018 IEEE/CVF

Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018, pp. 1475-14758, doi: 10.1109/CVPRW.2018.00188.

King, A., M. Bhandarkar, S., and Hopkinson, B. M. (2019). Deep Learning for Semantic Segmentation of Coral Reef Images Using Multi-View Information. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2019, pp. 1-10

Kingma, D. P., and Ba, J. (2015). Adam: A Method for Stochastic Optimization. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, eds. Y. Bengio and Y. LeCun Available at: http://arxiv.org/abs/1412.6980 [Accessed April 5, 2022].

Kiørboe, T. (2000). Colonization of marine snow aggregates by invertebrate zooplankton: Abundance, scaling, and possible role. *Limnology and Oceanography* 45, 479–484. doi: 10.4319/lo.2000.45.2.0479.

Kiørboe, T., Grossart, H.-P., Ploug, H., Tang, K., and Auer, B. (2004). Particle-associated flagellates: Swimming patterns, colonization rates, and grazing on attached bacteria. *Aquatic Microbial Ecology* 35, 141–152. doi: 10.3354/ame035141.

Kodera, S. M., Edwards, C. B., Petrovic, V., Pedersen, N. E., Eynaud, Y., and Sandin, S. A. (2020). Quantifying life history demographics of the scleractinian coral genus Pocillopora at Palmyra Atoll. *Coral Reefs* 39, 1091–1105. doi: 10.1007/s00338-020-01940-8.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research: Oceans* 114. doi: 10.1029/2009JC005303.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, eds. F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc.), 1097–1105. Available at: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

Lewis, J. B. (1970). Spatial Distribution and Pattern of Some Atlantic Reef Corals. *Nature* 227, 1158–1159. doi: 10.1038/2271158a0.

Lieberman, D., Lieberman, M., Peralta, R., and Hartshorn, G. S. (1985). Mortality Patterns and Stand Turnover Rates in a Wet Tropical Forest in Costa Rica. *Journal of Ecology* 73, 915–924. doi: 10.2307/2260157.

Macukow, B. (2016). Neural Networks – State of Art, Brief History, Basic Models and Architecture. in *Computer Information Systems and Industrial Management*, eds. K. Saeed and W. Homenda (Cham: Springer International Publishing), 3–14. doi: 10.1007/978-3-319-45378-1_1.

Marhaver, K. L., Vermeij, M. J. A., Rohwer, F., and Sandin, S. A. (2013). Janzen-Connell effects in a broadcast-spawning Caribbean coral: distance-dependent survival of larvae and settlers. *Ecology* 94, 146–160. doi: 10.1890/12-0985.1.

McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5, 115–133. doi: 10.1007/BF02478259.

Mizuno, K., Terayama, K., Hagino, S., Tabeta, S., Sakamoto, S., Ogawa, T., Sugimoto, K., Fukami, H. (2020). An efficient coral survey method based on a large-scale 3-D structure model obtained by Speedy Sea Scanner and U-Net segmentation. *Sci Rep* 10, 12416. doi: 10.1038/s41598-020-69400-5.

Mohammadi, H., Samadzadegan, F., and Reinartz, P. (2019). 2D/3D information fusion for building extraction from high-resolution satellite stereo images using kernel graph cuts. *International Journal of Remote Sensing* 40, 5835–5860. doi: 10.1080/01431161.2019.1584417.

Moore, C., Barnard, A., Fietzek, P., Lewis, M. R., Sosik, H. M., White, S., Zielinski, O. (2009). Optical tools for ocean monitoring and research. *Ocean Sci.*, 24.

Morel, A., and Bricaud, A. (1986). Inherent optical properties of algal cells including picoplankton: Theoretical and experimental results. *Can. Bull. Fish. Aquat. Sci.* 214, 521–559.

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. *Frontiers in Marine Science* 4. Available at: https://www.frontiersin.org/articles/10.3389/fmars.2017.00041 [Accessed October 23, 2022].

Newman, M. J. H., Paredes, G. A., Sala, E., and Jackson, J. B. C. (2006). Structure of Caribbean coral reef communities across a large gradient of fish biomass. *Ecology Letters* 9, 1216–1227. doi: 10.1111/j.1461-0248.2006.00976.x.

Nguyen, T., Liquet, B., Mengersen, K., and Sous, D. (2021). Mapping of Coral Reefs with Multispectral Satellites: A Review of Recent Papers. *Remote Sensing* 13, 4470. doi: 10.3390/rs13214470.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnology and Oceanography: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462.

Pizarro, O., Eustice, R. M., and Singh, H. (2009). Large Area 3-D Reconstructions From Underwater Optical Surveys. *IEEE Journal of Oceanic Engineering* 34, 150–169. doi: 10.1109/JOE.2009.2016071.

Platt, T., and Denman, K. L. (1978). *The structure of pelagic marine ecosystems*. Marine Ecology Laboratory, Bedford Institute of Oceanography.

Ploug, H., and Grossart, H.-P. (2000). Bacterial growth and grazing on diatom aggregates: Respiratory carbon turnover as a function of aggregate size and sinking velocity. *Limnology and Oceanography* 45, 1467–1475. doi: 10.4319/lo.2000.45.7.1467.

Quéré, C. L., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Da Cunha, L. C., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., Wolf-Gladrow, D. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x.

Reynolds, R. A., Stramski, D., Wright, V. M., and Woźniak, S. B. (2010). Measurements and characterization of particle size distributions in coastal waters. *Journal of Geophysical Research: Oceans* 115. doi: 10.1029/2009JC005930.

Rietkerk, M., and van de Koppel, J. (2008). Regular pattern formation in real ecosystems. *Trends in Ecology & Evolution* 23, 169–175. doi: 10.1016/j.tree.2007.10.013.

Risović, D. (1993). Two-component model of sea particle size distribution. *Deep Sea Research Part I: Oceanographic Research Papers* 40, 1459–1473. doi: 10.1016/0967-0637(93)90123-K.

Runyan, H., Petrovic, V., Edwards, C., Pedersen, N., Alcantar, E., Kuester, F., Sandin, S. A. (2022). Automated 2D, 2.5D, and 3D Segmentation of Coral Reef Pointclouds and Orthoprojections. *Frontiers in Robotics and AI* 9. doi: 10.3389/frobt.2022.884317.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. doi: 10.48550/arXiv.1409.0575.

Sandin, S. A., Edwards, C. B., Pedersen, N. E., Petrovic, V., Pavoni, G., Alcantar, E., Chancellor, K. S., Fox, M. D., Stallings, B., Sullivan, C. J., Rotjan, R. D., Ponchio, F., Zglicynski, B. J. (2020). "Chapter Seven - Considering the rates of growth in two taxa of coral across Pacific islands," in *Advances in Marine Biology* Population Dynamics of the Reef Crisis., ed. B. M. Riegl (Academic Press), 167–191. doi: 10.1016/bs.amb.2020.08.006.

Sandin, S. A., Smith, J. E., DeMartini, E. E., Dinsdale, E. A., Donner, S. D., Friedlander, A. M., Konotchick, T., Malay, M., Maregos, J. E., Obura, D., Pantos, O., Pauley, G., Richie, M., Rohwer, F., Schroeder, R. E., Walsh, S., Jackson, J. B. C., Knowlton, N., Sala, E. (2008). Baselines and Degradation of Coral Reefs in the Northern Line Islands. *PLOS ONE* 3, e1548. doi: 10.1371/journal.pone.0001548.

Saralioglu, E., and Gungor, O. (2020). Semantic segmentation of land cover from high resolution multispectral satellite images by spectral-spatial convolutional neural network. *Geocarto International*, 1–21. doi: 10.1080/10106049.2020.1734871.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27, 379–423, 623–656.

Sheldon, R. W. (1972). Size Separation of Marine Seston by Membrane and Glass-Fiber Filters1. *Limnology and Oceanography* 17, 494–498. doi: 10.4319/lo.1972.17.3.0494.

Shi, W., and Wang, M. (2019). Characterization of Suspended Particle Size Distribution in Global Highly Turbid Waters From VIIRS Measurements. *Journal of Geophysical Research: Oceans* 124, 3796–3817. doi: 10.1029/2018JC014793.

Sieburth, J. M., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions 1. *Limnology and Oceanography* 23, 1256–1263. doi: 10.4319/lo.1978.23.6.1256.

Slade, W. H., and Boss, E. (2015). Spectral attenuation and backscattering as indicators of average particle size. *Applied Optics* 54, 7264. doi: 10.1364/AO.54.007264.

Smith, J. E., Brainard, R., Carter, A., Grillo, S., Edwards, C., Harris, J., Lewis, L., Obura, D., Rohwer, F., Sala, E., Vroom, P. S., Sandin, S. A. (2016). Re-evaluating the health of coral reef communities: baselines and evidence for human impacts across the central Pacific. *Proceedings of the Royal Society B: Biological Sciences* 283. doi: 10.1098/rspb.2015.1985.

Song, A., and Choi, J. (2020). Fully Convolutional Networks with Multiscale 3D Filters and Transfer Learning for Change Detection in High Spatial Resolution Satellite Images. *Remote Sensing* 12, 799. doi: 10.3390/rs12050799.

Stemmann, L., and Boss, E. (2012). Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. *Annual Review of Marine Science* 4, 263–290. doi: 10.1146/annurev-marine-120710-100853.

Stemmann, L., Jackson, G. A., and Ianson, D. (2004). A vertical model of particle size distributions and fluxes in the midwater column that includes biological and physical processes -- Part I model formulation. *Deep Sea Research Part I: Oceanographic Research Papers* 51, 865–884.

Stimson, J. (1974). An Analysis of the Pattern of Dispersion of the Hermatypic Coral Pocillopora Meandrina Var. Nobilis Verril. *Ecology* 55, 445–449. doi: 10.2307/1935234.

Stramski, D., Bricaud, A., and Morel, A. (2001). Modeling the inherent optical properties of the ocean based on the detailed composition of the planktonic community. *Appl. Opt., AO* 40, 2929–2945. doi: 10.1364/AO.40.002929.

Stramski, D., and Kiefer, D. A. (1991). Light scattering by microorganisms in the open ocean. *Progress in Oceanography* 28, 343–383. doi: 10.1016/0079-6611(91)90032-H.

Turner, M. G. (1989). Landscape Ecology: The Effect of Pattern on Process. *Annual Review of Ecology and Systematics* 20, 171–197. doi: 10.1146/annurev.es.20.110189.001131.

Vardi, T., Williams, D., and Sandin, S. (2012). Population dynamics of threatened elkhorn coral in the northern Florida Keys, USA. *Endangered Species Research* 19, 157–169. doi: 10.3354/esr00475.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017). Attention Is All You Need. doi: 10.48550/arXiv.1706.03762.

Vodrahalli, K., and Bhowmik, A. K. (2017). 3D computer vision based on machine learning with deep neural networks: A review. *Journal of the Society for Information Display* 25, 676–694. doi: 10.1002/jsid.617.

Ward, B. A., Dutkiewicz, S., Jahn, O., and Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnology and Oceanography* 57, 1877–1891. doi: 10.4319/lo.2012.57.6.1877.

Weinberg, S. (1981). A comparison of coral reef survey methods. *Bijdragen tot de Dierkunde* 51, 199–218.

Werbos, P. J. (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks* 1, 339–356. doi: 10.1016/0893-6080(88)90007-X.

Williams, I. D., Couch, C. S., Beijbom, O., Oliver, T. A., Vargas-Angel, B., Schumacher, B. D., Brainard, R. E. (2019). Leveraging Automated Image Analysis Tools to Transform Our Capacity to Assess Status and Trends of Coral Reefs. *Frontiers in Marine Science* 6. Available at: https://www.frontiersin.org/articles/10.3389/fmars.2019.00222 [Accessed November 14, 2022].

Woodward, G., Ebenman, B., Emmerson, M., Montoya, J. M., Olesen, J. M., Valido, A., Warren, P. H. (2005). Body size in ecological networks. *Trends in Ecology & Evolution* 20, 402–409. doi: 10.1016/j.tree.2005.04.005.

Woźniak, S. B., Stramski, D., Stramska, M., Reynolds, R. A., Wright, V. M., Miksic, E. Y., Cichocka, M., Cieplak, A. (2010). Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. *Journal of Geophysical Research: Oceans* 115. doi: 10.1029/2009JC005554.

Yuval, M., Alonso, I., Eyal, G., Tchernov, D., Loya, Y., Murillo, A. C., Treibitz, T. (2021). Repeatable Semantic Reef-Mapping through Photogrammetry and Label-Augmentation. *Remote Sensing* 13, 659. doi: 10.3390/rs13040659.

Zvuloni, A., Artzy-Randrup, Y., Stone, L., Kramarsky-Winter, E., Barkan, R., and Loya, Y. (2009). Spatio-Temporal Transmission Patterns of Black-Band Disease in a Coral Community. *PLOS ONE* 4, e4993. doi: 10.1371/journal.pone.0004993.

CHAPTER 1

EVALUATION OF PARTICLE SIZE DISTRIBUTION METRICS TO ESTIMATE THE RELATIVE CONTRIBUTIONS OF DIFFERENT SIZE FRACTIONS BASED ON MEASUREMENTS IN ARCTIC WATERS.

Hugh Runyan, Rick A. Reynolds, and Dariusz Stramski

This chapter is presented as a published paper as it appears in the journal *Journal of Geophysical Research: Oceans* in 2020 under the title "Evaluation of Particle Size Distribution Metrics to Estimate the Relative Contributions of Different Size Fractions Based on Measurements in Arctic Waters."

ABSTRACT

The size distribution of suspended particles influences several processes in aquatic ecosystems, including light propagation, trophic interactions, and biogeochemical cycling. The shape of the particle size distribution (PSD) is commonly modeled as a single-slope power law in oceanographic studies, which can be used to further estimate the relative contributions of different particle size classes to particle number, area, and volume concentration. We use a data set of 168 high size-resolution PSD measurements in Arctic oceanic waters to examine variability in the shape of the PSD over the particle diameter range 0.8 to 120 μm. An average value of -3.6 ± 0.33 was obtained for the slope of a power law fitted over this size range, consistent with other studies. Our analysis indicates, however, that this model has significant limitations in adequately parameterizing the complexity of the PSD, and thus performs poorly in predicting the relative contributions of different size intervals such as those based on picoplankton, nanoplankton, and microplankton size classes. Similarly, median particle size was also generally a poor indicator of these size class contributions. Our results suggest that alternative percentile diameters derived from the cumulative distribution functions of particle number, cross-sectional area, and volume concentration may provide better metrics to capture the overall shape of the PSD and to quantify the contributions of different particle size classes.

AIM

The objective of this study was to evaluate how empirical PSDs collected from *in situ* water samples in the arctic related to the power law function commonly used to parametrize PSDs, and then investigate alternative parametrizations.

METHODS

PSDs from multiple research cruises in the Arctic Ocean, obtained by instrument measurements on water samples collected *in situ* from multiple depths, were combined to create an Arctic PSD dataset. These data were investigated with power law and other parametrizations—specifically size classes and cumulative distribution functions.

RESULTS

Analyses indicated that the power law model has significant limitations in adequately parameterizing the complexity of the PSD, and thus, with our Arctic dataset, performs poorly in predicting the relative contributions of different particle size classes.

Further, results suggest that percentile diameters derived from the cumulative distribution functions of particle number, cross-sectional area, and volume concentration better capture the overall shape of the PSD, and more accurately quantify the contributions of different particle size classes.

JGR Oceans

RESEARCH ARTICLE
10.1029/2020JC016218

Key Points:
- Measurements of Arctic seawater are used to examine variability in the shape of the size distribution of suspended particles
- A power law model performs poorly in predicting the relative contributions of three size classes to the particle size distribution
- Specific percentiles of the cumulative number, area, and volume distributions are strongly correlated with size class contributions

Correspondence to:
R. A. Reynolds,
rreynolds@ucsd.edu

Citation:
Runyan, H., Reynolds, R. A., & Stramski, D. (2020). Evaluation of particle size distribution metrics to estimate the relative contributions of different size fractions based on measurements in Arctic waters. *Journal of Geophysical Research: Oceans*, *125*, e2020JC016218. https://doi.org/10.1029/2020JC016218

Received 6 MAR 2020
Accepted 22 APR 2020
Accepted article online 5 MAY 2020

Evaluation of Particle Size Distribution Metrics to Estimate the Relative Contributions of Different Size Fractions Based on Measurements in Arctic Waters

Hugh Runyan[1], Rick A. Reynolds[1], and Dariusz Stramski[1]

[1]Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA

Abstract The size distribution of suspended particles influences several processes in aquatic ecosystems, including light propagation, trophic interactions, and biogeochemical cycling. The shape of the particle size distribution (PSD) is commonly modeled as a single-slope power law in oceanographic studies, which can be used to further estimate the relative contributions of different particle size classes to particle number, area, and volume concentration. We use a data set of 168 high size-resolution PSD measurements in Arctic oceanic waters to examine variability in the shape of the PSD over the particle diameter range 0.8 to 120 μm. An average value of -3.6 ± 0.33 was obtained for the slope of a power law fitted over this size range, consistent with other studies. Our analysis indicates, however, that this model has significant limitations in adequately parameterizing the complexity of the PSD, and thus performs poorly in predicting the relative contributions of different size intervals such as those based on picoplankton, nanoplankton, and microplankton size classes. Similarly, median particle size was also generally a poor indicator of these size class contributions. Our results suggest that alternative percentile diameters derived from the cumulative distribution functions of particle number, cross-sectional area, and volume concentration may provide better metrics to capture the overall shape of the PSD and to quantify the contributions of different particle size classes.

Plain Language Summary The particle size distribution (PSD) describes how the concentration of particles changes with particle size, and it is an important characteristic of suspended oceanic particles that influences ocean ecology and biogeochemistry. We collected an extensive set of measurements of the PSD from Arctic waters to examine how different size classes of particles contribute to the total concentration of particle number, cross-sectional area, and volume. A model of the PSD frequently employed for oceanic studies is found to have strong limitations in representing these measurements of natural samples, and consequently performs poorly in estimating the relative contributions of individual size classes. We show that an alternative approach of describing the PSD based on specific percentile diameters derived from the cumulative distribution function of size-dependent particle concentration provides a better means to characterize the shape of the PSD for oceanic particle assemblages and provides superior performance in estimating size class contributions.

1. Introduction

Suspended particles in seawater play a key role in mediating numerous biogeochemical and ecological processes within the ocean. An important characteristic of these assemblages is the particle size distribution (PSD), which quantifies the concentration of particles as a function of particle size. Knowledge of the PSD and other particle characteristics such as composition or shape is needed for understanding numerous physical, chemical, and biological processes that involve particles. Examples of such size-dependent processes include rates of particle aggregation and sinking (Burd, 2013; Jackson, 1995; Stemmann et al., 2004), particle colonization and remineralization rates (Kiørboe, 2000; Kiørboe et al., 2004; Ploug & Grossart, 2000), and planktonic metabolic processes and trophic interactions (Brown et al., 2004; Chisholm, 1992; Gillooly et al., 2001; Hansen et al., 1997; Jennings & Warr, 2003; Woodward et al., 2005). The PSD also plays a critical role in determining the light scattering and absorption properties of seawater, and thus the penetration of light within the ocean (Agagliate et al., 2018; Baker & Lavelle, 1984; Morel & Bricaud, 1986; Stemmann & Boss, 2012; Stramski et al., 2001; Stramski & Kiefer, 1991).

Table 1
List of Notation

Symbol	Description, [typical units]
D	Equivalent spherical diameter, [μm]
$N(D)$, $A(D)$, $V(D)$	Particle size distribution based on number, cross-sectional area, or volume concentration, [m^{-3}, μm^2 m^{-3}, or μm^3 m^{-3}]
$N'(D)$, $A'(D)$, $V'(D)$	Density functions of $N(D)$, $A(D)$, or $V(D)$, [m^{-3} μm^{-1}, μm^2 m^{-3} μm^{-1}, μm^3 m^{-3} μm^{-1}]
N_t, A_t, V_t	Total concentration of particle number, area, or volume over size range D = 0.8 to 120 μm, [m^{-3}, μm^2 m^{-3}, μm^3 m^{-3}]
$\zeta_{(N, A, or V)}$	Slope of power law fit to $N'(D)$, $A'(D)$, or $V'(D)$, [unitless]
$CDF_{(N, A, or V)}(D)$	Cumulative distribution function of $N'(D)$, $A'(D)$, or $V'(D)$, [unitless]
$D_{(N, A, or V)}(X)$	Diameter corresponding to the X^{th} percentile value of $CDF_N(D)$, $CDF_A(D)$, or $CDF_V(D)$, [μm]
$f_{(N, A, or V), (pico, nano, or micro)}$	Fractional contribution to N_t, A_t, or V_t of the picoplankton, nanoplankton, or microplankton size class, [unitless]

Note. Symbols appearing with a hat above them, e.g., $\hat{f}$, indicate a model-derived value.

The PSD can be defined as the numerical concentration of particles per unit particle size interval, denoted here by $N'(D)$ and expressed in units of m^{-3} μm^{-1} (Jonasz & Fournier, 2007):

$$N(D) = N'(D)\,\Delta D, \tag{1}$$

where $N(D)$ is the number of particles per unit volume (m^{-3}) in the size interval $D \pm 0.5\Delta D$, and D is the midpoint diameter of each size class in μm (see also Table 1 for a list of notation used throughout the paper). Depending on the technique of size measurement, the particle diameter typically represents the volume-equivalent or area-equivalent spherical diameter. The number-based distribution, $N(D)$, can be transformed to other distributions based on the concentration of particle cross-sectional area, $A(D)$, or volume, $V(D)$, through assumptions of particle shape such as a sphere. To account for variations in the width of measured size intervals, the different forms of PSD or density functions, $N'(D)$, $A'(D)$, and $V'(D)$, are calculated by dividing the respective values of $N(D)$, $A(D)$, and $V(D)$ for each size bin by its width.

The size distribution of marine particle assemblages is continually varying in time and space as several competing processes add or remove particles from a given volume of seawater, or by conversions of particles from one type to another which are accompanied by changes in particle size. Characterizing and predicting variability in the PSD of oceanic waters is thus a major research challenge. Recent advances in particle imaging and other optically-based measurement technologies based on light scattering including diffraction have increased capabilities to measure the PSD in oceanic waters (Agrawal & Pottsmith, 2000; Graham & Nimmo-Smith, 2010; Jackson et al., 1997; Moore et al., 2009; Picheral et al., 2010), but in situ measurements alone cannot provide the spatial or temporal resolution needed to characterize global oceanic ecosystems. For this reason, efforts to develop remote-sensing approaches for estimation of the PSD from airborne or satellite measurements of ocean color have been pursued (Bowers et al., 2007; Kostadinov et al., 2009; Shi & Wang, 2019). These approaches generally rely on empirical parameterizations or simplified descriptions of the PSD in order to quantify relationships between the PSD and the optical properties of seawater.

Single metrics such as the mean or median particle diameter derived from the distribution of particle number, area, or volume concentration are one means to characterize the PSD (Bernard et al., 2007; Briggs et al., 2013; Slade & Boss, 2015; Woźniak et al., 2010). Alternatively, parametrizations that describe the shape of the PSD in seawater have also been proposed, such as power law models (Bader, 1970), Gaussian or log-normal distributions (Jonasz, 1983, 1987), and the gamma function (Risović, 1993). The power law model, often referred to as the Junge distribution, is the most commonly utilized and previous studies have offered justification for its applicability to marine assemblages (Kiefer & Berwald, 1992; Platt & Denman, 1978; Sheldon et al., 1972). The density function of particle number concentration for this model can be written as

$$N'(D) = N'_o\,(D/D_o)^{\zeta_N} \tag{2}$$

where D_o is a reference diameter (μm), N'_o (m^{-3} μm^{-1}) is the value at D_o, and ζ_N is the dimensionless

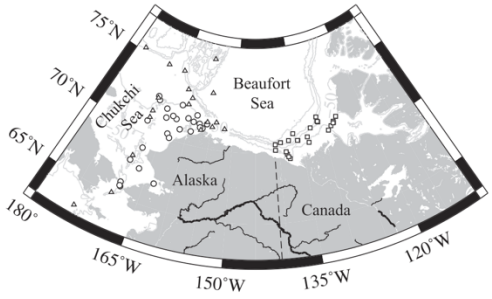


Figure 1. Map of study area depicting sampling locations for the MALINA (□), ICESCAPE (○), and Mirai (△) cruises.

slope of the distribution. Reported values for ζ_N in oceanic waters span a wide range and can vary for different size regions of the PSD (Jonasz & Fournier, 2007), but the majority of observations fall in the range of −4 to −3 (Buonassissi & Dierssen, 2010; Reynolds et al., 2010, 2016; Xi et al., 2014).

Although planktonic organisms exhibit a continuum of sizes, it has long been recognized that different particle size ranges tend to have different physiological capabilities as well as different ecological and biogeochemical roles (Le Quéré et al., 2005; Stemmann & Boss, 2012; Ward et al., 2012). The PSD is thus often further aggregated into broad size classes representing different planktonic "functional types" based on characteristic cell size (IOCCG, 2014; Mouw et al., 2017). The most common size grouping of plankton used in pelagic studies includes three planktonic size classes based on particle diameter (Sieburth et al., 1978); picoplankton (diameter range 0.2–2 μm), nanoplankton (2–20 μm), and microplankton (20–200 μm). Estimation of the relative contributions of these three size classes to the PSD is commonly used in ecological studies and its determination from remote sensing is also a subject of recent research efforts (IOCCG, 2014; Kostadinov et al., 2010).

Observations in various marine environments suggest that the PSD often exhibits a complex shape in response to the varying physical and biological processes that operate in aquatic ecosystems (Jonasz & Fournier, 2007; Reynolds et al., 2010; Sheldon et al., 1972), which reduces the ability of single descriptors of the size distribution to accurately quantify the role of different size ranges. The main objectives of this study are to evaluate variability in the relative contributions of different size classes to the particle number, area, and volume concentration and to explore the utility of relatively simple measures derived from the generally complex shapes of the measured PSD to estimate these contributions. For this purpose, we utilize 168 measurements of PSDs collected at 87 sites in the Arctic Ocean to characterize the shape of the PSD and the contribution of different particle size classes to overall particle number, area, and volume concentrations. These measurements, obtained with an electrical impedance approach (Coulter Counter), provide high size resolution measurements of the PSD for particle diameters spanning the range from 0.8 to 120 μm, enabling calculation of the relative contribution of particles in different size classes to total particle concentration. These size intervals were chosen to approximate the picoplankton, nanoplankton, and microplankton size range. We compare the measured particle number distributions with the single-slope power law parametrization and evaluate this model's capability to quantify the relative contributions of these three size classes. We then examine relationships between these size fraction contributions and the median and other percentile diameters derived from the cumulative distribution functions of particle number, area, and volume concentration.

2. Materials and Methods

2.1. Study Area and Sampling

Measurements were obtained on four expeditions to the western Arctic Ocean. The MALINA (MAckenzie LIght aNd cArbon) cruise occurred in the southeastern Beaufort Sea from 31 July to 24 August 2009 on the CCGS Amundsen. The station grid bracketed the outflows of the Mackenzie River, with transects extending from the delta to the southernmost limit of the pack ice outside the continental shelf. Two cruises associated with the NASA ICESCAPE (Impacts of Climate on EcoSystems and Chemistry of the Arctic Pacific Environment) program utilized the USCGC Healy to sample the Chukchi Sea and western Beaufort Sea during two successive years; from 18 June through 16 July 2010 and from 28 June through 24 July 2011. Sampling on these cruises included transects where measurements were done from open water across the ice edge to several kilometers within consolidated pack ice. A fourth cruise took place onboard the RV Mirai from 26 August to 18 September 2017 as part of the Japanese ArCS (Arctic Challenge for Sustainability) program, with sampling conducted from the Bering Strait to the southern limit of the ice edge at about 76.5°N. From these four cruises, 168 samples were collected for measurements of the PSD (Figure 1).

Water samples were obtained from two or three depths at each station using a CTD-Rosette equipped with Niskin bottles. The near surface layer (nominally 1–3 m depth) was always sampled, with additional depths corresponding to features such as maxima in chlorophyll *a* fluorescence or the optical beam attenuation coefficient, turbid layers within 3–5 m above the bottom on the shelf, or depths up to 300 m offshore. Water from an entire Niskin bottle or combination of bottles was withdrawn by opening the bottom closure and draining the contents into 20-L carboys to ensure collection of all particles and to minimize sampling error related to particle settling within the bottle. All measurements were made on board the research vessels and began within 1 hr of sampling.

2.2. Measurements of the PSD

The PSD was measured on seawater samples using a Coulter counter, an instrument that has been used for decades in the study of oceanic PSDs (Brun-Cottan, 1971; Carder et al., 1971; Jackson et al., 1997; Jonasz, 1983; Kitchen et al., 1975; Parsons, 1969; Sheldon et al., 1972). Water samples were withdrawn from the carboy after mixing, and the PSD was measured with a Beckman-Coulter Multisizer III using 0.2-μm filtered seawater as the diluent and blank. Samples were measured with a combination of two aperture sizes (30 and 200 μm), which provided the capability to count particles in the size range $D = 0.7$ to 120 μm where D represents volume equivalent spherical diameter. Both apertures were calibrated using suspensions of NIST-traceable microsphere standards of known size. For each seawater sample, multiple (>25) replicate measurements of the PSD were acquired for each aperture size and summed together to increase sampling volume and reduce the statistical error of particle counts. Total sample volumes after this summation averaged about 1.2 cm^3 for the 30-μm aperture and 180 cm^3 for the 200-μm aperture.

Each measurement with a given aperture provided the number of particles per unit volume, N (m^{-3}), within discrete size bins. Size classes consisted of 256 bins with logarithmically increasing bin width over the measured range of each aperture, ensuring high size resolution. The density function of the number concentration as a function of diameter, $N'(D)$ (m^{-3} μm^{-1}), was calculated by dividing the concentration of particles within each size bin by the bin width. To create the final distribution, measurements of $N'(D)$ from both apertures were merged at an overlapping size bin that shared a similar midpoint and bin width ($D = 4.8 \pm 0.03$ μm). The magnitude of $N'(D)$ as determined by the 30-μm aperture was adjusted to match the value measured by the 200-μm aperture at this overlapping bin, and the resulting merged distribution was converted back to $N(D)$ through multiplication of $N'(D)$ for each bin by the bin width. Because of high noise levels often observed with the smallest size bins of the 30-μm aperture, the distributions were truncated to a lower limit of $D = 0.8$ μm. The final merged distributions consist of 383 size bins spanning the range of D from 0.8 to 120 μm, with bin widths varying from about 0.01 to 1.6 μm.

2.3. Measurements of Particle Mass Concentration

Measurements of the particle dry mass concentration and organic carbon content were also obtained from the same samples to further characterize the bulk particle assemblage. The mass concentration of dried suspended particulate matter per unit volume of water, SPM (g m^{-3}), was measured using a standard gravimetric technique (van der Linde, 1998). Particles were collected on prerinsed, precombusted 25-mm glass-fiber filters (Whatman GF/F) that were weighed prior to use. Following filtration under low vacuum, sample filters and edges were rinsed with deionized water to remove residual sea salt, dried at 60°C, and stored sealed until analysis. The mass of particles collected on the filters was determined with a Mettler-Toledo MT5 microbalance with 1-μg precision. Two to three replicate filters were typically measured for each sample and averaged.

The concentration of particulate organic carbon, POC (mg m^{-3}), was obtained using a method consistent with established protocols (e.g., Knap et al., 1996). Water samples were filtered through precombusted 25-mm GF/F filters; filters were transferred to clean glass scintillation vials and dried at 60°C, then stored until post cruise analysis. Prior to analysis, filters were exposed to concentrated acid fumes (HCl) to remove inorganic carbon, and organic carbon concentration of each filter was determined with standard CHN analysis involving high temperature combustion of sample filters (Parsons et al., 1984). For MALINA, POC was measured from combustion of the same filters used in SPM determination. A number of unused filters from each lot of precombusted filters were used to quantify the background carbon content of filters and subtracted from the sample data. These blank filters were treated exactly like sample filters except that no

sample water was passed through them. Sample filtration volumes were large enough to ensure that contributions of the blank filter or adsorbed DOC were small relative to the particulate carbon content of the sample filter. Duplicate or triplicate samples were taken for each station and averaged to produce the final result of POC.

2.4. Analysis of Measured PSDs

The concentrations of particle cross-sectional area, $A(D)$ ($\mu m^2\ m^{-3}$), and particle volume, $V(D)$ ($\mu m^3\ m^{-3}$), within each size bin were calculated from the final merged $N(D)$ by assuming spherical particles and the relations $A(D) = N(D)\ \pi D^2/4$ and $V(D) = N(D)\ \pi D^3/6$. Density functions in terms of particle area concentration, $A'(D)$ ($\mu m^2\ m^{-3}\ \mu m^{-1}$), and particle volume concentration, $V'(D)$ ($\mu m^3\ m^{-3}\ \mu m^{-1}$), were computed in a manner similar to that of $N'(D)$ by dividing each value of $A(D)$ and $V(D)$ by the bin width.

To characterize the PSDs, cumulative distribution functions for particle number, $CDF_N(D)$, area, $CDF_A(D)$, and volume, $CDF_V(D)$, concentrations were calculated from the respective density functions, $N'(D)$, $A'(D)$, and $V'(D)$. For example, in the case of the particle number concentration, the cumulative distribution function $CDF_N(D)$ was computed according to

$$CDF_N(D) = \int_{0.8}^{D} N'(D)dD \Big/ \int_{0.8}^{120} N'(D)dD, \tag{3}$$

where the integration limits represent D in μm. The $CDF_A(D)$ and $CDF_V(D)$ functions were calculated in an analogous manner by replacing $N'(D)$ in equation 3 with $A'(D)$ and $V'(D)$, respectively. The particle diameters corresponding to specific percentiles of the cumulative distribution functions for particle number, area, and volume concentrations were then determined for each sample. The integral in the denominator of equation 3 is equivalent to the total particle number concentration over the measured size range of 0.8 to 120 μm, N_t (m^{-3}). Similarly, the total concentrations of particle area, A_t ($\mu m^2\ m^{-3}$), and volume, V_t ($\mu m^3\ m^{-3}$), can be calculated using the expression in the denominator of equation 3 and replacing $N'(D)$ with $A'(D)$ and $V'(D)$, respectively.

Other metrics characterizing the relative shapes of the PSDs were calculated by quantifying the fractional contributions of discrete size ranges to the values of N_t, A_t, and V_t. The size ranges of these classes were chosen to approximate the traditional plankton size classification scheme of Sieburth et al. (1978); picoplankton (f_{pico}; $0.8 \le D \le 2$ μm), nanoplankton (f_{nano}; $2 < D \le 20$ μm), and microplankton (f_{micro}; $20 < D \le 120$ μm). The fractional contributions of these size classes to the total number concentration were calculated according to

$$f_{N,pico} = \int_{0.8}^{2} N'(D)dD \Big/ \int_{0.8}^{120} N'(D)dD, \tag{4a}$$

$$f_{N,nano} = \int_{2}^{20} N'(D)dD \Big/ \int_{0.8}^{120} N'(D)dD, \tag{4b}$$

$$f_{N,micro} = \int_{20}^{120} N'(\mathrm{D})dD \Big/ \int_{0.8}^{120} N'(D)dD. \tag{4c}$$

The contributions of the three size classes to the particle area and volume distributions were calculated in an analogous manner by replacing $N'(D)$ in equation 4a-4c with $A'(D)$ and $V'(D)$, respectively. These fractional contributions are closely tied to the cumulative distribution functions described above. For example, equation 4a is equivalent to the cumulative distribution function of number concentration provided in equation 3 with an upper integration limit of $D = 2$ μm for the numerator. Similarly, the other size classes can be related to the CDF_N through the relations $f_{N,micro} = 1 - CDF_N(D = 20\ \mu m)$ and $f_{N,nano} = CDF_N(D = 20\ \mu m) - CDF_N(D = 2\ \mu m)$.

A final means to characterize the overall size distribution was determined by fitting a power law model to the measured data of the density function of particle number concentration (equation 2). This calculation was performed as a linear fit to $\log_{10}$-transformed data of $N'(D)$ and D over the entire measured size range. Prior to fitting, size bins for the 200-μm aperture that exhibited counts fewer than 20 particles (generally size bins corresponding to the largest diameters) were aggregated until that threshold was reached. We use the symbol $\hat{N}(D)$ to indicate the fitted values of the density function of particle number concentration. The

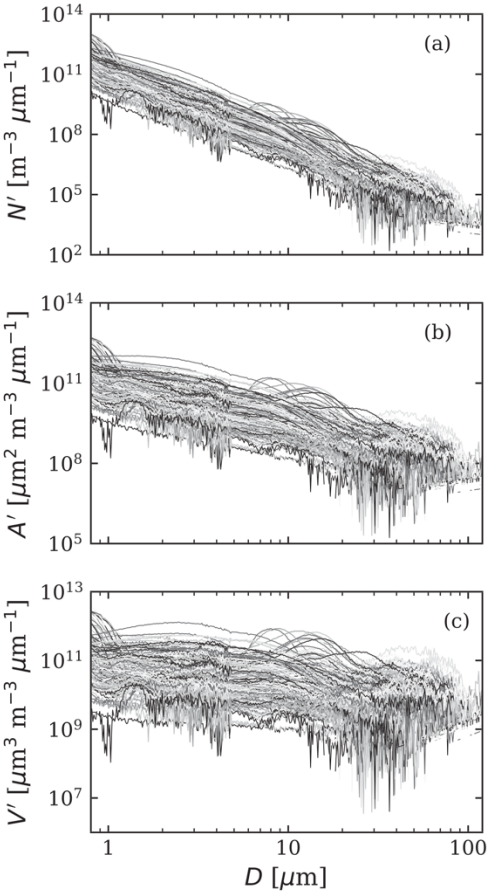


Figure 2. Measured density functions for the concentration of particle (a) number, N', (b) cross-sectional area, A', and (c) volume, V', as a function of equivalent spherical diameter D.

fitted power law functions of $\hat{A}'(D)$ and $\hat{V}'(D)$ were derived in a similar manner except that $A'(D)$ and $V'(D)$ replaced $N'(D)$, and ζ_A and ζ_V replaced ζ_N, in equation 2.

2.5. Statistical Analyses

We evaluated the capability of different metrics derived from the measured PSDs to estimate the fractional contributions of the picoplankton, nanoplankton, and microplankton size classes. Model I regression analysis was used to parameterize linear relationships, and nonlinear relationships were determined using the Python implementation of the Levenberg-Marquardt algorithm. For each relationship, the goodness-of-fit between the individual observations, O_i, and fitted model predictions, P_i, was characterized through the coefficient of determination R^2. To further assess model performance, scatterplots depicting model-derived versus measured values were subjected to Model II (reduced major axis) regression. We report values for the slope, intercept, and correlation coefficient R resulting from this analysis. Other statistical parameters of model performance include the median values of the ratio of model-derived to measured data, MdR, median values of the model bias, MdB = median (P_i – O_i), and the median absolute percent difference, $MdAPD$ = 100 × median|P_i – O_i/O_i|, between the model-derived and measured values. The root mean square deviation between model predictions and observations, $RMSD$, is also provided. This analysis was performed on the same data set that was used to parameterize the modeled relationships.

3. Results and Discussion

3.1. General Features of the PSD

Figure 2 illustrates the density functions for particle number, area, and volume concentration obtained for the entire data set (n = 168 observations). For some PSDs, portions of the density function exhibit rapid bin-to-bin variations that reflect poor counting statistics associated with low particle concentrations; this is generally most prevalent in the region D > 20 μm but also occasionally observed in the region D = 2 to 4.8 μm representing the range of the 30-μm aperture near the merge point with the 200-μm aperture. Particle number concentrations, $N'(D)$, in the smallest diameter size bins vary five orders of magnitude and range from 10^8 to more than 10^{13} m^{-3} μm^{-1}, while concentrations in the largest diameter size bins span the range 10^2 to more than 10^7 m^{-3} μm^{-1}. Particle area concentrations, $A'(D)$, in the smallest diameter size bins range from 10^8 to nearly 10^{13} μm^2 m^{-3} μm^{-1}, while concentrations in the largest size bins range from 10^{-5} to 10^9 μm^2 m^{-3} μm^{-1}. Particle volume concentrations, $V'(D)$, in the smallest diameter size bins range from 10^8 to more than 10^{11} μm^3 m^{-3} μm^{-1} while concentrations in the largest size bins range from fewer than 10^7 to nearly 10^{12} μm^3 m^{-3} μm^{-1}. The highest observed total concentrations generally occurred in coastal areas, plankton blooms, or near the bottom, while lower concentrations were generally associated with offshore areas or subsurface samples obtained below the chlorophyll a maximum but well above the bottom.

In all samples, the density functions of particle number, $N'(D)$, and area, $A'(D)$, concentration over the entire size range decreased in a generally logarithmic fashion with increasing diameter, and thus appear approximately linear when plotted with logarithmic scaling of both axes. This was also observed for nearly all density functions of particle volume, $V'(D)$, although three samples exhibited an increase in particle volume concentration with diameter. The calculated slopes ζ_N of $N'(D)$ range from −4.44 to −2.87 with an average value of −3.60 ± 0.33 (mean ± standard deviation) and a median value of −3.55 (Figure 3).

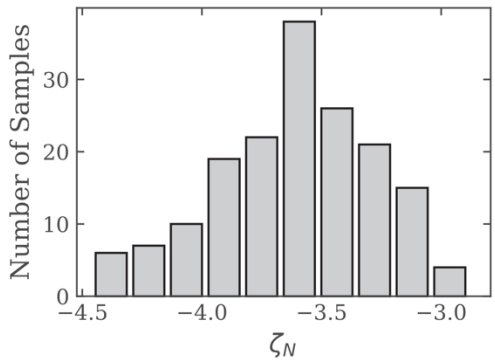


Figure 3. Histogram of the power law exponent ζ_N obtained by fitting a power law function to all measurements of $N'(D)$ (equation 2).

These average values and range of variability are consistent with previous studies of coastal and oceanic waters (Buonassissi & Dierssen, 2010; Jackson et al., 1997; Reynolds et al., 2010, 2016; Stemmann et al., 2008). Corresponding median values for the slopes ζ_A of $A'(D)$ and ζ_V of $V'(D)$ are −1.55 and −0.55, respectively. These values are in accordance with the assumption of spherical particles, for which $\zeta_A = \zeta_N - 2$ and $\zeta_V = \zeta_N - 3$, as discussed in Bader (1970).

Although the power law relationship may adequately characterize the overall relationship between particle abundance and size, significant departures from this model were observed in numerous samples from our data set. These consisted of discernible changes in the slope of the relationship for specific regions of the PSD, as well as obvious peaks in the distribution corresponding to specific particle, mostly likely planktonic, populations. Such features are illustrated in specific examples extracted from the data set.

Figure 4 depicts the measured density functions and corresponding cumulative distribution functions for samples representing the two extremes of particle mass concentration within our data set. The distributions shown in Figures 4a and 4b were obtained from a surface sample collected 9 km off the Alaskan coast near Utqiagvik (formerly known as Barrow). This extremely turbid sample (beam attenuation coefficient = 9.91 m^{-1} at 660 nm) exhibits the highest mass concentration of particles of any surface sample in the data set (SPM = 12.7 g m^{-3}). Despite the high SPM value, the particulate organic carbon content is low, and this sample exhibits the minimum POC/SPM value in our data set (0.024 g g^{-1}), suggesting that the sample is dominated by inorganic particles (Woźniak et al., 2010). The particle number size distribution for this sample is generally featureless with a steep overall slope across the distribution ($\zeta_N = -4.36$) indicating a proportionally high contribution of small particles to the PSD. However, deviations from the single slope obtained from the overall fit occur in portions of the size range

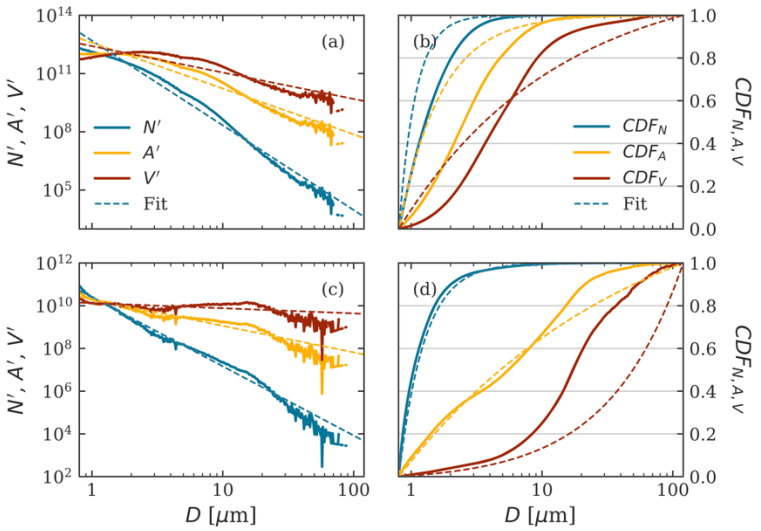


Figure 4. Example particle size distributions for samples representing the highest (a,b) and lowest (c,d) particle mass concentrations observed at near-surface depths. Panels (a,c) depict the measured distributions N', A', and V'; panels (b,d) depict the associated cumulative distribution functions, *CDF*. In all panels, dashed lines indicate the modeled distributions derived from a power law fit to the measured N'.

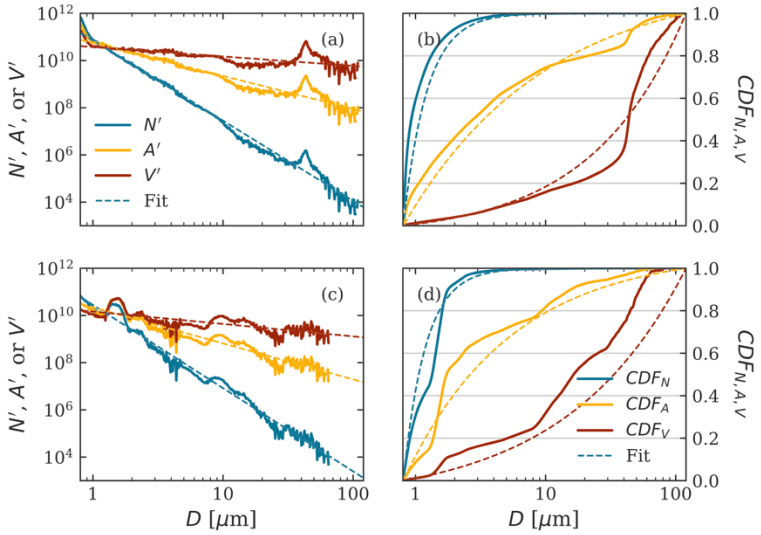


Figure 5. Similar to Figure 4, but for samples representing the subsurface chlorophyll *a* fluorescence maximum at two locations.

and these discrepancies are clearly discernable in a comparison of the cumulative distribution functions representing the measured $N'(D)$, $A'(D)$, and $V'(D)$ against those generated by the power law fit (Figure 4b). For example, the cumulative distribution function derived from the power law fit, $\widehat{CDF}_N$, is observed to overestimate the contribution of small particles resulting in an artificially steep cumulative distribution. This pattern becomes increasingly more pronounced in the power law predictions of the area and volume distribution and additionally illustrates that the power law model also overestimates the contribution of large particles.

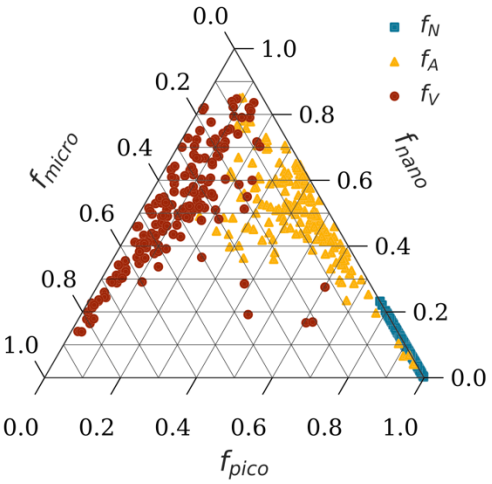


Figure 6. Ternary diagram illustrating the fractional contribution of three size classes (f_{pico}, f_{nano}, f_{micro}; equations 4a–4c) to the total concentration of particle number, N_t, cross-sectional area, A_t, and volume, V_t, as indicated by the legend.

In contrast, a sample obtained off the continental shelf in the western Beaufort Sea (Figures 4c and 4d) represents the lowest surface value of SPM obtained in our data set (0.044 g m^{-3}) and a relatively high POC/SPM ratio (0.47 g g^{-1}), suggesting that organic particles had a predominant contribution to particle mass concentration. This sample exhibits a flatter overall slope ($\zeta_N = -3.24$) than the example dominated by inorganic particles and has a discernable peak in the distribution centered near $D = 11$ μm (Figure 4c). This peak is also visible in Figure 4d as an increase in the slope of $CDF_A(D)$ and $CDF_V(D)$ in the same size range. The power law model provides a somewhat better description of the measured $N'(D)$ than in the sample illustrated in Figures 4a and 4b, as indicated by closer agreement between the $CDF_N(D)$ and the fitted line in Figure 4d. However, departures from the model are still evident upon examination of the $CDF_A(D)$ and $CDF_V(D)$. In the latter, the power law fit dramatically overestimates the contribution of large particles to volume concentration.

Figure 5 depicts the measured density functions and cumulative distribution functions for two samples collected from the subsurface

Table 2
Values of the Coefficient of Determination, R^2, for Relationships Between the Fractional Contributions of Three Size Classes to the Total Particle Number Concentration ($f_{N,pico}$, $f_{N,nano}$, $f_{N,micro}$), With the Corresponding Fractional Contributions to Total Particle Cross-Sectional Area ($f_{A,pico}$, $f_{A,nano}$, $f_{A,micro}$) and Volume ($f_{V,pico}$, $f_{V,nano}$, $f_{V,micro}$) Concentration

	$f_{A,pico}$	$f_{A,nano}$	$f_{A,micro}$	$f_{V,pico}$	$f_{V,nano}$	$f_{V,micro}$
$f_{N,pico}$	0.74			0.46		
$f_{N,nano}$		0.82			0.42	
$f_{N,micro}$			0.85			0.61

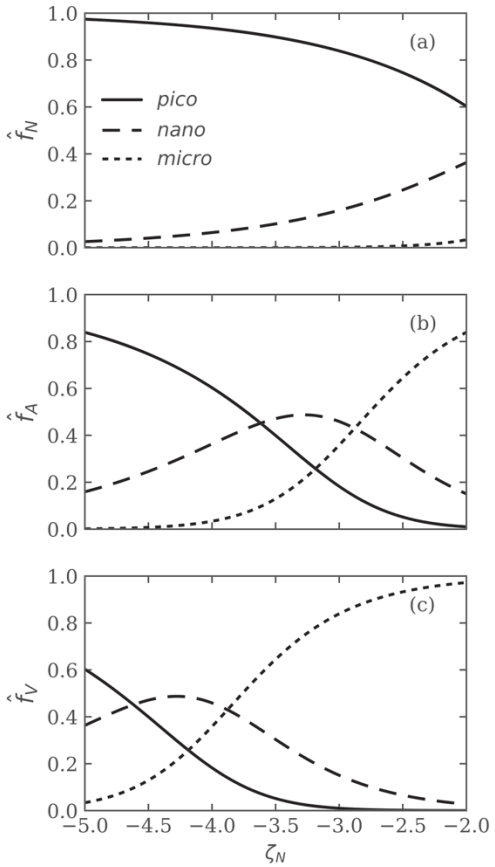


Figure 7. The theoretical fractional contributions of the picoplankton, $\hat{f}_{pico}$, nanoplankton, $\hat{f}_{nano}$, and microplankton, $\hat{f}_{micro}$, size classes to total particle (a) number, (b) cross-sectional area, and (c) volume concentration over the particle diameter range 0.8 to 120 μm calculated as a function of the power law slope ζ_N.

chlorophyll *a* fluorescence maximum. The particle assemblages in both samples are predominantly organic, with POC/SPM ratios exceeding 0.34. The distributions illustrated in Figures 5a and 5b are from a sample collected at a depth of 21 m in the Chukchi Sea. Phytoplankton pigment measurements indicate a high chlorophyll *a* concentration for this sample (2.3 mg m^{-3}) and suggest that the phytoplankton community was composed primarily of microplanktonic dinoflagellates, consistent with observation of a discrete peak in the PSD occurring near $D = 40$ μm (Figure 5a). In contrast, the distributions depicted in Figures 5c and 5d are from 56 m deep in the Beaufort Sea and with a much lower chlorophyll *a* concentration (0.5 mg m^{-3}). The phytoplankton community composition of this sample was largely composed of picophytoplankton species, consistent with a prominent peak in the size distribution centered near 1.6 μm (Figure 5c). This sample is additionally characterized by the presence of a second maximum around 10 μm and an additional feature around and above 30 μm, suggesting a diverse mix of plankton populations. Both samples shown in Figure 5 exhibit significant disagreements between the derived power law fits and measured data, such that the fitted values either overestimate or underestimate the measured data in different size ranges.

3.2. Relative Contributions of Size Classes to Particle Number, Area, and Volume Concentration

For each measured PSD, we quantified the fractional contribution of three different planktonic size ranges (f_{pico}, f_{nano}, and f_{micro}) to the total particle number, area, and volume concentration (Figure 6). In terms of particle number concentration, the contribution of the small-sized particles overwhelmingly dominates the numerical concentration with $f_{N,pico}$ averaging 90 ± 5.1% and always greater than 75%. The corresponding values of $f_{N,nano}$ range from 0.3% to 23%, with $f_{N,micro}$ never exceeding 0.2%. The fractional contributions of picoplankton and nanoplankton size classes to particle area concentration exhibit more variability than particle number with both $f_{A,pico}$ and $f_{A,nano}$ spanning a broad range (9–95% and 4–85%, respectively), while $f_{A,micro}$ again contributes only a relatively small fraction (8.5 ± 7.4%, always <35%). In contrast to particle number and area, the picoplankton size fraction has the smallest contributions to particle volume concentration ($f_{V,pico}$ in the range 1–62%, with only 7 samples >25%). The particle volume distribution is generally dominated by $f_{V,nano}$ (51 ± 17%) and $f_{V,micro}$ (39 ± 19%).

The fractional contribution of a size class in a given distribution, $N'(D)$, $A'(D)$, or $V'(D)$, is generally correlated with its contribution to other distributions. Table 2 summarizes the determination coefficients between size classes among the different measures of particle size. The contribution of the three size classes calculated from $N'(D)$ are more strongly correlated with the corresponding contributions determined from $A'(D)$ than with $V'(D)$. For example, the fractional contribution of the picoplankton size class to the total number concentration of particles, $f_{N,pico}$, is well correlated with its contribution to total particle area $f_{A,pico}$ ($R^2 = 0.74$) but to a lesser extent with its contribution to total particle volume $f_{V,pico}$ ($R^2 = 0.46$). The determination coefficients in Table 2 indicate that translating particle size classes from one measure of the PSD to another is not straightforward owing to the convolution of particle number and the area or volume weighting as a function of diameter.

3.3. Estimation of Particle Size Class Contributions From the Power Law Model

Figure 7 illustrates theoretical results obtained with calculations to show how varying the slope parameter of a single-slope power law model leads to changes in the predicted fractional contributions of the picoplankton, nanoplankton, and microplankton size classes to the total particle concentration in terms of number, area, and volume. In these calculations, the slope parameter ζ_N was varied between the values of -5 to -2, a broader range than observed in our field data set from the Arctic seas (Figure 3). The predicted picoplankton contribution to particle number concentration is dominant over the entire range of ζ_N, with contributions of the microplankton size class always low (<3.5%, Figure 7a). This pattern is consistent with our experimental data, in which $f_{N,micro}$ averaged about 0.03 ± 0.03% and never exceeded 0.18% (Figure 6). In contrast, for the volume distribution picoplankton are dominant only for the smallest (steepest) values of $\zeta_N \leq -4.75$, and microplankton contribute more than 50% over most of the range of ζ_N (Figure 7c). The patterns observed for the area concentration distribution are intermediate between the number and volume distributions (Figure 7b). Although the nanoplankton size fraction can be appreciable and even the largest contributor to area or volume over a narrow range of ζ_N, the value of $\hat{f}_{nano}$ never exceeds 0.5 in any of the three distributions.

We examined how well the single-slope power law approximation to our measurements of the particle number distribution predicts the fractional contributions of each size class to total particle number, area, and volume concentration. The fitted values of the slope parameter ζ_N (equation 2) determined from each

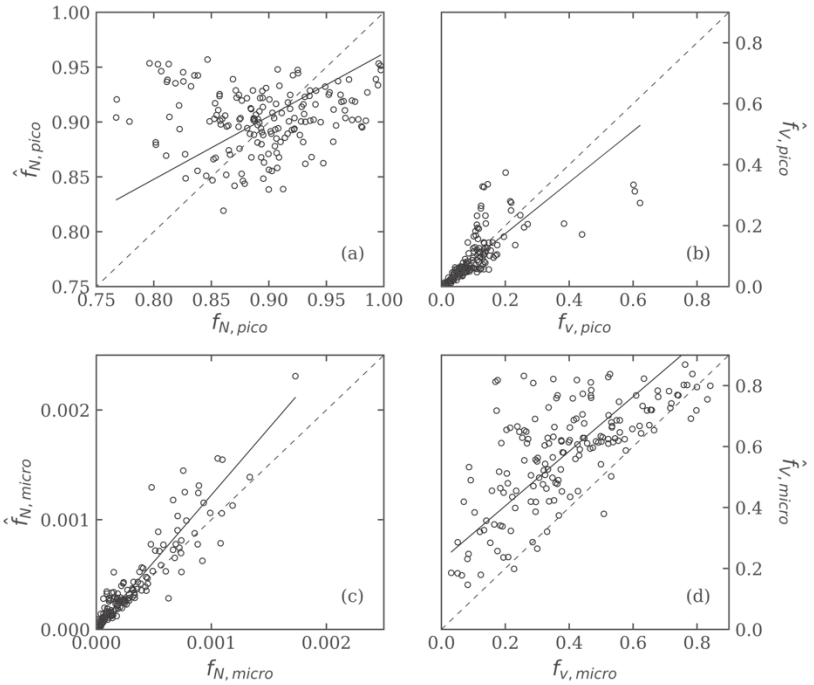


Figure 8. Scatter plots comparing measured values of the fractional contributions of picoplankton, f_{pico}, and microplankton, f_{micro}, size classes to those derived from a power law fit ($\hat{f}_{pico}$, $\hat{f}_{micro}$). (a,c) Fractional contribution to particle number concentration. (b,d) Fractional contribution to particle volume concentration. In each panel, the dashed line indicates the 1:1 line and the solid line represents the Model II regression line.

measured size distribution, $N'(D)$, were used to estimate the predicted (modeled) contribution of each size fraction ($\hat{f}_{N,pico}$, $\hat{f}_{N,nano}$, $\hat{f}_{N,micro}$). These modeled contributions were then compared with the actual contributions ($f_{N,pico}$, $f_{N,nano}$, $f_{N,micro}$) computed from equations 4a–4c using the measured data. Similar calculations were made on the basis of $A'(D)$ and $V'(D)$ distributions, which provided the comparison of modeled $\hat{f}_{A,pico}$, $\hat{f}_{A,nano}$, $\hat{f}_{A,micro}$ with measured $f_{A,pico}$, $f_{A,nano}$, $f_{A,micro}$ as well as the comparison of modeled $\hat{f}_{V,pico}$, $\hat{f}_{V,nano}$, $\hat{f}_{V,micro}$ with measured $f_{V,pico}$, $f_{V,nano}$, $f_{V,micro}$.

Scatter plots depicting four of these comparisons between the modeled and observed size class contributions are illustrated in Figure 8, and Table 3 provides a statistical evaluation for these relationships in addition to other comparisons not illustrated. In general, the correlations between the power law model predictions and actual observed values are strongest for the microplankton size class, that is, moderate to high correlation with the correlation coefficient R ranging from 0.68 for the particle volume distribution to 0.93 for the number distribution (Table 3, Figures 8c and 8d). However, the power law predictions of $\hat{f}_{N,micro}$, $\hat{f}_{A,micro}$, and $\hat{f}_{V,micro}$ are all characterized by significant positive bias (the median ratio of predicted to measured data, *MdR*, in the range 1.26 to 1.61) and relatively large values of median absolute percent difference between predicted and measured data (*MdAPD* between about 30% and 60%, Table 3). The weakest correlations between the power law model predictions and measured data are observed for the nanoplankton size class. Specifically, for these size fractions, the correlation coefficient R is always weak, ranging from 0.09 to 0.37 depending on the type of PSD (Table 3). For the picoplankton size class, the correlation between the power law predictions and measured data ranges from very weak (R = 0.1 for the particle number distribution; Figure 8a) to moderate or moderately high (R = 0.63 and 0.72 for area and volume distribution, respectively; Table 3). It is notable, however, that the case with moderately high correlation coefficient of 0.72 for the volume distribution is characterized by significant negative bias (*MdR* = 0.8) and increased *MdAPD* of 28% (Table 3, Figure 8b).

Overall, these results indicate that the role of the nanoplankton and microplankton size classes cannot be predicted well by the power law under any of the circumstances encountered in our data set, with typical prediction errors ranging from 16% to 61%. The best performance of the power law in terms of minimal error and bias is the prediction of the contribution of picoplankton size class to the particle number concentration; however, this represents a case in which picoplankton contribution is always dominant and the value of $\hat{f}_{N,pico}$ varies over a small range regardless of the power law slope. In addition, in this case, the correlation coefficient between the predicted $\hat{f}_{N,pico}$ and measured $f_{N,pico}$ is extremely low (R = 0.1). Thus, the general conclusion that stems from the analysis of our Arctic data set is that the power law fits to the measured PSDs do not provide acceptable estimates of the fractional contributions of picoplankton, nanoplankton, and microplankton size classes to total particle number, area, or volume concentration.

Table 3
Statistical Characterization of the Performance of a Power Law Model to Estimate the Fractional Contribution of Three Size Classes (f_{pico}, f_{nano}, f_{micro}) to the Particle Size Distribution Based on Particle Number (N), Cross-Sectional Area (A), and Volume (V) Concentration

Variable	Slope	Intercept	*R*	*MdR*	*MdB*	*MdAPD* [%]	*RMSD*
$f_{N,pico}$	0.57	0.39	0.10	1.00	<0.001	3.71	0.055
$f_{N,nano}$	0.57	0.04	0.09	1.00	<0.001	33.4	0.055
$f_{N,micro}$	1.22	0.00	0.93	1.26	<0.001	29.2	0.000
$f_{A,pico}$	0.86	0.08	0.63	1.02	0.005	16.0	0.128
$f_{A,nano}$	0.40	0.24	0.22	0.91	−0.044	18.2	0.151
$f_{A,micro}$	1.22	0.02	0.81	1.61	0.029	61.3	0.068
$f_{V,pico}$	0.84	0.01	0.72	0.80	−0.009	28.3	0.067
$f_{V,nano}$	0.56	0.05	0.37	0.69	−0.148	32.9	0.239
$f_{V,micro}$	0.90	0.23	0.68	1.51	0.158	50.5	0.235

Note. All comparisons are based on n = 168 measurements of the particle size distribution. Modeled values were determined by fitting the measured size distribution to a single slope power law function (equation 2) and calculating the predicted contributions of each size class. The slope, intercept, and correlation coefficient R of a Model II linear regression between modeled and measured variables is provided. Other statistical descriptors include the median ratio, *MdR*, median bias, *MdB*, and median absolute percent difference, *MdAPD*, between modeled and measured values. *RMSD* is the root mean square deviation. The values provided in rows 1, 7, 3, and 9 correspond to the relationships illustrated in Figures 8a to 8d, respectively.

3.4. Estimation of Particle Size Class Contributions From Percentiles of the Cumulative Distribution Functions

The cumulative distribution functions for particle number, $CDF_N(D)$, cross-sectional area, $CDF_A(D)$, and volume concentration, $CDF_V(D)$, obtained for all our measurements with equation 3 are depicted in Figure 9. For reference, predicted cumulative distribution functions, $\widehat{CDF}$, are also illustrated for the power law model with slopes of ζ_N = −4.5, −3.5, and −2.5, which bracket the experimental median value of −3.55 obtained in this study. The $CDF_N(D)$ generally shows a rapid increase with increasing particle diameter as counts of large particles are always very low relative to the number of smaller particles. In contrast, the $CDF_A(D)$ and to greater extent the $CDF_V(D)$ show a less rapid increase, and generally a greater range in the

particle diameter associated with a given percentile as specified by a given value of $CDF_A(D)$ or $CDF_V(D)$. These patterns are evident from the size-dependent probability distributions of specific percentile values (Figure 10).

We examined relationships between specific percentile values of particle diameter obtained from the $CDF_N(D)$, $CDF_A(D)$, and $CDF_V(D)$ with the fractional contributions of the picoplankton, nanoplankton, and microplankton size classes to total particle concentrations N_t, A_t, and V_t (Table 4). This analysis aimed at identifying a relatively simple set of percentile diameters, which can serve as reasonable proxies for the contributions of the three particle size classes to diverse particulate assemblages encountered in natural waters, especially in the investigated Arctic waters. The median particle size has been previously used as a single descriptor of the PSD (e.g., Woźniak et al., 2010), but in our data set, this metric was generally a poor predictor of the three size class contributions with determination coefficients $R^2 < 0.6$ (see the values for $D_N(50)$, $D_A(50)$, and $D_V(50)$ in Table 4). The sole exception was that the median diameter based on the particle volume distribution, $D_V(50)$, was a reasonable predictor of the microplankton size class contribution to the total volume concentration, $f_{V,micro}$ ($R^2 = 0.85$).

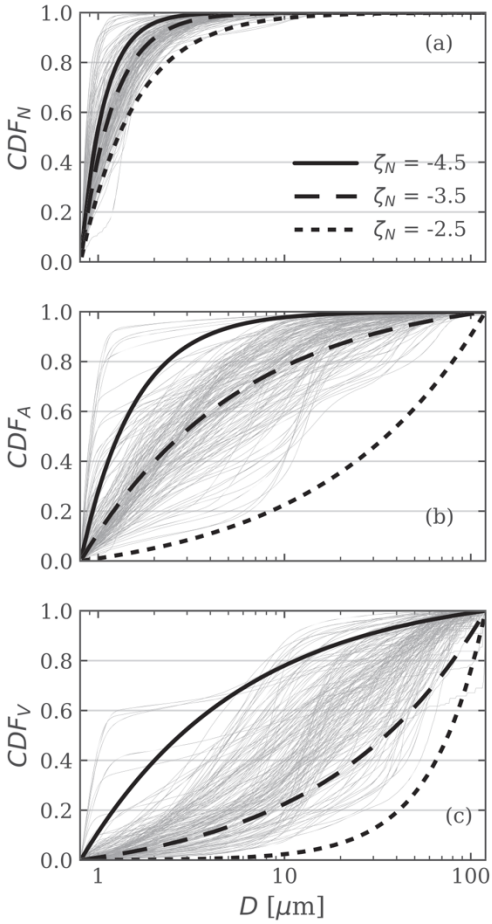


Figure 9. Measured cumulative distribution functions, *CDF*, of particle (a) number, (b) cross-sectional area, and (c) volume concentration. In each panel, all measured distributions are shown in gray and the distribution predicted by a power law model corresponding to different values of the slope ζ_N are shown for comparison.

The percentile diameters $D_N(80)$, $D_N(80)$, and $D_N(99.9)$ from the cumulative distribution $CDF_N(D)$ exhibited the highest determination coefficients with the respective size class contributions $f_{N,pico}$, $f_{N,nano}$, and $f_{N,micro}$ to the particle number concentration N_t (Table 4). The determination coefficients between all percentile diameters of the cumulative distribution function $CDF_A(D)$ with $f_{A,pico}$ and $f_{A,nano}$ were generally low and only $f_{A,micro}$ exhibited strong relationships, with the highest determination coefficient observed for $D_A(90)$ ($R^2 = 0.94$). Similarly, the microplankton contribution to the particle volume concentration, $f_{V,micro}$, was the only fraction demonstrating a high determination coefficient associated with percentiles derived from the cumulative distribution $CDF_V(D)$, with the percentile diameter $D_V(60)$ yielding the strongest relationship.

These four best-performing relationships, i.e. $f_{N,pico}$ versus $D_N(80)$, $f_{N,nano}$ versus $D_N(80)$, $f_{A,micro}$ versus $D_A(90)$, and $f_{V,micro}$ versus $D_V(60)$, are illustrated in Figure 11. The first three relationships are described by Model I linear regression models (Figures 11a–11c), while the fourth relationship is parameterized as a second-order polynomial function (Figure 11d). The fitted equations for these relationships are

$$\hat{f}_{N,pico} = -0.178\, D_N(80) + 1.17, \tag{5a}$$

$$\hat{f}_{N,nano} = 0.178\, D_N(80) - 0.17, \tag{5b}$$

$$\hat{f}_{A,micro} = 0.006\, D_A(90) - 0.02, \tag{5c}$$

$$\hat{f}_{V,micro} = -0.00012\, D_V(60)^2 + 0.020\, D_V(60) + 0.027. \tag{5d}$$

The size class contributions predicted by these relationships based on percentile diameters, $\hat{f}_{N,pico}$, $\hat{f}_{N,nano}$, $\hat{f}_{A,micro}$, $\hat{f}_{V,micro}$, are compared with measured values, $f_{N,pico}$, $f_{N,nano}$, $f_{A,micro}$, $f_{V,micro}$, in Figure 12. All four comparisons of modeled versus measured values exhibit high correlation ($R \geq 0.95$) with slope values of the Model II linear regression near one (Table 5). In addition, the statistics characterizing model bias and random error are all very good (Table 5) and greatly

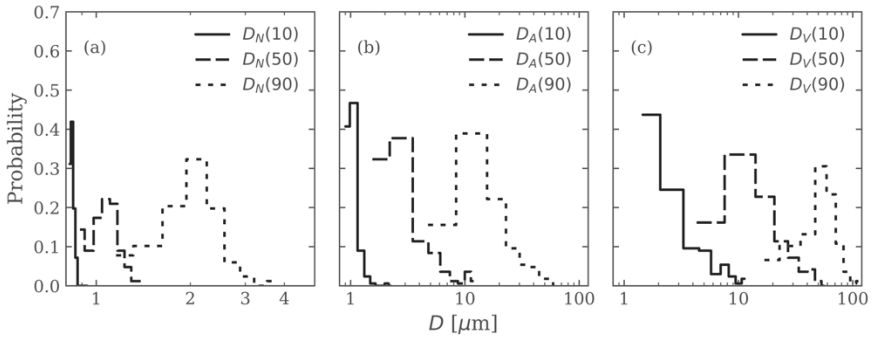


Figure 10. Frequency distributions of the particle diameter *D* representing the 10th, 50th, and 90th percentiles of the particle (a) number, (b) cross-sectional area, and (c) volume concentration. Note the different scaling of the *x*-axis in (a).

improved compared to the power law model results provided in Table 3. Specifically, the *MdR* values for the four relationships shown in Figure 11 indicate that the median bias is no greater than 1% and *MdAPD* generally less than 10%, with the highest value of only 13% for the case when $f_{N,nano}$ is predicted from $D_N(80)$ (Table 5). These results indicate that the use of specific percentile diameters derived from the cumulative distribution functions of particle size offers a highly promising approach for estimating the fractional contributions of diverse size classes, and warrants further investigation.

Table 4
Values of the Determination Coefficient, R^2, Between Diameters Representing Specific Percentiles (Given in Percent) of the Cumulative Distribution Function With the Corresponding Fractional Contributions of Three Size Classes to the Distribution

$CDF_N(D)$	$f_{N,pico}$	$f_{N,nano}$	$f_{N,micro}$	$CDF_A(D)$	$f_{A,pico}$	$f_{A,nano}$	$f_{A,micro}$	$CDF_V(D)$	$f_{V,pico}$	$f_{V,nano}$	$f_{V,micro}$
[μm]				[μm]				[μm]			
$D_N(1)$	<0.01	<0.01	<0.01	$D_A(1)$	0.30	0.38	<0.01	$D_V(1)$	0.23	0.01	0.11
$D_N(5)$	0.30	0.30	0.03	$D_A(5)$	0.45	0.47	0.01	$D_V(5)$	0.26	0.04	0.18
$D_N(10)$	0.36	0.36	0.01	$D_A(10)$	0.45	0.46	0.02	$D_V(10)$	0.32	0.11	0.34
$D_N(20)$	0.32	0.32	0.01	$D_A(20)$	0.37	0.36	0.03	$D_V(20)$	0.32	0.27	0.56
$D_N(25)$	0.37	0.37	0.01	$D_A(25)$	0.40	0.36	0.04	$D_V(25)$	0.30	0.33	0.62
$D_N(30)$	0.41	0.41	<0.01	$D_A(30)$	0.44	0.36	0.08	$D_V(30)$	0.29	0.38	0.67
$D_N(40)$	0.46	0.46	<0.01	$D_A(40)$	0.52	0.35	0.17	$D_V(40)$	0.28	0.46	0.77
$D_N(50)$	0.53	0.53	<0.01	$D_A(50)$	0.57	0.31	0.31	$D_V(50)$	0.28	0.54	0.85
$D_N(60)$	0.66	0.66	<0.01	$D_A(60)$	0.58	0.26	0.47	$D_V(60)$	0.25	0.59	**0.89**
$D_N(70)$	0.81	0.81	<0.01	$D_A(70)$	0.47	0.11	0.71	$D_V(70)$	0.18	0.63	0.86
$D_N(75)$	0.88	0.88	0.01	$D_A(75)$	0.38	0.05	0.81	$D_V(75)$	0.15	0.61	0.80
$D_N(80)$	**0.94**	**0.94**	0.02	$D_A(80)$	0.32	0.02	0.87	$D_V(80)$	0.10	0.56	0.69
$D_N(90)$	0.87	0.87	0.08	$D_A(90)$	0.26	<0.01	**0.94**	$D_V(90)$	0.04	0.52	0.57
$D_N(95)$	0.54	0.53	0.17	$D_A(95)$	0.23	<0.01	0.86	$D_V(95)$	0.04	0.43	0.49
$D_N(99)$	0.32	0.31	0.45	$D_A(99)$	0.13	<0.01	0.57	$D_V(99)$	0.04	0.23	0.27
$D_N(99.9)$	0.13	0.13	0.83	$D_A(99.9)$	0.04	0.01	0.31	$D_V(99.9)$	0.07	0.06	0.12

Note. Relationships between percentile diameters derived from the cumulative distribution of particle number concentration, $CDF_N(D)$, with the fractional contributions of picoplankton, $f_{N,pico}$, nanoplankton, $f_{N,nano}$, and microplankton, $f_{N,micro}$, size classes to the total particle number concentration are shown in columns 2 to 4. Similarly, relationships for the cumulative distributions of particle area, $CDF_A(D)$, and volume, $CDF_V(D)$, concentration are provided in columns 6 to 8 and columns 10 to 12, respectively. Values in boldface correspond to the four best-performing relationships given in equations 5a–5d and illustrated in Figure 11.

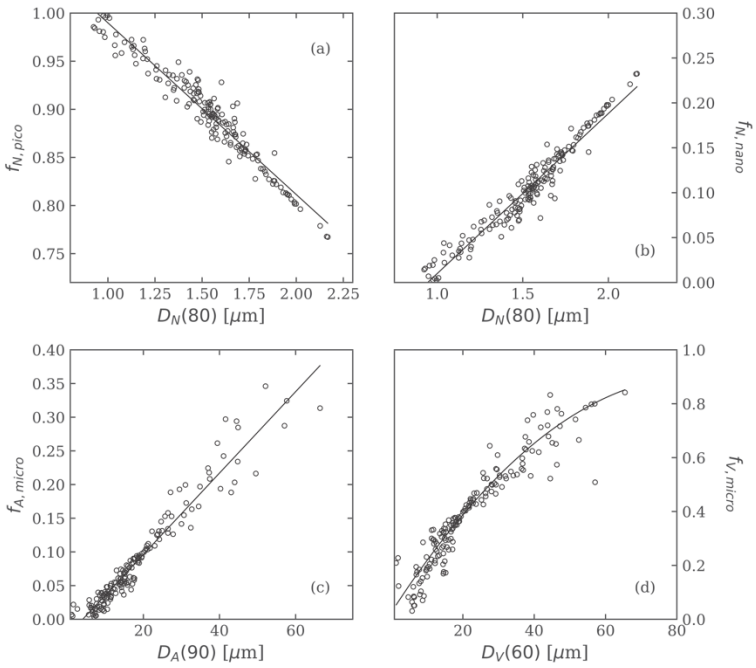


Figure 11. Example relationships between specific percentiles of the PSD and the fractional contribution of three size classes to the entire distribution. (a,b) Fractional contribution to total particle number concentration N_t of the picoplankton, $f_{N,pico}$, and nanoplankton, $f_{N,nano}$, size classes as a function of the 80th percentile particle diameter, $D_N(80)$. (c) Fractional contribution to total particle cross-sectional area concentration A_t of the microplankton size class, $f_{A,micro}$, as a function of the 90th percentile particle diameter, $D_A(90)$. (d) Fractional contribution to total particle volume concentration V_t of the microplankton size class, $f_{V,micro}$, as a function of the 60th percentile particle diameter, $D_V(90)$. The solid line in panels (a–c) depict a Model I linear regression fitted to the observations and in panel (d) depicts a fitted polynomial function obtained through a nonlinear least squares method.

The cumulative distribution functions depicted in Figure 9 indicate that the $CDF_V(D)$ exhibits the greatest range of variation between PSDs, followed by the $CDF_A(D)$ and then the $CDF_N(D)$. However, percentile values derived from $CDF_N(D)$ generally had strong determination coefficients with the fractional size class contributions, especially for picoplankton and nanoplankton size classes (Table 4). This is likely because all PSDs in our data set exhibited a very small value of $f_{N,micro}$ (Figure 6), so the problem of determining three size classes from a single variable was reduced to the simpler problem of predicting two classes ($f_{N,pico}$ and $f_{N,nano}$) that must sum to a value nearly equal to 1. Because $f_{V,pico}$ was also generally quite small, similar reasoning suggests that percentile diameters from the $CDF_V(D)$ should also show strong correlations with $f_{V,nano}$ and $f_{V,micro}$. Although relatively strong determination coefficients are observed with $f_{V,micro}$ (e.g., $R^2 \geq 0.8$ for diameters corresponding to 50th to 75th percentile range), similarly strong correlations are not seen for $f_{V,nano}$. This can be attributed to the observation that although values of $f_{V,pico}$ are generally small ($10 \pm 8\%$), they are still appreciably greater than zero so that the closure of the budget of contributions to particle volume concentration still requires an accurate estimation of three size classes.

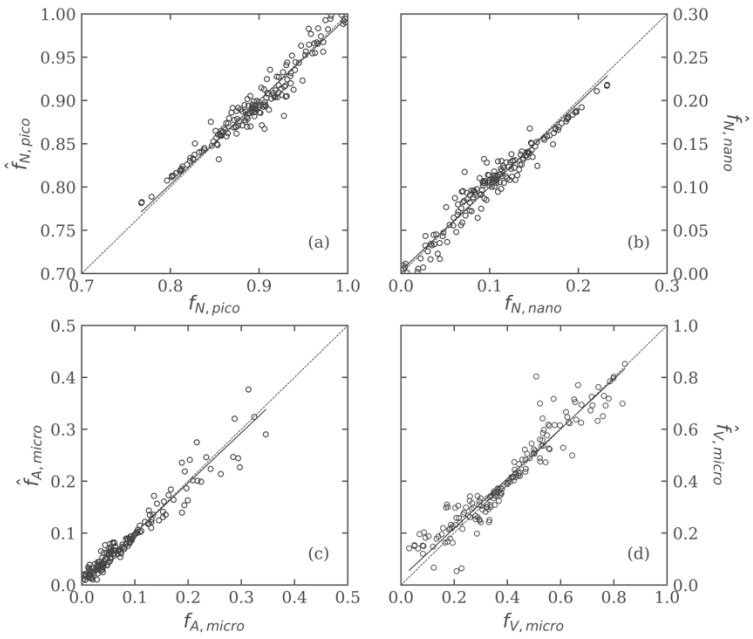


Figure 12. Comparison of model-derived with measured values for the relationships depicted in Figures 11a–11d and described in equations 5a–5d. In each panel, the dotted line represent the 1:1 line and the solid line represents the Model II regression line.

The determination coefficients for a given size fraction are observed to be generally highest at the percentile values where the range of corresponding diameters most closely match the diameter used to demarcate one size fraction from another and decrease as these values move further away from these limits. For example, the highest determination coefficient for $f_{V,micro}$ is observed with the 60th percentile diameter of the volume distribution, $D_V(60)$, so this percentile diameter was selected as the best proxy for predicting $f_{V,micro}$ (Table 4, Figure 11d). The mean value of $D_V(60)$ in our data set is 22.3 ± 13.3 μm, which is very close to the value of 20 μm that defines the lower boundary of the microplankton size range (Figure 13). Percentiles higher than 60% corresponded to diameters significantly larger than 20 μm, for example, the average value of $D_V(90)$ was 54.6 ± 19.9 μm. Thus, the portion of $CDF_V(D)$ corresponding to relatively high values of this function ($CDF_V(D) > 0.6$) carried more information about the distribution of particles within the microplankton size fraction than the magnitude of $f_{V,micro}$ relative to magnitudes of $f_{V,pico}$ and $f_{V,nano}$. Similar reasoning also explains why $D_N(80)$, which has an average value of 1.53 ± 0.3 μm (Figure 13), most accurately predicted the contributions $f_{N,pico}$ and $f_{N,nano}$ (Table 4, Figures 11a and 11b), and $D_A(90)$ with a value of 18.1 ± 11.9 μm (Figure 13) was the best proxy for predicting $f_{A,micro}$ (Table 4, Figure 11c). These results imply that different combinations of PSD size ranges and desired size intervals will result in different percentiles correlating best with those size classes.

Table 5
Similar to Table 3, but for Predicted Values Calculated From Percentiles of the Cumulative Distribution Function Using the Relationships in Equations 5a–5d

Variable	Slope	Intercept	*R*	*MdR*	*MdB*	*MdAPD* [%]	*RMSD*
$f_{N,pico}$	0.97	0.03	0.97	1.00	0.001	0.89	0.0123
$f_{N,nano}$	0.97	<0.01	0.97	0.99	−0.001	7.84	0.0123
$f_{A,nano}$	0.97	<0.01	0.97	1.00	<0.001	13.35	0.0183
$f_{V,micro}$	0.96	0.03	0.95	1.01	0.005	7.48	0.0603

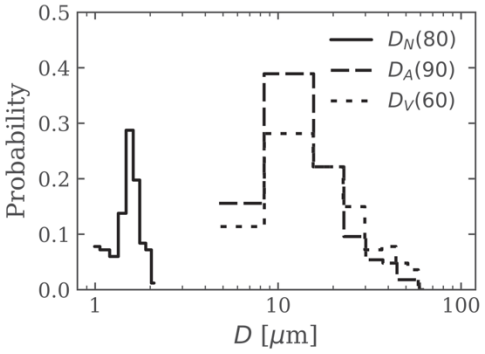


Figure 13. Frequency distribution of the three percentile diameters used in the relationships depicted in Figure 11.

4. Conclusions

Our measurements on seawater samples from the Arctic Ocean indicate that the size distribution of suspended particle assemblages in this environment exhibits a wide range of variation in both the magnitude and shape of the distribution of particle number, cross-sectional area, and volume concentration. The high size resolution of the Coulter Counter measurements reveals that changes in the slope of the size distribution generally occur over the measured size range of 0.8 to 120 μm, and additionally shows the frequent occurrence of peaks in the distribution associated with planktonic populations. In our data set, these deviations from an idealized single-slope power law parameterization are a consistent feature in nearly all measured size distributions of natural particle assemblages suspended in seawater.

Analysis of this data set was used to quantify the relative contributions of three particle size classes (picoplankton, nanoplankton, and microplankton) to the integrated size distribution. The number of particles per unit volume always exhibits an overall strong decrease with increasing particle diameter, and thus the contribution of particles in the picoplankton size range to the total number concentration of particles is always high (77% to 99%) and larger particles have only small contributions. The size distributions based on particle cross-sectional area and volume are characterized by an increasing contribution of nanoplankton and microplankton size classes at the expense of decreasing picoplankton contribution. Importantly, however, the complex and varied shapes exhibited by the PSD result in a large degree of variability among the contributions of the size classes to particle area and volume concentration and also limits the straightforward translation of a size class contribution from one distribution to another.

In many oceanic studies, simple one or two parameter models such as the power law distribution are used to describe the shape of the PSD in aquatic ecosystems. Our field measurements indicate that while a power law model may describe the overall general trend of the PSD in natural marine particle assemblages, significant departures from a single power law slope commonly occur throughout the size range. As a consequence, application of this model can yield considerable error in predicting the fractional contributions of individual size classes to the total number, area, or volume concentration.

Alternative metrics such as mean or median particle size have also been used as a simple means to characterize PSDs. We examined the utility of specific percentile diameters derived from the cumulative distribution function of particle number, area, or volume concentration to estimate the relative contribution of the three size classes to these distributions. Although median particle diameter was generally a poor predictor of size class contributions, other percentile diameters were observed to be strongly correlated with the contributions of individual size classes. These results suggest that the development and use of such metrics can lead to significantly improved estimates of individual size class contributions compared to the power law model. In this study, the best-performing percentile diameters for discriminating the contributions of different size classes were generally those which displayed a marked variation among individual PSDs, and exhibited a frequency distribution that straddled the diameter used to delineate individual size classes. For the particle size classes chosen in this study, these best-performing percentiles consisted of $D_N(80)$ for estimating both $f_{N,pico}$ and $f_{N,nano}$, $D_A(90)$ for the estimation of $f_{A,micro}$, and $D_V(60)$ for the estimation of $f_{V,micro}$. The choice of percentile diameters for developing a given approach will depend on the overall size range in question, as well as the desired size class definitions. These results expand on our previous studies of the optical properties of particle suspensions in which metrics such as $D_V(50)$ and $D_V(90)$ were used as indicators for the relative contributions of small and large particles (Koestner et al., 2020; Woźniak et al., 2010).

Our analysis utilizes size classes based on traditional size classifications of oceanic plankton (i.e., picoplankton, nanoplankton, microplankton), but it should be emphasized that our measurements encompass all particles suspended in seawater including living and nonliving particles, and both organic and minerogenic material. Although our analysis is based on measurements collected in Arctic seas, the wide diversity of particle assemblage types encountered in this region likely spans that of many marine habitats including the

pelagic, coastal, and estuarine environments. It is important to recognize, however, that the measured PSDs in this study are limited to a finite range of 0.8 to 120 μm, and thus contributions of smaller or larger particles outside of this size range are not assessed. Submicron particles (e.g., colloids or nanoparticles) as well as much larger particles (e.g., plankton chains or colonies, aggregates) can make a variable but significant contribution to both particle mass and size concentration in certain oceanic environments. In addition, marine particles can exhibit complex, nonspherical shapes that violate the assumptions employed in this study. The uncertainties in our results that arise from these limitations will depend on both the environment and the intended use of the PSD.

Optical measurements obtained from autonomous underwater vehicles or from above water platforms (e.g., aircraft and satellites) have the potential to extend observations of the PSD to the synoptic spatial and temporal scales necessary for studies of particle dynamics in the global ocean. Strong linkages between the PSD and the optical properties of seawater have long been recognized (e.g., Brown & Gordon, 1974; Jerlov, 1976; Kullenberg, 1974; Morel, 1973), and numerous studies have demonstrated relationships between specific metrics derived from the PSD with the light absorption and scattering characteristics of the particle suspension (Bowers et al., 2009; Briggs et al., 2013; Ciotti et al., 2002; Slade & Boss, 2015; Twardowski et al., 2001; Woźniak et al., 2010). These studies have typically utilized metrics such as mean or median particle size to characterize the PSD. Our results suggest that alternative metrics of the PSD, in particular, specific percentiles from the cumulative distribution functions of particle number, area, or volume concentration and chosen to target specific size class definitions, may provide a better means to represent the complex shape of the oceanic PSD. To take full advantage of this approach, further work is needed to establish quantitative relationships between these alternative percentile diameters and suitable proxies derived from measurements of seawater optical properties.

Acknowledgments

We thank Marcel Babin, Kevin Arrigo, Shigeto Nishino, Toru Hirawake, and all cruise participants for assistance in the field. Support for this work was provided by the National Aeronautics and Space Administration (NNX07AR20G, NNX10AG05G), the Japan Aerospace Exploration Agency (GCOM-C 311), and the National Science Foundation (OPP-1822021). Any opinions, findings, and conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of these agencies. The authors declare no conflicts of interest with respect to the results of this paper. Data sets used in this study are publicly available in the LEFE-CYBER database (http://www.obs-vlfr.fr/proof/index2.php; MALINA cruise), the NASA SeaWiFS Bio-optical Archive and Storage System (https://seabass.gsfc.nasa.gov/; ICESCAPE 2010 and 2011 cruises), and the Data and Sample Research System for Whole Cruise Information database of the Japan Agency for Marine-Earth Sciences (https://doi.org/10.17596/0001879; MR1705-C cruise). We thank two anonymous reviewers for providing comments on the manuscript.

References

Agagliate, J., Lefering, I., & McKee, D. (2018). Forward modelling of inherent optical properties from flow cytometry estimates of particle size and refractive index. *Applied Optics*, *57*(8), 1777–1788. https://doi.org/10.1364/AO.57.001777

Agrawal, Y. C., & Pottsmith, H. C. (2000). Instruments for particle size and settling velocity observations in sediment transport. *Marine Geology*, *168*, 89–114. https://doi.org/10.1016/S00253227(00)00044-X

Bader, H. (1970). The hyperbolic distribution of particle sizes. *Journal of Geophysical Research*, *75*, 2822–2830. https://doi.org/10.1029/JC075i015p02822

Baker, E. T., & Lavelle, J. W. (1984). The effect of particle size on the light attenuation coefficient of natural suspensions. *Journal of Geophysical Research*, *89*, 8197–8203.

Bernard, S., Shillington, F. A., & Probyn, T. A. (2007). The use of equivalent size distributions of natural phytoplankton assemblages for optical modeling. *Optics Express*, *15*(5), 1995–2007. https://doi.org/10.1364/OE.15.001995

Bowers, D. G., Binding, C. E., & Ellis, K. M. (2007). Satellite remote sensing of the geographical distribution of suspended particle size in an energetic shelf sea. *Estuarine, Coastal and Shelf Science*, *73*, 457–466. https://doi.org/10.1016/j.ecss.2007.02.005

Bowers, D. G., Braithwaite, K. M., Nimmo-Smith, W. A. M., & Graham, G. W. (2009). Light scattering by particles suspended in the sea: The role of particle size and density. *Continental Shelf Research*, *29*, 1748–1755. https://doi.org/10.1016/j.csr.2009.06.004

Briggs, N. T., Slade, W. H., Boss, E., & Perry, M. J. (2013). Method for estimating mean particle size from high-frequency fluctuations in beam attenuation or scattering measurements. *Applied Optics*, *52*(27), 6710–6725. https://doi.org/10.1364/AO.52.006710

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, *85*, 1771–1789. https://doi.org/10.1890/03-9000@10.1002

Brown, O. B., & Gordon, H. R. (1974). Size-refractive index distribution of clear coastal water particulates from light scattering. *Applied Optics*, *13*(12), 2874–2881. https://doi.org/10.1364/AO.13.002874

Brun-Cottan, J. C. (1971). Etude de la granulométrie des particules marines, mesures effectuées avec un compteur Coulter. *Cahiers Océanographiques*, *23*, 193–205.

Buonassissi, C. J., & Dierssen, H. M. (2010). A regional comparison of particle size distributions and the power law approximation in oceanic and estuarine surface waters. *Journal of Geophysical Research*, *115*, C10028. https://doi.org/10.1029/2010JC006256

Burd, A. B. (2013). Modeling particle aggregation using size class and size spectrum approaches. *Journal of Geophysical Research: Oceans*, *118*, 3431–3443. https://doi.org/10.1002/jgrc.20255

Carder, K. L., Beardsley, G. F. Jr., & Pak, H. (1971). Particle size distributions in the eastern equatorial Pacific. *Journal of Geophysical Research*, *76*, 5070–5077.

Chisholm, S. W. (1992). Phytoplankton size. In P. G. Falkowski, & A. D. Woodhead (Eds.), *Primary productivity and biogeochemical cycles in the sea*, (pp. 213–237). New York, N. Y: Plenum Press.

Ciotti, Á., Lewis, M. R., & Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnology and Oceanography*, *47*, 404–417. https://doi.org/10.4319/lo.2002.47.2.0404

Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M., & Charnov, E. L. (2001). Effects of size and temperature on metabolic rate. *Science*, *293*(5538), 2248–2251. https://doi.org/10.1126/science.1061967

Graham, G. W., & Nimmo-Smith, W. A. (2010). The application of holography to the analysis of size and settling velocity of suspended cohesive sediments. *Limnology and Oceanography: Methods*, *8*, 1–15.

Hansen, P. J., Bjørnsen, P. K., & Hansen, B. W. (1997). Zooplankton grazing and growth: Scaling within the 2-2,-µm body size range. *Limnology and Oceanography*, *42*, 687–704. https://doi.org/10.4319/lo.1997.42.4.0687

IOCCG (2014). Phytoplankton functional types from space. In S. Sathyendranath (Ed.), *Reports of the International Ocean Colour Coordinating Group*, (Vol. 15). Dartmouth, N.S: International Ocean Colour Coordinating Group. https://doi.org/10.1111/imre.12086

Jackson, G. A. (1995). Comparing observed changes in particle size spectra with those predicted using coagulation theory. *Deep-Sea Research Part II*, *42*, 159–184.

Jackson, G. A., Maffione, R., Costello, D. K., Alldredge, A. L., Logan, B. E., & Dam, H. G. (1997). Particle size spectra between 1 µm and 1 cm at Monterey Bay determined using multiple instruments. *Deep Sea Research Part I: Oceanographic Research Papers*, *44*, 1739–1767.

Jennings, S., & Warr, K. J. (2003). Smaller predator-prey body size ratios in longer food chains. *Proceedings of the Royal Society of London - Series B: Biological Sciences*, *270*(1522), 1413–1417. https://doi.org/10.1098/rspb.2003.2392

Jerlov, N. G. (1976). *Marine optics*. Amsterdam: Elsevier.

Jonasz, M. (1983). Particle-size distributions in the Baltic. *Tellus Series B: Chemical and Physical Meteorology*, *35*, 346–358. https://doi.org/10.3402/tellusb.v35i5.14624

Jonasz, M. (1987). Nonspherical sediment particles: Comparison of size and volume distributions obtained with an optical and a resistive particle counter. *Marine Geology*, *78*, 137–142. https://doi.org/10.1016/0025-3227(87)90072-7

Jonasz, M., & Fournier, G. (2007). *Light scattering by particles in water: Theoretical and experimental foundations*. San Diego, CA: Academic Press.

Kiefer, D. A., & Berwald, J. (1992). A random encounter model for the microbial planktonic community. *Limnology and Oceanography*, *37*, 457–467. https://doi.org/10.4319/lo.1992.37.3.0457

Kiørboe, T. (2000). Colonization of marine snow aggregates by invertebrate zooplankton: Abundance, scaling, and possible role. *Limnology and Oceanography*, *45*, 479–484. https://doi.org/10.4319/lo.2000.45.2.0479

Kiørboe, T., Grossart, H.-P., Ploug, H., Tang, K., & Auer, B. (2004). Particle-associated flagellates: Swimming patterns, colonization rates, and grazing on attached bacteria. *Aquatic Microbial Ecology*, *35*, 141–152. https://doi.org/10.3354/ame035141

Kitchen, J. C., Menzies, D., Pak, H., & Zaneveld, J. R. V. (1975). Particle size distributions in a region of coastal upwelling analyzed by characteristic vectors. *Limnology and Oceanography*, *20*, 775–783.

Knap, A. H., Michaels, A., Close, A. R., Ducklow, H., & Dickson, A. G. (1996). *Protocols for the Joint Global Ocean Flux Study (JGOFS) core measurements (Vol. 19). Reprint of the IOC Manuals and Guides (Vol. 29)*. Paris: UNESCO.

Koestner, D., Stramski, D., & Reynolds, R. A. (2020). Assessing the effects of particle size and composition on light scattering through measurements of size-fractionated seawater samples. *Limnology and Oceanography*, *65*, 173–190. https://doi.org/10.1002/lno.11259

Kostadinov, T. S., Siegel, D. A., & Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *Journal of Geophysical Research*, *114*, C09015. https://doi.org/10.1029/2009JC005303

Kostadinov, T. S., Siegel, D. A., & Maritorena, S. (2010). Global variability of phytoplankton functional types from space: Assessment via the particle size distribution. *Biogeosciences*, *7*, 3239–3257. https://doi.org/10.5194/bg-7-3239-2010

Kullenberg, G. (1974). Observed and computed scattering functions. In N. G. Jerlov, & E. Steeman-Nielsen (Eds.), *Optical aspects of oceanography*, (pp. 25–45). London: Academic Press.

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, *11*, 2016–2040. https://doi.org/10.1111/j.1365-2486.2005.1004.x

Moore, C., Barnard, A., Fietzek, P., Lewis, M. R., Sosik, H. M., White, S., & Zielinski, O. (2009). Optical tools for ocean monitoring and research. *Ocean Science*, *5*, 661–684.

Morel, A. (1973). Diffusion de la lumière par les eaux de mer; résultats expérimentaux et approche théorique. AGARD Lect. Ser., 3.1.1.–3.1.76. North Atlantic Treaty Organization.

Morel, A., & Bricaud, A. (1986). Inherent optical properties of algal cells including picoplankton: Theoretical and experimental results. *Canadian Bulletin of Fisheries and Aquatic Sciences*, *214*, 521–559.

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Frontiers in Marine Science*, *4*, 41. https://doi.org/10.3389/fmars.2017.00041

Parsons, T. R. (1969). The use of particle size spectra in determining the structure of a plankton community. *Journal of the Oceanographic Society of Japan*, *25*, 172–181.

Parsons, T. R., Maita, Y., & Lalli, C. M. (1984). *A manual of chemical and biological methods for seawater analysis*. Oxford, England: Pergamon Press.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., & Gorsky, G. (2010). The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnology and Oceanography: Methods*, *8*, 462–473.

Platt, T., & Denman, K. L. (1978). The structure of pelagic marine ecosystems. *Conseil permanent international pour l'exploration de la mer, première réunion*, *173*, 60–65.

Ploug, H., & Grossart, H.-P. (2000). Bacterial growth and grazing on diatom aggregates: Respiratory carbon turnover as a function of aggregate size and sinking velocity. *Limnology and Oceanography*, *45*, 1467–1475. https://doi.org/10.4319/lo.2000.45.7.1467

Reynolds, R. A., Stramski, D., & Neukermans, G. (2016). Optical backscattering of particles in Arctic seawater and relationships to particle mass concentration, size distribution, and bulk composition. *Limnology and Oceanography*, *61*, 1869–1890. https://doi.org/10.1002/lno.10341

Reynolds, R. A., Stramski, D., Wright, V. M., & Woźniak, S. B. (2010). Measurements and characterization of particle size distributions in coastal waters. *Journal of Geophysical Research*, *115*, C08024. https://doi.org/10.1029/2009JC005930

Risović, D. (1993). Two-component model of sea particle size distribution. *Deep Sea Research Part I: Oceanographic Research Papers*, *40*, 1459–1473. https://doi.org/10.1016/0967-0637(93)90123-K

Sheldon, R. W., Prakash, A., & Sutcliffe, W. H. Jr. (1972). The size distribution of particles in the ocean. *Limnology and Oceanography*, *17*, 327–340.

Shi, W., & Wang, M. (2019). Characterization of suspended particle size distribution in global highly turbid waters from VIIRS measurements. *Journal of Geophysical Research: Oceans*, *124*, 3796–3817. https://doi.org/10.1029/2018JC014793

Sieburth, J. M., Smetacek, V., & Lenz, J. (1978). Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnology and Oceanography*, *23*, 1256–1263. https://doi.org/10.4319/lo.1978.23.6.1256

Slade, W. H., & Boss, E. (2015). Spectral attenuation and backscattering as indicators of average particle size. *Applied Optics*, *54*(24), 7264–7277. https://doi.org/10.1364/AO.54.007264

Stemmann, L., & Boss, E. (2012). Plankton and particle size and packaging: From determining optical properties to driving the biological pump. *Annual Review of Marine Science*, *4*, 263–290. https://doi.org/10.1146/annurev-marine-120710-100853

Stemmann, L., Eloire, D., Sciandra, A., Jackson, G. A., Guidi, L., Picheral, M., & Gorsky, G. (2008). Volume distribution for particles between 3.5 to 2000 μm in the upper 200 m region of the South Pacific gyre. *Biogeosciences*, *5*, 299–310. https://doi.org/10.5194/bg-5-299-2008

Stemmann, L., Jackson, G. A., & Ianson, D. (2004). A vertical model of particle size distributions and fluxes in the midwater column that includes biological and physical processes — Part I model formulation. *Deep Sea Research Part I: Oceanographic Research Papers*, *51*, 865–884.

Stramski, D., Bricaud, A., & Morel, A. (2001). Modeling the inherent optical properties of the ocean based on the detailed composition of planktonic community. *Applied Optics*, *40*(18), 2929–2945. https://doi.org/10.1364/ao.40.002929

Stramski, D., & Kiefer, D. A. (1991). Light scattering by microorganisms in the open ocean. *Progress in Oceanography*, *28*, 343–383. https://doi.org/10.1016/0079-6611(91)90032-H

Twardowski, M. S., Boss, E., Macdonald, J. B., Pegau, W. S., Barnard, A. H., & Zaneveld, J. R. V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *Journal of Geophysical Research*, *106*, 14,129–14,142.

van der Linde, D.W. (1998). Protocol for the determination of total suspended matter in oceans and coastal zones (technical note I.98.182). Ispra, Italy: Joint Research Centre.

Ward, B. A., Dutkiewicz, S., Jahn, O., & Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnology and Oceanography*, *57*, 1877–1897. https://doi.org/10.4319/lo.2012.57.6.1877

Woodward, G., Ebenman, B., Emmerson, M., Montoya, J. M., Olesen, J. M., Valido, A., & Warren, P. H. (2005). Body size in ecological networks. *Trends in Ecology & Evolution*, *20*(7), 402–409. https://doi.org/10.1016/j.tree.2005.04.005

Woźniak, S. B., Stramski, D., Stramska, M., Reynolds, R. A., Wright, V. M., Miksic, E. Y., et al. (2010). Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. *Journal of Geophysical Research*, *115*, C08027. https://doi.org/10.1029/2009JC005554

Xi, H., Larouche, P., Tang, S., & Michel, C. (2014). Characterization and variability of particle size distributions in Hudson Bay, Canada. *Journal of Geophysical Research: Oceans*, *119*, 3392–3406. https://doi.org/10.1002/2013JC009542

CHAPTER 1 ACKNOWLEDGEMENTS

Chapter 1, in full, is a reprint of material as it appears in *Journal of Geophysical Research: Oceans*, Runyan, H.; Reynolds, R.A.; Stramski, D.; 2020. The dissertation author was the primary investigator and author of this paper.

CHAPTER 2

A DECADE OF UNDERWATER CORAL REEF PHOTOGRAMMETRY: A DISCUSSION OF LESSONS LEARNED, BEST PRACTICES, AND COMMUNICATION STANDARDS

Hugh Runyan, Clinton B. Edwards, Nicole Pedersen, Vid Petrovic, Falko Kuester, and Stuart A. Sandin

ABSTRACT

Newly available digital imaging and image processing technology, and the associated field methodology, offer a means to increase observational capacity and increase return on the investment of limited research resources. A small field team can sometimes leverage these developments to collect substantially more data than was previously possible. Along these lines, the Sandin and Kuester labs have been experimenting with image surveys of large tracts of coral reefs for more than a decade. In this report, these collaborative research groups discuss the reasoning behind the image survey procedures they use, with the hope of contributing to understanding of these methods in the coral research community. Additionally, an evaluation of the resulting image data products is presented, along with recommendations for reporting standards that facilitate transparent communication between research groups.

2.1 INTRODUCTION

Making measurements underwater with sufficient accuracy and replication to answer questions of ecological interest has been a challenge for coral researchers since the time of Darwin. Increasing personnel is one solution, but beyond the resource capacity of most research groups. Field deployments are resource intensive, requiring training and expertise of personnel as well as funds for travel and operations. A means to increase observational capacity and maximize field investments is offered by technology; it is possible for a small field team to leverage newly available sensors to collect much more data than was previously possible. In coral reef science, advances in digital photography and computational efficiency have made it possible to conduct image-based surveys of large areas of reef, enabling the measuring and tracking of thousands of individual coral colonies (Weinberg, 1981; Gracias and Santos-Victor, 2000; Lirman et al., 2007). With the photogrammetric method structure from motion (SFM), those images can be used to estimate an accurate 3D model of the benthic community, enabling geometrically consistent and precise analysis of the entire reef area back in the lab (Pizarro et al., 2009; Smith et al., 2016; Edwards et al., 2017; Ferrari et al., 2017; Kodera et al., 2020; Sandin et al., 2020). As coral grows mere centimeters or millimeters a year, accurate measurements are particularly important. Repeated and co-registered image-based sampling can provide unique data streams for quantifying this change through time. These digital twins can be archived, shared, and used as a basis for collaboration. The research groups contributing to this report have surveyed nearly 2000 reef sites around the world in the past decade, building an image-based repository of these reefs. In doing so, we have experimented with a variety of approaches, eventually settling on the best practices for efficient data collection and accurate and precise descriptions of benthic community structure. We present the approaches with associated rationale in this report.

2.1.1 BIOLOGICAL VALUE OF LARGE-AREA IMAGE SURVEYS

In terrestrial biology, researchers have taken advantage of standardized large-area sampling of individual organisms, particularly in the context of long-term study sites, including Barro Colorado Island (Hubbell and Foster, 1992), the Hubbard Brook experimental forest , and others (Condit et al., 2000). Among many other applications, these data have been used to identify or evaluate fundamental ecological dynamics, such as dispersion patterns (Hubbell, 1979; Lieberman et al., 1985; Condit et al., 2000), structuring mechanisms e.g. recruitment patterns, habitat preference and availability, dispersal probabilities, resource limitation (Hubbell, 1979; Connell, 1985; Turner, 1989; Condit et al., 2000; Rietkerk and van de Koppel, 2008), and space use (Harms et al., 2000; Marhaver et al., 2013).

However, to date, most spatial studies of coral communities focusing on site (Goreau, 1959; Kenyon et al., 2010), island (Newman et al., 2006; Sandin et al., 2008), or region (Smith et al., 2016), have relied on percent cover data. A complete understanding of population-level demography as well as community-level succession requires data collected at a finer scale. A handful of studies have accomplished this with laborious in-water monitoring protocols (Lewis, 1970; Stimson, 1974; Bradbury and Young, 1981; Bak et al., 1982; Hughes, 1984; Carlon and Olson, 1993; Bak and Nieuwland, 1995; Connell et al., 1997; Fong and Glynn, 1998; Hughes and Tanner, 2000; Jolles et al., 2002; Karlson et al., 2007; Zvuloni et al., 2009; Vardi et al., 2012; Deignan and Pawlik, 2015; Doropoulos et al., 2015; Edmunds, 2015), but their taxonomic, spatial, and chronological comprehensiveness was necessarily limited by the demanding methodology.

2.1.2 AN INTRODUCTION TO PHOTOGRAMMETRY

Photogrammetry—the computation of the geometric structure of the contents of a photograph—has been pursued since the invention of the camera. For measurements of flat scenes taken from a known perspective, measurements between objects are straightforward; in scenes with unknown and varying depth, additional geometric consideration must be made. The basic principle for conducting measurement on perspective imagery is triangulation: roughly, an object's location in 3D space has three unknown variables and thus requires three measurements (i.e., photographs) to be fully determined. The problem can also be formulated as one of the object's size, distance, and angular relationship to the center of the camera's sensor. With a single image from a known location with known lens geometry, it is possible to calculate the angle between objects in an image and the camera that took it. However, range and scale remain unknown—it is not generally possible to determine if an object is small or merely far away. These remaining variables become solvable with multiple views of the same object.

The task of photogrammetry requires estimating the relative locations and orientations of the camera for each image. Intersecting rays projected from the multiple image/camera orientations to points visible in multiple images can then be used to calculate their location relative to the cameras. These perspective geometric equations are relatively straightforward, but solving them requires a lot of calculation that must be repeated for every point one desires to locate. To create a cm-resolution 3D model of a single square meter of reef surface, these tasks would need to be repeated 100 x 100 = 10,000 times (assuming no two points are directly above or below each other). Only in the last few decades have computers become fast enough, and digital cameras cheap and sharp enough, for this process to be practical for biological researchers. The specific technique of using a sequence of overlapping images to infer 3D structure is referred to as structure from motion (SFM). It has become common in underwater fields like archeology and biology, and has

been demonstrated to be reliably capable of producing accuracies of approximately 1 cm (often better) with proper procedure (Figueira et al., 2015; Rossi et al., 2020).

2.1.3 A DISCUSSION OF IMAGE COLLECTION CONSIDERATIONS

The quality of a photographic survey, and therefore the resulting photogrammetric 3D reconstruction, is determined by 1) the quality of the imagery, and 2) the extent to which it comprehensively covers the target area. An example of a survey image is shown in Figure 2.1. The 3D model cannot contain any information that was not captured in the underlying imagery. The biological question being pursued—specifically, what measurements are required to answer it—determine what resolution and coverage is necessary. Planning an image survey requires balancing the desired measurements against available resources.

2.1.3.1 IMAGE QUALITY

Image quality (i.e., the level of detail captured) is influenced by a multitude of factors. Camera sensor resolution and the clarity of the lens and waterproof housing limit the highest possible level of detail. The term ground sampling distance (GSD) describes the linear distance on land corresponding to a single camera pixel at some range, orientation, focal length, sensor resolution, viewing angle, etc.—the detail of an image is physically limited to the number of sensor pixels divided by the area of target surface the image covers. GSD and resolution are functionally equivalent for the purposes of this discussion, but are inversely related: GSD decreases and resolution increases as a sensor is moved closer to its target and vice versa. As resolution increases, the distance along the benthos covered by a single pixel shrinks. Resolution is influenced by focal length (informally referred to as zoom): a longer focal length produces an image that is zoomed

in, meaning the sensor pixels are spread across a smaller area. Small details that are blurry at a short focal length, because they comprise only a few pixels, will become much clearer at longer focal lengths where many more pixels are devoted to the same area.

It is important to briefly distinguish between sharpness and resolution. As previously mentioned, resolution refers to the physical relationship between the size of the target area and the number of sensor pixels. However, a high-resolution image can still be blurry (not sharp) if, for example, the lens is out of focus. A perfectly sharp image maximally utilizes the available sensor pixels. Underwater, sharpness is limited by water clarity—interactions between suspended particles in the water (and the water particles themselves) and the light traveling to the camera sensor.

2.1.3.2 CAMERA SETTINGS AND CHARACTERISTICS

The underwater photography environment is challenging: light levels are low, so, in order to keep electronic noise low (resulting from a low gain/ISO value), it would be desirable to use a relatively large aperture combined with slow shutter speeds. Large apertures let more light in, and slow shutter speeds allow more time to collect light. Unfortunately, for reef survey applications, shutter speeds must be kept high (<1/250 s) to ensure sharp images while the diver is moving. Additionally, large apertures introduce a focus problem known as depth of field, which is defined as the range of depths—distances from the lens—that are in focus in an image. Larger apertures result in narrower depth of field, meaning less of the image is in focus if the image contains objects that are varying distances from the camera. At 1.5 m from the benthos, a large aperture like f/2 that collects a lot of light will, when combined with a wide-angle 24 mm focal length lens, produce an image with a depth of field of only 0.47 m; 0.2 m in front of the focal plane (where the camera

is focused) to 0.27 m behind it. In this configuration, the bottom of any coral with more vertical relief than 0.27 m would be blurry, as well as anything more than 0.2 m taller. Depth of field expands as apertures get smaller: at 24mm focal length f/4 gives 1.02 m, f/5.6 gives 1.61 m, and f/8 gives 2.96 m.

A lot of light or high electronic noise levels are required to take images with a high shutter speed and a small aperture, because that combination captures little of the available light. To mitigate these shortcomings, we use the largest aperture we can get away with at a given reef site and make sure to use cameras with large sensors: large sensors have a high number of pixels while also using larger pixels, which collect more light and therefore have less electronic noise. Noise also scales with resolution; with more pixels, noise in pixel values occupy smaller areas of the image.

In most situations, image quality will be better with larger sensors found on DSLR-style and mirrorless cameras (35 mm wide), as opposed to smaller, cheaper, more mobile solutions like GoPros (6.16 mm wide). Larger sensors generally have more pixels, but even if both cameras have the same number of pixels, the larger pixels of a larger sensor have a better signal to noise ratio, and their larger size mitigates physical interference between light entering adjacent pixels. However, smaller sensors do have one advantage: they produce an equivalent field of view at shorter focal lengths, and shorter focal lengths have wider depths of field than longer ones. This means that an equivalent depth of field is achieved with a larger aperture, which captures more light, allowing faster shutter speeds. In places with very low light or exceptional vertical reef structure, a smaller sensor/larger aperture approach allows surveyors to achieve a suitably large depth of field. Balancing all these settings is confusing and difficult to explain without extensive visuals. The two important points are 1) larger sensors tend to be better, and 2) if it becomes

impossible to maintain a large enough depth of field due to very low light, a good solution is to use a smaller sensor. We encourage all researchers that use cameras as scientific instruments to investigate these subjects, as there is a large volume of good explanatory information available (generally targeted to amateur and professional photographers), and to experiment. Choosing the right camera equipment for a particular application and then optimally using it is the only way to capture as much information about the reef as is possible.

Finally, it is important to note that we use a grey card to white balance the image, because it is critical for color consistency between times of day, lighting conditions, water conditions, and locations. Color underwater is quite complicated, varying with time of day, cloud cover, depth from the surface and between the camera/target, and suspended particles in the water. Cameras designed for use above water do not always handle it very well, so it is important to calibrate against something of a known color. Grey cards are convenient because they can be used to white balance the camera in the water before an image survey. Color cards would provide even more accurate color, but require additional processing to implement, so we do not use them. Algorithmic approaches to correcting the color of underwater images also exist (Akkaynak and Treibitz, 2019).

2.1.3.3 IMAGE COVERAGE

Once individual image quality has been optimized through choice and operation of cameras, researchers must determine how to cover the targeted survey area. Common software for reconstructing 3D models from image surveys like Agisoft Metashape use the photogrammetric method structure from motion. SFM works best if there is > 90% overlap between adjacent images (Harwin et al. 2015). To insure >90% overlap everywhere without leaving holes, we plan a safety factor. We have found that swimming a grid pattern, as shown in Figure 2.2, followed by a second

redundant perpendicular gridded pass, best insures adequate overlap everywhere within the study site. An illustration of progressively-increasingly image overlap is illustrated in Figure 2.3 by darker shades of yellow.

With the constraints of SCUBA bottom time at 10 m depth, we have settled on 10 m x 10 m sites as the largest plot size reliably manageable while still achieving desired image resolution, with 1 m between grid passes and the camera on a 1 image/s interval timer. At 1.5 m above the bottom, a 24mm focal length lens (with an angle of view that is 74 degrees horizontally and 53 degrees vertically) will be able to see a rectangle of reef that is ~2.24 m wide and ~1.5 m high. High and wide here are relative to a camera held in its normal orientation facing the horizon—facing down at a reef, the high direction is parallel to the direction the diver is swimming and the wide direction is perpendicular. A diver swimming 0.2 m/s will cover 0.2 meters between image frames. This results in ~7 images seeing any one spot on the reef as the diver passes over. If the diver swims intervals of 1 m between imaging passes, there is also side-to-side overlap, so most points are captured on three sequential passes as well, resulting in a total of 21 images capturing the same spot. We swim a second perpendicular grid, bringing the average number of images to 42, which is ~four times the minimum number—10—corresponding to 90% overlap. This level of redundancy ensures the full plot area is adequately surveyed even if the second set of passes cannot be completed.

Figure 2.4 illustrates the impact of adequate overlap coverage. It shows a small 3D model of a few corals built from a few hundred images. The first model was calculated with the entire set of images. Each sequential model after it was calculated with a shrinking, randomly-selected fraction of the images: 1/2, 1/3rd, 1/5th, and then 1/10th. As number of images decreased, holes

began appearing in the model, until it is almost entirely unreconstructable from the few available views.

2.1.3.4 ALTERNATIVES TO DIVER IMAGING

In situ image surveys capture a lot of information, but they do require multiple divers to actually be at the site and manually swim a survey pattern with a camera. Other technologies do not have this disadvantage: airplanes can capture images of an entire island in a day, satellites capture images of the entire world on a time scale of hours to weeks depending on the orbit and sensor, camera-wielding robots can stay underwater much longer than humans, and ship-based sonar can survey entire islands in hours to days.

Satellite data products have many uses in the natural sciences, but their ground sampling distance—on the order of meters at the smallest—is much too large to measure coral growth that amounts to only millimeters or centimeters per year. An additional shortcoming of imaging from above the ocean surface is the depth and clarity of the water over the targeted benthos. Aerial photography (from a manned or unmanned aerial vehicle (UAV)) can get much closer to the reef than satellites and so has correspondingly smaller GSD (i.e., better resolution), but is similarly depth-limited by water clarity, and is still orders of magnitude further away from the reef than a diver. In shallow areas, UAV photography can be sufficient if distortion from surface waves is removed (Chirayath and Instrella, 2016; Chirayath Ved and Earle Sylvia A., 2016; Chirayath, 2017).

It is possible to conduct underwater surveys with robots (automated underwater vehicles — AUVs) instead of humans (Noguchi et al., 2022), but expense and imprecise underwater navigation/obstacle avoidance currently render that approach impractical for most research

scenarios, especially as it generally requires substantial ship and personnel resources to deploy the robot, monitor its piloted or automated progress, and then retrieve it. These systems cannot be fired with cruise missiles from urban universities to faraway atolls; they must be attended by well-equipped research vessels.

Imaging platforms towed from boats enable *in situ* imaging at much faster rates over much larger areas than a diver limited by SCUBA equipment and human physiology. However, it is risky and difficult to use towed platforms within 3-5+ m above the benthos, so resolution is much lower than what is achievable with a diver.

All of these platforms offer interesting and valuable avenues by which to pursue ocean science. However, they are not currently of matching the efficiency and capability of human divers with handheld cameras.

2.1.3.5 IMAGING WITH SOUND INSTEAD OF LIGHT

Acoustic SONAR (an acronym for sound navigation and ranging, meaning imaging with sound waves instead of light, often just written as sonar) has one significant advantage over optical cameras: sound can propagate many times further than light underwater. However, that is only true of low frequencies. While ultrasound imaging devices like those used in hospitals can approach the resolution of cameras, they are still not as sharp, cost much more, and have a useful range of only a few meters due to the rapid attenuation of the high frequencies necessary to image small features. Acoustic imaging systems that do not attenuate beyond detectability in the distance from a ship to the reef and back again use sound wavelengths that are too long for high-resolution imagery comparable to a camera. Acoustic imaging also cannot detect color. Due to these reasons, acoustic imaging is not a suitable replacement for diver-based imaging procedures.

2.2 FIELD PROTOCOL

What follows is our current field protocol methodology and a discussion of the significance of its components. The objective of this image survey procedure is to capture an entire 10 m x 10 m patch of reef with sufficient overlap (10x+) for accurate photogrammetric reconstruction, as well as the information required to scale it and orient it to the surface accurately. Here, we discuss, study design and sampling approach.

2.2.1 STUDY DESIGN

We start by choosing target islands or regions. If we are new to an area, we generally select a specific number of approximately evenly distributed target areas along the forereef of the island/region (e.g., 12 per island), and then select the exact sites randomly on arrival for the dive. We generally seek out sites 10 m deep on forereef away from passes, though some studies call for sites at a range of depths or in lagoons or passes. Operating at 10 m depth allows the 60 minutes required to survey a site without complicated and dangerous decompression protocols. Additionally, we find 10 m to be a representative depth for studying reefs, and it is deep enough for us to cover large areas without entering water too shallow for diving, especially in the presence of large waves.

If we are returning to a site, we exactly relocate it using GPS coordinates, left-behind steel stakes, and a printed out 2D orthographic projection (orthoprojection) of the site. Exactly relocating sites is necessary to study change through time.

2.2.2 SAMPLING APPROACH

All image surveys are done with a team of at least two divers. A third may be useful in difficult conditions, such as strong current. At the beginning of the dive, one diver places one heavy corner tile, measures 10 m offshore and places tiles 5 m and 10 m in that direction. The diver then returns to the first corner and measures 10 m alongshore to place the third corner tile, and then from there measures 10 m offshore, placing a tile at the 5 m midpoint and then the final fourth corner. This process is difficult to explain, but it is clear when depicted visually as in Figure 2.5. The exact order of operations is not critical, but we find this version to be the fastest and most precise.

The other diver hammers a steel stake into the reef outside of each of the two midpoint tiles and places four scale bars (used to calibrate the scale of the 3D model) inside the box created by the four corners, and then records the depth of the tiles on the scale bars and each of the six corners/edges. It is critical that all tiles and scale bars do not move during the image survey—if they do, they will not be reconstructed properly in the 3D model.

This diver also places 1.5 m high reference floats ~1 m outboard of each of the four corners (depicted in Figure 2.2b), which give the divers a visual reference when swimming the image survey grid. Finally, the second diver attaches a surface float to the benthos outside of the plot and ascends to the surface to collect a GPS location with a waterproof device stored on the float.

One of the divers then slowly swims a grid over the site with 1 m horizontal distance between each pass with the camera on a 1 image/s interval timer. Beginning at the offshore edge, the diver swims the alongshore passes first and then, once the entire plot has been photographed, the grid pattern is repeated in the on-/offshore direction. The survey pattern is shown in Figure 2a. Starting offshore and moving onshore during the fast pass has the diver moving from deeper parts of the plot to the shallow areas as the survey progresses. Deeper areas consume more oxygen and

carry additional risks to divers in the case of emergency, so we advise surveying these areas first. Our pass spacing and swim speed ensure that, if time is not available for the second orthogonal grid, enough image overlap for an accurate 3D model is still available.

We use a rectangular plastic housing with two full frame DSLR cameras; one with a wider-angle lens (24mm) to ensure adequate coverage and substantial overlap between images, and the other with a longer focal length lens (usually 85 mm) to capture a magnified, more detailed view. We load the images from the survey onto a computer, and then reconstruct a 3D model of the site with SFM software (generally Agisoft Metashape, see (Burns and Delparte, 2017) for a comparison of software options). The result is a 3D pointcloud or a surface mesh. We prefer to work with pointclouds because they do not require algorithmic assumptions determining which points belong to the reef surface and which do not.

Finally, we use the 3D pointcloud software Viscore (Petrovic et al., 2014) to scale the model, compute the direction of gravity with the depth values recorded at the tiles, align it with previous models from the same location if we have any, and export an orthoprojection (in the direction of gravity, not normal to the reef surface) of the site. These operations are also possible in other 3D pointcloud software packages such as Open3D.

2.3 REPRESENTATIVE RESULTS

Figure 2.1 shows sample images from image surveys. Figure 2.4 shows a 3D model of a single coral, with the a), b), c), d), e) panels respectively showing the degradation of 3D reconstruction that results from successively less adequate survey coverage. This figure illustrates why we emphasize the necessity of heavily redundant survey imagery. Figures 2.6 and 2.7 show orthoprojections and the analyses they enable: Figure 2.6 shows the traced boundaries of a single

coral in successive years, while Figure 2.7 shows an untraced complete 2D orthoprojection of a 10 m x 10 m field site, and the same site after most organisms have been traced by coral experts. Figure 2.7 shows similar coral tracing, but on a 3D model.

2.4 DISCUSSION

2.4.1 ANALYSIS MEDIUM

A balance between study goals and available resources determines procedural approaches, including how to analyze the resulting image-derived data products—2D survey images, 3D models, or 2D orthoprojections thereof. Once a high-quality 3D model of the site is made, researchers must decide whether to analyze it in 3D, or to make some 2D projection of it, either by translating the images to a common orientation and stitching them together, or by capturing a projected view of the 3D model. We find that a 2D orthographic projection (orthoprojection) is generally most advantageous. The reasons for this are:

1) Our image surveys are conducted from the top down, so most 3D information that is discarded when translating from 3D to 2D (the vertical sides and undersides) is not well-captured to begin with. A 2D orthoprojection is approximately analogous to what the camera was able to see.
2) Geometry is much simpler in 2D. 2D surface areas and distances are easy to calculate and do not require determining where in the 3D pointcloud the surface of the reef is. This makes 2D analysis much mathematically simpler, much less computationally demanding, and less prone to subjective (sometimes incorrect) algorithmic assumptions. 2D files are much smaller than full 3D pointclouds and therefore much easier to store, and faster and easier to work with and share. Importantly, while image pixels form a regular and ordered grid,

pointclouds are irregular and unordered: points are not a consistent distance from one another, nor is it generally possible to use their listed order to determine which are nearest each other.

3) Studies have shown 2D surface area correlates well with 3D surface area and volume (House et al., 2018), though it varies by genera/growth form and definition of 3D surface area.

When capturing a 2D projection of a 3D model, a projection angle must be chosen. The two projection angles most likely to be useful are: 1) normal to gravity (straight down from the surface), and 2) normal to the surface of the reef. We have found that the orientation of gravity is preferable, because it is consistent, light-seeking corals tend to orient themselves to the surface rather than the plane of the bottom to which they're attached, and topographically complex reefs produce normal vectors that are difficult to calculate and sometimes intuitively strange. Moving a selected site only a handful of meters one way or another near large features (e.g., spur and groove formations) can significantly alter the resulting average surface normal, which in turn alters projected surface area of corals.

An alternative means to create a 2D view of a 3D site is directly stitching adjacent images together. However, stitching many images together with mm precision is challenging and error prone. The inherent distortion in each image caused by the lens elements must be determined (it is different for every lens), perfectly corrected for in each image, and then their exact relative positions must be calculated so they can be transformed to a common perspective and then pieced together. The 3D model derivation process entails calculating many of the variables necessary to perform this task (camera location and orientation for each image), but has additional advantages,

such as consulting numerous views of an area before collapsing them to a single structure. Stitching the images together rather than relying on the 3D model as the best available geometric understanding of the scene often results in imperfections: boundaries between images that show misalignments across them or features that appear on both sides or neither.

Once a 2D projection is created, it can be traced, point subsampled, or otherwise analyzed by a variety of means. Researchers can use simple software like PhotoGrid or Photoshop, or experiment with neural-network-assisted tools like CoralNet (Beijbom, 2015) or Taglab (Pavoni et al., 2020). Figure 2.6 shows the growth of a single coral in successive years as depicted by a 2D orthorpojection, while Figure 2.7 shows an example of an entirely traced orthoprojection. Figure 2.8 shows a similarly-traced *Porites* (a coral genus) in a 3D pointcloud model.

2.4.2 REPORTING STANDARDS

Large-area imagery data products enable digital collaboration and database pooling amongst coral researchers around the world. They also enable dialogue, through published research, between disparate researchers working from hypothetically similar data products. However, variability in e.g. survey procedures, camera equipment and settings, and color correction can make it difficult to determine how comparable the data products actually are. We propose the following reporting standards to facilitate communication about and interpretability of large area imagery data products:

1) Survey procedure
2) Camera equipment and settings
3) Height above the benthos at which the imagery was collected and resulting ground sample distance (resolution)

4) Sample images from the survey and of the resulting 2D/3D data products
5) 3D reconstruction software

2.5 CONCLUSIONS

Large area imagery products are a significant step forward in the ability of researchers to access accurate quantitative measurements of the coral reefs they study. If the necessary resources can be marshalled to conduct image surveys of suitable scale and resolution, they offer the ability to, in a single dive, capture entire benthic ecosystems to study for years into the future. Collecting such data enables precise analysis of change in these ecosystems over time, as well as an invaluable means to archive coral reefs as they change for both research and public conservation outreach. Our labs have successfully used these methods to collect the data underlying analyses in numerous publications that have contributed to scientific understanding of coral reefs (citations). We hope this procedural discussion will widen use of and expediate communicate about large area imaging methods and data products. As technology continues to develop in the future, we look forward to participating in a continually-evolving dialogue on methods to collect measurements for studying coral reefs.

FIGURES

Figure 2.1: Example of survey image.

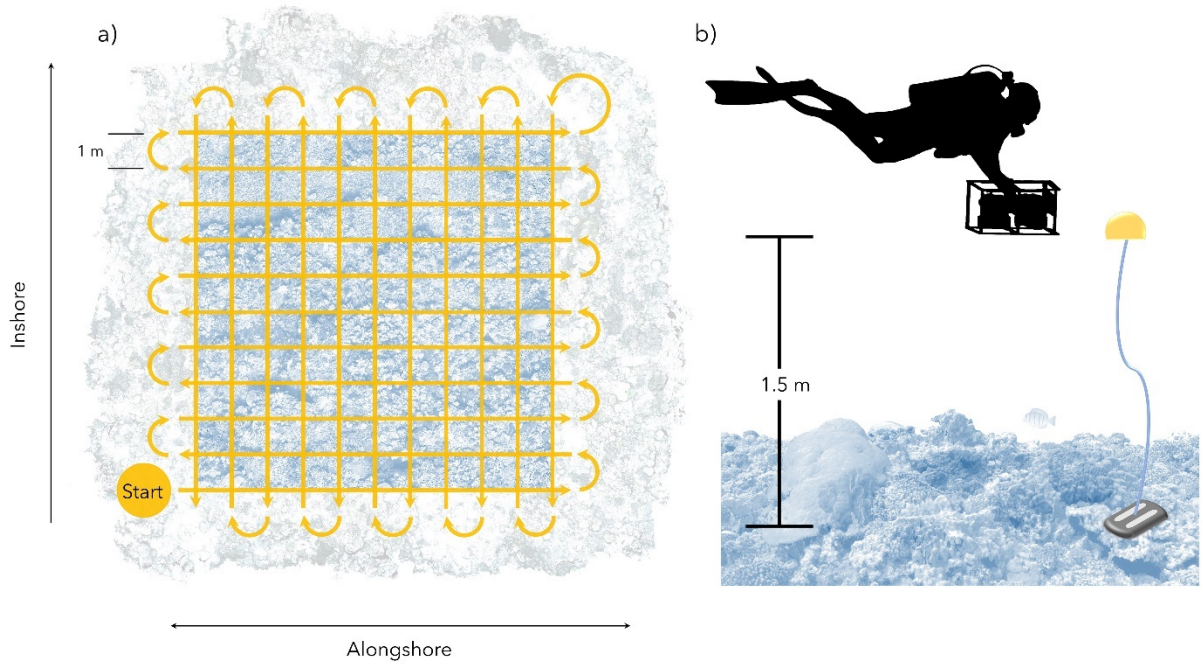


Figure 2.2: a) Swim pattern for image survey, b) Arrangement of diver, cameras, and height markers during image survey.

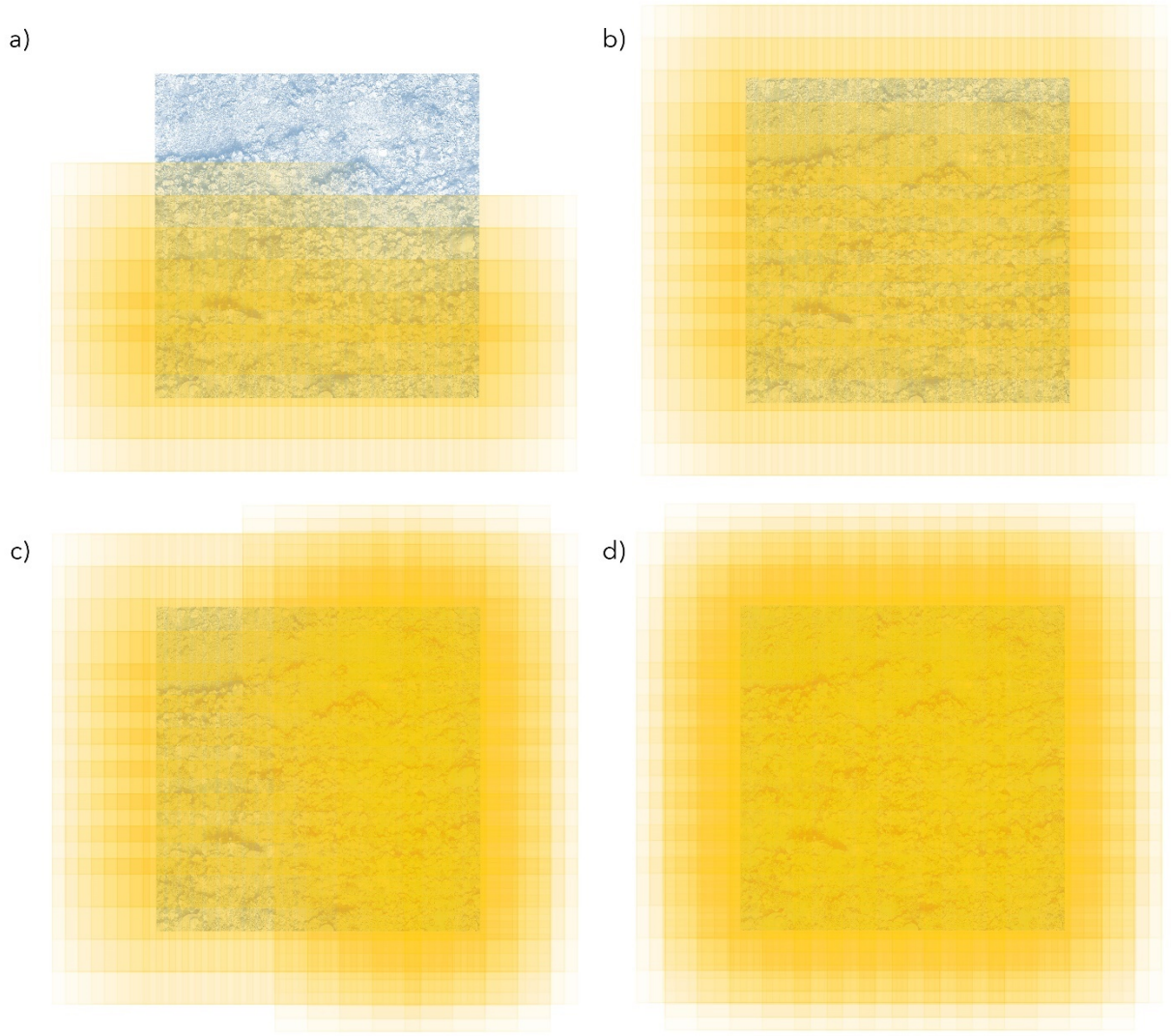
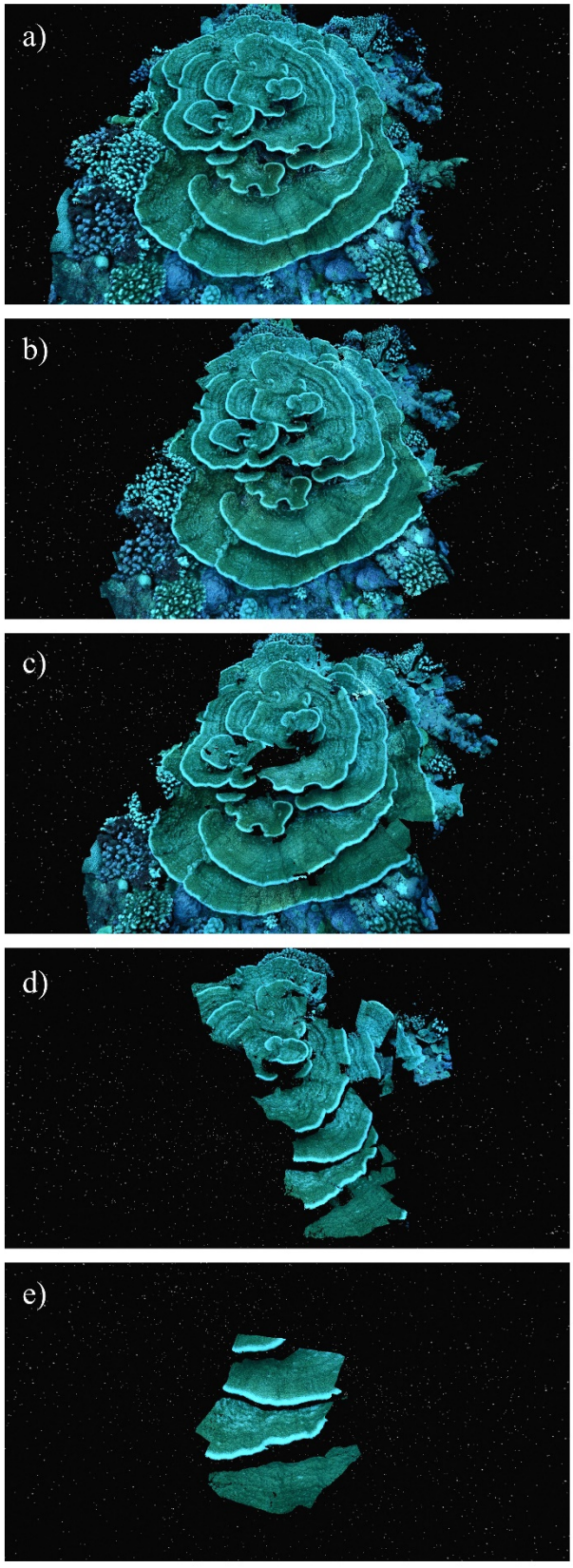


Figure 2.3: Sequential image overlap at four stages during image survey: after a) half of the first set of passes is completed, b) the first is finished, c) the second perpendicular set of passes is completed, and d) the second is finished.

Figure 2.4: The degradation of a 3D model as the number of overlapping images is reduced. Panel a) shows the full model built with all images, while the model in b) was derived from a randomly selected ½ of them, c) from 1/3, d) from 1/5, and e from 1/10th


a)
b)
c)
d)
e)

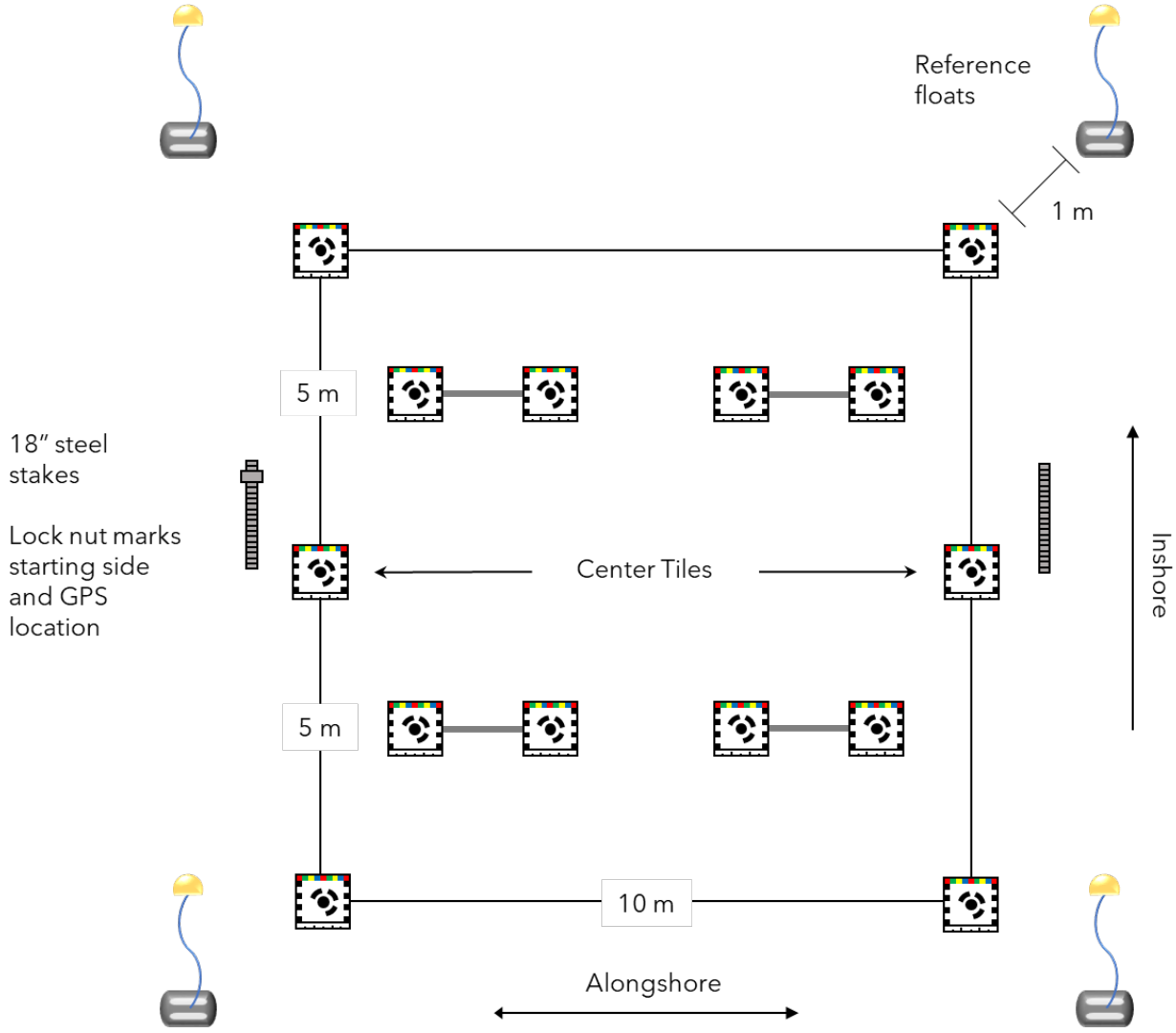


Figure 2.5: Equipment arrangement at field site

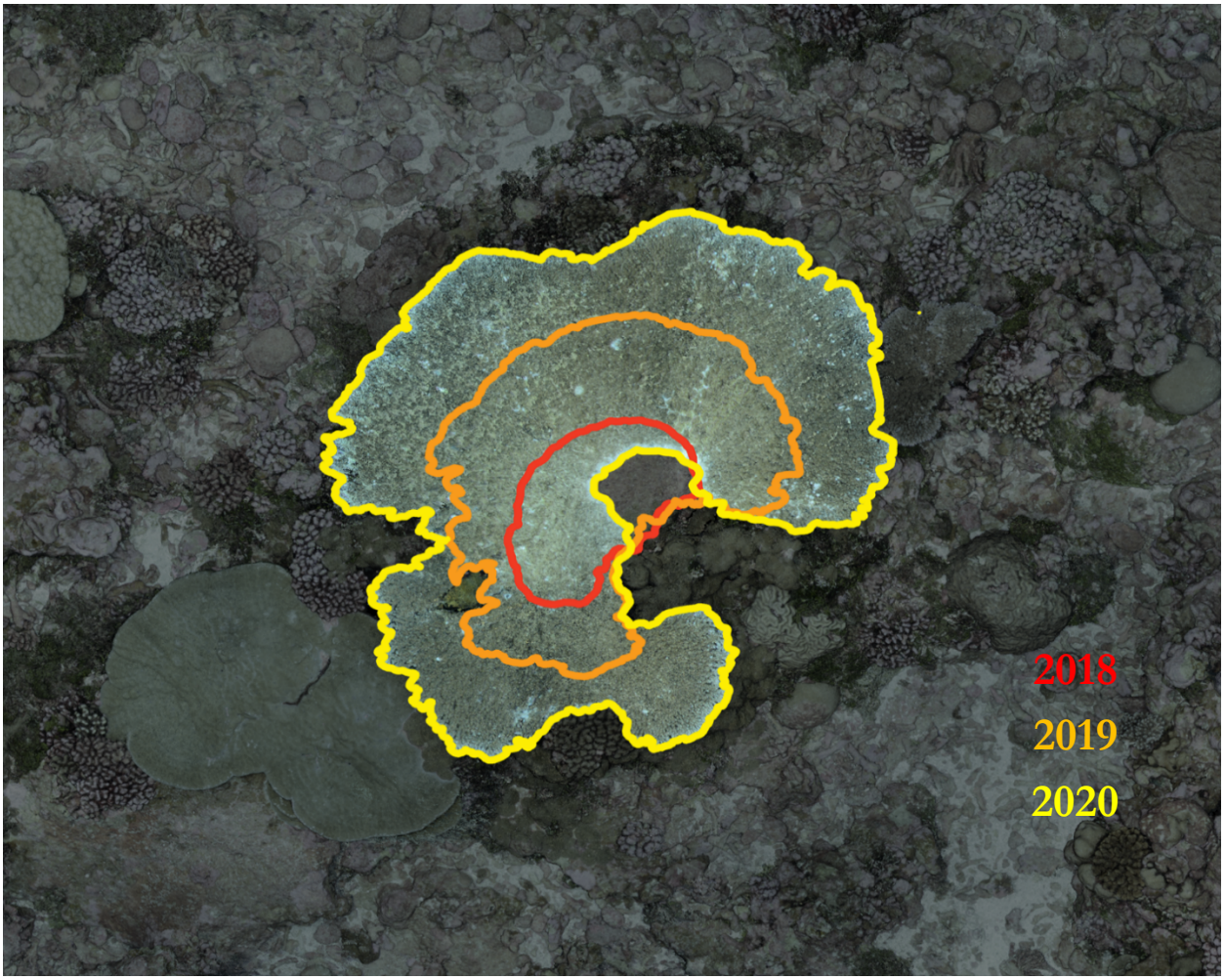
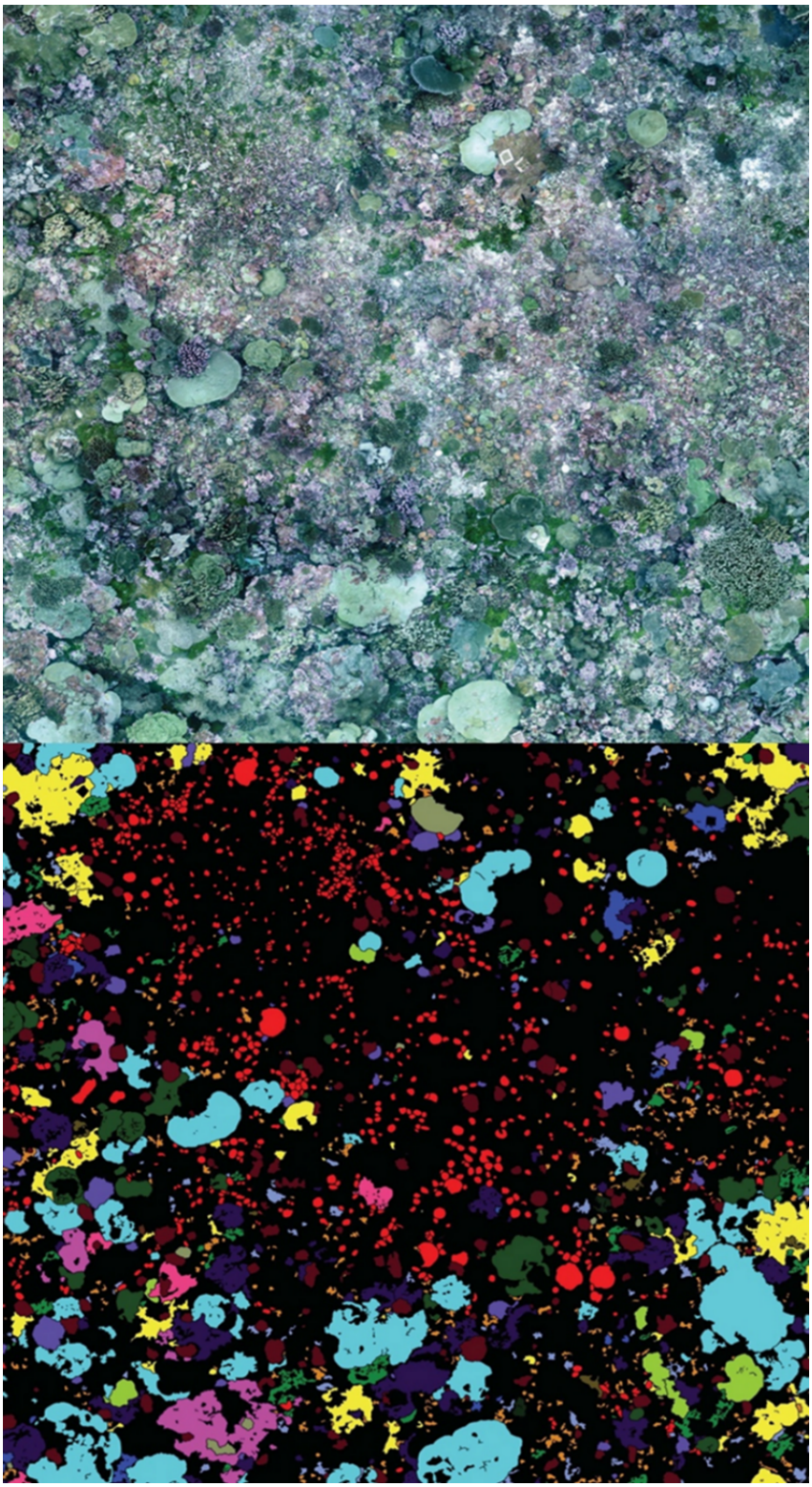


Figure 2.6: Growth of a single coral in successive years. Red is 2018, orange is 2019, yellow is 2020.

Figure 2.7: An orthoprojection before and after being traced.



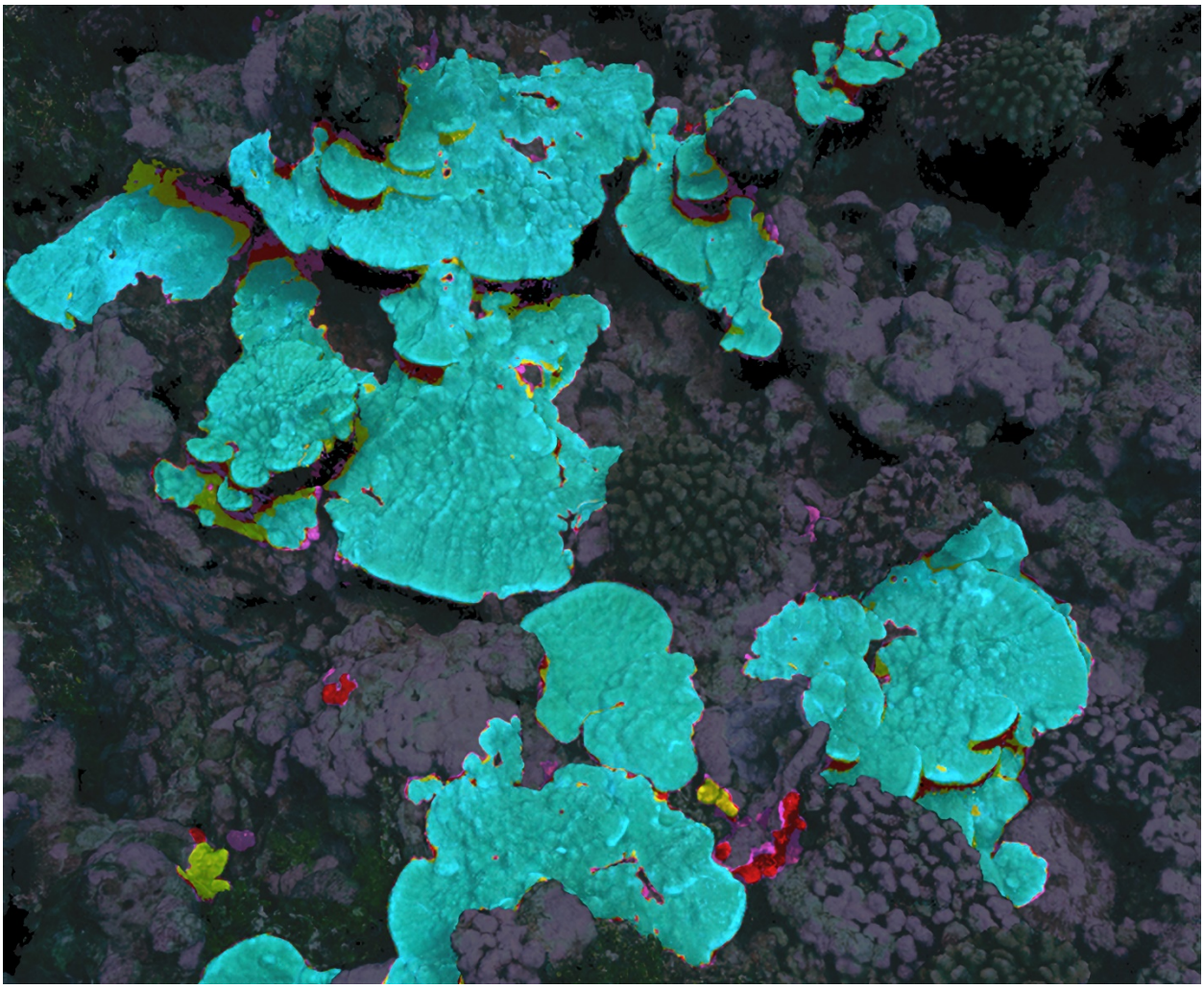


Figure 2.8: An example of an image-survey-derived 3D model. Colors represent tracing of the genus *Porites* by multiple coral experts—blue is agreement amongst all tracers, and other colors show areas experts disagreed.

REFERENCES

Akkaynak, D., and Treibitz, T. (2019). Sea-Thru: A Method for Removing Water From Underwater Images. in, 1682–1691. Available at: https://openaccess.thecvf.com/content_CVPR_2019/html/Akkaynak_Sea-Thru_A_Method_for_Removing_Water_From_Underwater_Images_CVPR_2019_paper.html [Accessed November 15, 2022].

Bak, R. P. M., and Nieuwland, G. (1995). Long-Term Change in Coral Communities Along Depth Gradients Over Leeward Reefs in the Netherlands Antilles. *Bulletin of Marine Science* 56, 609–619.

Bak, R. P. M., Termaat, R. M., and Dekker, R. (1982). Complexity of coral interactions: Influence of time, location of interaction and epifauna. *Mar. Biol.* 69, 215–222. doi: 10.1007/BF00396901.

Beijbom, O. (2015). Automated Annotation of Coral Reef Survey Images. Available at: https://escholarship.org/uc/item/0rd0r3wd [Accessed April 13, 2018].

Bradbury, R., and Young, P. (1981). The Effects of a Major Forcing Function, Wave Energy, on a Coral Reef Ecosystem. *Marine Ecology-progress Series - MAR ECOL-PROGR SER* 5, 229–241. doi: 10.3354/meps005229.

Burns, J. H. R., and Delparte, D. (2017). Comparison of commercial structure-from-motion photogrammetry software used for underwater three-dimensional modeling of coral reef environments. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLII-2/W3, 127–131. doi: 10.5194/isprs-archives-XLII-2-W3-127-2017.

Carlon, D. B., and Olson, R. R. (1993). Larval dispersal distance as an explanation for adult spatial pattern in two Caribbean reef corals. *Journal of Experimental Marine Biology and Ecology* 173, 247–263. doi: 10.1016/0022-0981(93)90056-T.

Chirayath, V. (2017). Fluid Lensing and Applications to Remote Sensing of Aquatic Environments. in (8-12 May 2017, South Africa). Available at: https://ntrs.nasa.gov/search.jsp?R=20170004512.

Chirayath, V., and Instrella, R. (2016). FluidCam 1&2 - UAV-based Fluid Lensing Instruments for High-Resolution 3D Subaqueous Imaging and Automated Remote Biosphere Assessment of Reef Ecosystems. in Available at: http://adsabs.harvard.edu/abs/2016AGUOSIS53A..01C.

Chirayath Ved and Earle Sylvia A. (2016). Drones that see through waves – preliminary results from airborne fluid lensing for centimetre-scale aquatic conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems* 26, 237–250. doi: 10.1002/aqc.2654.

Condit, R., Ashton, P. S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbel, S. P., Foster, R.B., Itoh, A., Lafrankie, J. V., Lee, H. S., Losos, E., Manokaran, N., Sukumar, R., Yamakura, T. (2000). Spatial Patterns in the Distribution of Tropical Tree Species. *Science* 288, 1414–1418. doi: 10.1126/science.288.5470.1414.

Connell, J. H. (1985). The consequences of variation in initial settlement vs. post-settlement mortality in rocky intertidal communities. *Journal of Experimental Marine Biology and Ecology* 93, 11–45. doi: 10.1016/0022-0981(85)90146-7.

Connell, J. H., Hughes, T. P., and Wallace, C. C. (1997). A 30-Year Study of Coral Abundance, Recruitment, and Disturbance at Several Scales in Space and Time. *Ecological Monographs* 67, 461–488. doi: 10.1890/0012-9615(1997)067[0461:AYSOCA]2.0.CO;2.

Deignan, L. K., and Pawlik, J. R. (2015). Perilous proximity: Does the Janzen–Connell hypothesis explain the distribution of giant barrel sponges on a Florida coral reef? *Coral Reefs* 34, 561–567. doi: 10.1007/s00338-014-1255-x.

Doropoulos, C., Ward, S., Roff, G., González-Rivero, M., and Mumby, P. J. (2015). Linking Demographic Processes of Juvenile Corals to Benthic Recovery Trajectories in Two Common Reef Habitats. *PLOS ONE* 10, e0128535. doi: 10.1371/journal.pone.0128535.

Edmunds, P. J. (2015). A quarter-century demographic analysis of the Caribbean coral, Orbicella annularis, and projections of population size over the next century. *Limnology and Oceanography* 60, 840–855. doi: 10.1002/lno.10075.

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, 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., Atwood, T. B., Byrne, M. (2017). 3D photogrammetry quantifies growth and external erosion of individual coral colonies and skeletons. *Sci Rep* 7, 1–9. doi: 10.1038/s41598-017-16408-z.

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

Fong, P., and Glynn, P. W. (1998). A dynamic size-structured population model: does disturbance control size structure of a population of the massive coral Gardineroseris planulata in the Eastern Pacific? *Marine Biology* 130, 663–674. doi: 10.1007/s002270050288.

Goreau, T. F. (1959). The Ecology of Jamaican Coral Reefs I. Species Composition and Zonation. *Ecology* 40, 67–90. doi: 10.2307/1929924.

Gracias, N. R., and Santos-Victor, J. (2000). Underwater Video Mosaics as Visual Navigation Maps. *Computer Vision and Image Understanding* 79, 66–91. doi: 10.1006/cviu.2000.0848.

Harms, K. E., Wright, S. J., Calderón, O., Hernández, A., and Herre, E. A. (2000). Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* 404, 493–495. doi: 10.1038/35006630.

House, J. E., Brambilla, V., Bidaut, L. M., Christie, A. P., Pizarro, O., Madin, J. S., Dornelas, M. (2018). Moving to 3D: relationships between coral planar area, surface area and volume. *PeerJ* 6, e4280. doi: 10.7717/peerj.4280.

Hubbell, S. P. (1979). Tree Dispersion, Abundance, and Diversity in a Tropical Dry Forest. *Science* 203, 1299–1309. doi: 10.1126/science.203.4387.1299.

Hubbell, S. P., and Foster, R. B. (1992). Short-Term Dynamics of a Neotropical Forest: Why Ecological Research Matters to Tropical Conservation and Management. *Oikos* 63, 48–61. doi: 10.2307/3545515.

Hughes, T. P. (1984). Population Dynamics Based on Individual Size Rather than Age: A General Model with a Reef Coral Example. *The American Naturalist* 123, 778–795. doi: 10.1086/284239.

Hughes, T. P., and Tanner, J. E. (2000). Recruitment Failure, Life Histories, and Long-Term Decline of Caribbean Corals. *Ecology* 81, 2250–2263. doi: 10.2307/177112.

Jolles, A. E., Sullivan, P., Alker, A. P., and Harvell, C. D. (2002). Disease Transmission of Aspergillosis in Sea Fans: Inferring Process from Spatial Pattern. *Ecology* 83, 2373–2378. doi: 10.1890/0012-9658(2002)083[2373:DTOAIS]2.0.CO;2.

Karlson, R., Cornell, H., and Hughes, T. (2007). Aggregation influences coral species richness at multiple spatial scales. *Ecology* 88, 170–7. doi: 10.1890/0012-9658(2007)88[170:AICSRA]2.0.CO;2.

Kenyon, J. C., Maragos, J. E., and Cooper, S. (2010). Characterization of Coral Communities at Rose Atoll, American Samoa. Available at: http://repository.si.edu//handle/10088/11728 [Accessed November 12, 2019].

Kodera, S. M., Edwards, C. B., Petrovic, V., Pedersen, N. E., Eynaud, Y., and Sandin, S. A. (2020). Quantifying life history demographics of the scleractinian coral genus Pocillopora at Palmyra Atoll. *Coral Reefs* 39, 1091–1105. doi: 10.1007/s00338-020-01940-8.

Lewis, J. B. (1970). Spatial Distribution and Pattern of Some Atlantic Reef Corals. *Nature* 227, 1158–1159. doi: 10.1038/2271158a0.

Lieberman, D., Lieberman, M., Peralta, R., and Hartshorn, G. S. (1985). Mortality Patterns and Stand Turnover Rates in a Wet Tropical Forest in Costa Rica. *Journal of Ecology* 73, 915–924. doi: 10.2307/2260157.

Lirman, D., Gracias, N. R., Gintert, B. E., Gleason, A. C. R., Reid, R. P., Negahdaripour, S., Kramer, P. (2007). Development and application of a video-mosaic survey technology to document the status of coral reef communities. *Environ Monit Assess* 125, 59–73. doi: 10.1007/s10661-006-9239-0.

Marhaver, K. L., Vermeij, M. J. A., Rohwer, F., and Sandin, S. A. (2013). Janzen-Connell effects in a broadcast-spawning Caribbean coral: distance-dependent survival of larvae and settlers. *Ecology* 94, 146–160. doi: 10.1890/12-0985.1.

Newman, M. J. H., Paredes, G. A., Sala, E., and Jackson, J. B. C. (2006). Structure of Caribbean coral reef communities across a large gradient of fish biomass. *Ecology Letters* 9, 1216–1227. doi: 10.1111/j.1461-0248.2006.00976.x.

Noguchi, Y., Humblet, M., Furushima, Y., Ito, S., and Maki, T. (2022). Wide-Area Three-Dimensional Imaging of Mesophotic Coral Reefs Using a Low-Cost AUV. *Marine Technology Society Journal* 56, 74–89. doi: 10.4031/MTSJ.56.4.4.

Pavoni, G., Corsini, M., Callieri, M., Fiameni, G., Edwards, C., and Cignoni, P. (2020). On Improving the Training of Models for the Semantic Segmentation of Benthic Communities from Orthographic Imagery. *Remote Sensing* 12, 3106. doi: 10.3390/rs12183106.

Petrovic, V., Vanoni, D. J., Richter, A. M., Levy, T. E., and Kuester, F. (2014). Visualizing high resolution three-dimensional and two-dimensional data of cultural heritage sites. *Mediterranean Archaeology and Archaeometry* 14, 93–100.

Pizarro, O., Eustice, R. M., and Singh, H. (2009). Large Area 3-D Reconstructions From Underwater Optical Surveys. *IEEE Journal of Oceanic Engineering* 34, 150–169. doi: 10.1109/JOE.2009.2016071.

Rietkerk, M., and van de Koppel, J. (2008). Regular pattern formation in real ecosystems. *Trends in Ecology & Evolution* 23, 169–175. doi: 10.1016/j.tree.2007.10.013.

Rossi, P., Castagnetti, C., Capra, A., Brooks, A. J., and Mancini, F. (2020). Detecting change in coral reef 3D structure using underwater photogrammetry: critical issues and performance metrics. *Appl Geomat* 12, 3–17. doi: 10.1007/s12518-019-00263-w.

Sandin, S. A., Edwards, C. B., Pedersen, N. E., Petrovic, V., Pavoni, G., Alcantar, E., Chancellor, K. S., Fox, M. D., Stallings, B., Sullivan, C. J., Rotjan, R. D., Ponchio, F., Zglicynski, B. J. (2020). "Chapter Seven - Considering the rates of growth in two taxa of coral across Pacific islands," in *Advances in Marine Biology* Population Dynamics of the Reef Crisis., ed. B. M. Riegl (Academic Press), 167–191. doi: 10.1016/bs.amb.2020.08.006.

Sandin, S. A., Smith, J. E., DeMartini, E. E., Dinsdale, E. A., Donner, S. D., Friedlander, A. M., Konotchick, T., Malay, M., Maregos, J. E., Obura, D., Pantos, O., Pauley, G., Richie, M., Rohwer, F., Schroeder, R. E., Walsh, S., Jackson, J. B. C., Knowlton, N., Sala, E. (2008). Baselines and Degradation of Coral Reefs in the Northern Line Islands. *PLOS ONE* 3, e1548. doi: 10.1371/journal.pone.0001548.

Smith, J. E., Brainard, R., Carter, A., Grillo, S., Edwards, C., Harris, J., Lewis, L., Obura, D., Rohwer, F., Sala, E., Vroom, P. S., Sandin, S. A. (2016). Re-evaluating the health of coral reef communities: baselines and evidence for human impacts across the central Pacific. *Proceedings of the Royal Society B: Biological Sciences* 283. doi: 10.1098/rspb.2015.1985.

Stimson, J. (1974). An Analysis of the Pattern of Dispersion of the Hermatypic Coral Pocillopora Meandrina Var. Nobilis Verril. *Ecology* 55, 445–449. doi: 10.2307/1935234.

Turner, M. G. (1989). Landscape Ecology: The Effect of Pattern on Process. *Annual Review of Ecology and Systematics* 20, 171–197. doi: 10.1146/annurev.es.20.110189.001131.

Vardi, T., Williams, D., and Sandin, S. (2012). Population dynamics of threatened elkhorn coral in the northern Florida Keys, USA. *Endangered Species Research* 19, 157–169. doi: 10.3354/esr00475.

Weinberg, S. (1981). A comparison of coral reef survey methods. *Bijdragen tot de Dierkunde* 51, 199–218.

Zvuloni, A., Artzy-Randrup, Y., Stone, L., Kramarsky-Winter, E., Barkan, R., and Loya, Y. (2009). Spatio-Temporal Transmission Patterns of Black-Band Disease in a Coral Community. *PLOS ONE* 4, e4993. doi: 10.1371/journal.pone.0004993.

CHAPTER 2 ACKNOWLEDGEMENTS

Chapter 2 contains material being prepared for publication, coauthored with Petrovic, V.; Edwards, C. B.; Pedersen, N.; Alcantar, E.; Kuester, F.; Sandin, S. A. The dissertation author was the primary researcher and author of this material.

This research was made possible through efforts of the 100 Island Challenge, and the various organizations that have supported this project, including the Scripps family and the Moore Family Foundation.

CHAPTER 3

AUTOMATED 2D, 2.5D, AND 3D SEGMENTATION OF CORAL REEF POINTCLOUDS AND ORTHOPROJECTIONS

Hugh Runyan, Vid Petrovic, Clinton B. Edwards, Nicole Pedersen, Esmeralda Alcantar, Falko Kuester, and Stuart A. Sandin.

This chapter is presented as a published paper as it appears in the journal *Frontiers in Robotics and AI* in 2022 under the title "Automated 2D, 2.5D, and 3D segmentation of coral reef pointclouds and orthoprojections."

ABSTRACT

Enabled by advancing technology, coral reef researchers increasingly prefer use of image-based surveys over approaches depending solely upon *in situ* observations, interpretations, and recordings of divers. The images collected, and derivative products such as orthographic projections and 3D models, allow researchers to study a comprehensive digital twin of their field sites. Spatio-temporally located twins can be compared and annotated, enabling researchers to virtually return to sites long after they have left them. While these new data expand the variety and specificity of biological investigation that can be pursued, they have introduced the much-discussed Big Data Problem: research labs lack the human and computational resources required to process and analyze imagery at the rate it can be collected. The rapid development of unmanned underwater vehicles suggests researchers will soon have access to an even greater volume of imagery and other sensor measurements than can be collected by diver-piloted platforms, further exacerbating data handling limitations. Thoroughly segmenting (tracing the extent of and taxonomically identifying) organisms enables researchers to extract the information image products contain, but is very time-consuming. Analytic techniques driven by neural networks offer the possibility that the segmentation process can be greatly accelerated through automation. In this study, we examine the efficacy of automated segmentation on three different image-derived data products: 3D models, and 2D and 2.5D orthographic projections thereof; we also contrast their relative accessibility and utility to different avenues of biological inquiry. The variety of network architectures and parameters tested performed similarly, ~80% IoU for the genus *Porites*, suggesting that the primary limitations to an automated workflow are 1) the current capabilities of neural network technology, and 2) consistency and quality control in image product collection and human training/testing dataset generation.

AIM

The objective of this study was to discuss differences between 2D, 2.5D (RGBD), and 3D image-survey-derived maps of coral reefs, and then evaluate the ability of 2D, 2.5D, and 3D neural networks to automatically taxonomically segment (i.e., label or annotate every pixel/point of) those maps.

METHODS

Image surveys of coral reefs were conducted in Palmyra from 2013 to 2020. Those images were used, with photogrammetry/structure from motion software, to derive 3D pointcloud maps of the reef site. The genus *Porites* was then traced in the 3D pointclouds by coral experts with the software Viscore, in order to create training data for the neural networks and a ground truth against which to evaluate automated prediction accuracy. The 3D pointclouds were then used to create top-down 2D and 2.5D orthographic projections. Publicly-available code for 2D, 2.5D, and 3D neural networks was adapted to the specific application of segmenting these 2D, 2.5D, and 3D data.

RESULTS

The 2D, 2.5D, and 3D segmentation neural networks performed similarly: ~80% IoU for the genus *Porites*.

frontiers | Frontiers in Robotics and AI

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/frobt.2022.884317

Automated 2D, 2.5D, and 3D Segmentation of Coral Reef Pointclouds and Orthoprojections

Hugh Runyan[1], Vid Petrovic[2], Clinton B. Edwards[1], Nicole Pedersen[1], Esmeralda Alcantar[1], Falko Kuester[2] and Stuart A. Sandin[1]*

[1]Sandin Lab, Scripps Institution of Oceanography, University of California San Diego, San Diego, CA, United States, [2]Cultural Heritage Engineering Initiative, Qualcomm Institute, University of California San Diego, San Diego, CA, United States

Enabled by advancing technology, coral reef researchers increasingly prefer use of image-based surveys over approaches depending solely upon *in situ* observations, interpretations, and recordings of divers. The images collected, and derivative products such as orthographic projections and 3D models, allow researchers to study a comprehensive digital twin of their field sites. Spatio-temporally located twins can be compared and annotated, enabling researchers to virtually return to sites long after they have left them. While these new data expand the variety and specificity of biological investigation that can be pursued, they have introduced the much-discussed Big Data Problem: research labs lack the human and computational resources required to process and analyze imagery at the rate it can be collected. The rapid development of unmanned underwater vehicles suggests researchers will soon have access to an even greater volume of imagery and other sensor measurements than can be collected by diver-piloted platforms, further exacerbating data handling limitations. Thoroughly segmenting (tracing the extent of and taxonomically identifying) organisms enables researchers to extract the information image products contain, but is very time-consuming. Analytic techniques driven by neural networks offer the possibility that the segmentation process can be greatly accelerated through automation. In this study, we examine the efficacy of automated segmentation on three different image-derived data products: 3D models, and 2D and 2.5D orthographic projections thereof; we also contrast their relative accessibility and utility to different avenues of biological inquiry. The variety of network architectures and parameters tested performed similarly, ~80% IoU for the genus *Porites*, suggesting that the primary limitations to an automated workflow are 1) the current capabilities of neural network technology, and 2) consistency and quality control in image product collection and human training/testing dataset generation.

Keywords: coral, segmentation, automated, pointcloud, orthoprojection

OPEN ACCESS

Edited by:
Nuno Gracias,
University of Girona, Spain

Reviewed by:
Matan Yuval,
University of Haifa, Israel
Pablo Gauterio Cavalcanti,
Federal Technological University of Paraná, Brazil

****Correspondence:***
Hugh Runyan
hurunyan@ucsd.edu

Specialty section:
This article was submitted to Robot and Machine Vision, a section of the journal Frontiers in Robotics and AI

Received: *26 February 2022*
Accepted: *06 May 2022*
Published: *27 May 2022*

Citation:
Runyan H, Petrovic V, Edwards CB, Pedersen N, Alcantar E, Kuester F and Sandin SA (2022) Automated 2D, 2.5D, and 3D Segmentation of Coral Reef Pointclouds and Orthoprojections. Front. Robot. AI 9:884317. doi: 10.3389/frobt.2022.884317

1 INTRODUCTION

Unmanned remote sensing platforms offer the prospect of collecting for monitoring and analysis vastly more image-based data than was previously possible. In particular, the high-resolution imagery made available by the IKONOS and MODIS satellites revolutionized macroecological studies across terrestrial habitats (Pfeifer et al., 2012). In the marine sciences, these technologies have

enabled observation of the entirety of the planet's oceans regularly. However, the opacity of water prevents similarly comprehensive viewing of the ocean floor beyond the shallowest areas. Capabilities of unmanned underwater vehicles (UAVs) for closer, finer-scaled inspection are not as well-developed as satellites but are advancing rapidly. In shallow, nearshore environments like coral reefs, ecological research groups are increasingly transitioning away from diver-recorded data to manned underwater imagery surveys—analogous to what a UAV would collect—that can be archived and analyzed (Weinberg, 1981; Gracias and Santos-Victor, 2000).

Using structure from motion (SfM) algorithms, 3D models of reef tracts can be derived from survey imagery (Pizarro et al., 2009; Smith et al., 2016; Edwards et al., 2017; Ferrari et al., 2017; Kodera et al., 2020; Sandin et al., 2020), enabling analysis in a data medium that much more closely resembles real-world coral environments, particularly in steep or geometrically-complex areas that are common in reefs but poorly captured with top-down imagery. A variety of metrics, including percent cover, growth, species composition, or disease or bleaching incidence can be extracted from 3D models, or 2D orthographic projections (orthoprojections) thereof, using random point sampling or full semantic segmentation (i.e., labeling per-point, per-taxa; hereafter referred to simply as segmentation). Image-based data increases the amount of information collected from a field site relative to *in situ* observation approaches, but it introduces new data extraction challenges: it is much simpler to amass a library of thousands of reef sites than it is to extract the ecological information contained in that library.

Neural networks and other computational/algorithmic processing and analytical strategies offer hope that such backlogs can be cleared. A number of studies have evaluated automated segmentation methods, both neural-network-based and otherwise, for classifying (applying a label to an entire scene) or segmenting (applying labels per-pixel) 2D images of coral (Beijbom et al., 2012) or orthoprojections of 3D models of coral reefs (Alonso et al., 2017, 2019; Alonso and Murillo, 2018; Yuval et al., 2021). Beyond coral applications, neural networks for top-down photographic media have been tested in 2D, 2.5D (RGBD), 3D, and hybrid formats in contexts such as satellite imagery with promising results (Mohammadi et al., 2019; Bachhofner et al., 2020; Saralioglu and Gungor, 2020; Song and Choi, 2020). More generically, many studies have examined 2D, 3D, and hybrid methods of segmentation, with the two most common applications being 1) scene analysis to inform automated decision-making for driverless cars, grasping arms, and other robotic applications, and 2) rapid or assisted interpretation of diagnostic medical imagery. Development in these fields is quite rapid, but as of this date no automated solution exists to segment corals with the accuracy desired by coral scientists hoping to reliably track growth on the scale of, in some cases, mere millimeters per year.

In this study, we compare the performance of 2D, 2.5D, and 3D segmentation neural networks on 10 m × 10 m reef pointclouds, and 2D and 2.5D orthoprojections thereof, from Palmyra Atoll in the Pacific. We aim to contribute to the effort of evaluating the utility of neural-network-expedited segmentation workflows, as well as mapping out scenarios where inexpensive and fast 2D analysis is adequate and those where benefits of 3D are substantial.

2 MATERIALS AND METHODS

2.1 Photographic Survey, 3D Model and 2D Orthoprojection Derivation

The details of our field sampling design have been discussed in detail elsewhere (Edwards et al., 2017; Fox et al., 2019; Kodera et al., 2020; Sandin et al., 2020). Briefly, the photographic surveys were conducted by a pair of divers within 100 m^2 plots positioned along the 10m isobath in oceanic fore reef habitats. Utilizing a custom-built frame containing two Nikon d7000 DSLR cameras mounted in tandem, the diver operating the camera collected imagery *via* a lawnmower pattern 1.5 m above the site, with cameras set to a 1 s interval timer and oriented straight down in the direction of gravity (as opposed to perpendicular to the surface plane of the reef). The first camera was equipped with a wide angle 18 mm lens to ensure the substantial overlap required for accurate 3D model reconstruction, while the second was equipped with a 55 mm lens to provide the visual detail needed to disambiguate complex species-level identifications. During each survey, a series of scale bars and boundary markers were placed throughout the plot, which are visible in the final reconstructed model. At each plot boundary marker, a second diver recorded depth information to establish orientation of the plot relative to the plane of the ocean surface.

The details of the technical processing software used to generate the 3D models have been described previously (Westoby et al., 2012; Sandin et al., 2020). Briefly, the models used in this study were created with Photoscan, now known as Metashape, which is developed by Agisoft LLC. Scale bars and depth measurements were used as ground control points to determine scale and orientation. The resulting model is a pointcloud: a list of points with XYZ spatial coordinates and corresponding RGB color values.

Pointcloud visualization and geometric analysis in this study was performed in the software package Viscore, developed at UC San Diego in the Cultural Heritage Engineering Initiative/Kuester Lab (Petrovic et al., 2014; Fox et al., 2019; Sandin et al., 2020). Viscore is a visual analysis platform that facilitates a number of 3D workflows useful for performing virtual fieldwork, including interactive alignment and inspection of time-series site representations, manual plot segmentation in 3D, import/export/filtering of point subsets, and generation of orthoprojected maps and digital elevation models.

2.2 Dataset

The models studied in this report were collected annually from 2013 to 2020 on Palmyra Atoll at a single 10 m × 10 m site. Top-down 2D orthoprojections were derived for each 3D model at a scale of 1 mm/pixel. Corresponding 2.5D versions of the orthoprojections were produced by including the Z depth value for each pixel. All *Porites* (a coral genus) within the study site were traced by hand in Viscore. Therefore, every

TABLE 1 | Coral segmenters, their experience, the plot(s) they traced, and IoU on the 3 m × 3 m subsection of the 2020 model relative to consensus.

Tracer	Experience	Years traced	IoU v. consensus on 2020 subsection [%]
Hugh Runyan (HR)	Non-expert	2013, 2014, 2018, 2019, 2020	88.6
Clint Edwards (CE)	Expert	2016	93.7
Nicole Pedersen (NP)	Expert	2015	94.1
Esmeralda Alcantar (AE)	Expert	2017	92.4
	-		Mean: 92.2

point/pixel had one of two possible labels: *Porites* or everything else. For ingestion into neural networks, 3D models were split into 1 m × 1 m patches along the X and Y axes (defined as the plane of constant depth), while orthoprojections were split into image patches 512 pixels on a side. This resulted in 100 1 m^2 3D patches and 400 512 × 512 pixel patches per year. Each 512^2 pixel patch corresponds to approximately 0.25 m^2. These dimensions were chosen because they had roughly similar memory impact in our experiments, resulting in the use of equivalent batch size (8) during training across all dimensionalities. 60% of patches from years 2013–2019 (420 pointcloud patches, 1,680 image patches) were randomly assigned to the training subset, while another 20% each was reserved for validation and testing. To test how the neural network models would perform on a model no part of which was in the training set, the entire 2020 model was reserved for separate testing.

Performance on data collected under different circumstances than that which is in the training set is a critical metric when trying to determine how useful an automated tool is. This concept is known as generalization. It is analogous to the difference between memorization and conceptual understanding—tools that generalize well perform consistently across different times, places, and contexts, while those that generalize poorly only perform well on inputs that closely resemble examples they have already seen. In the context of this study, a neural network that generalizes well would accurately segment *Porites* photographed anywhere in the world, with any camera, at any time of day, etc., because it has developed an understanding of how to distinguish *Porites*, while one that generalizes poorly would only segment accurately when presented with *Porites* from our single site on Palmyra between the years 2013 and 2019 because it has memorized how to make correct predictions in that specific context. Automated tools that generalize well are much more useful than those that do not.

Porites in each 10 m × 10 m plot were manually segmented in Viscore by one of the authors (NP, CE, HR, or EA). Viscore offers high-framerate 3D rotation and zooming, survey image draping, and point labeling with adjustable-size paintbrushes to make the manual segmentation process as easy and accurate as possible. Which author segmented which plot is shown in **Table 1**. The segmenters are separated into two categories: those with extensive coral identification experience ("experts"), and those with less ("non-experts"). All four of the authors also segmented the same 3 m × 3 m subsection of the 2020 plot to evaluate consistency between ourselves.

2.3 Segmentation With Neural Networks

Unsegmented 3D pointclouds and derived 2D orthoprojections can be used to investigate a variety of characteristics of ecological importance, such as 3D reef structure and qualitative observations of coral colony health. However, far more information is available if each organism is identified and segmented. Repeated surveys of a location depict the recruitment, growth, and death of individual corals, so they can be used to evaluate competition, successional states, equilibrium, anthropogenic impacts, and more. Manually segmenting a single square meter requires an average of 1 hour; at this rate, an expert working full time would need at least a century to map every organism in the ~2000 pointclouds thus far collected by the research labs contributing to this report. Expediting segmentation is therefore necessary to make full use of photographic survey products. Neural networks are the state-of-the-art solution for segmentation of visual media such as images, videos, and 3D models.

For this study, we tested SparseConvNet (SCN), a successful point cloud segmentation neural network architecture (Graham et al., 2018). SCN does not operate directly on pointclouds, but instead voxelizes (the 3D equivalent of a pixel) the input pointcloud to a chosen grid size. Voxels with side length 2 mm were used for this study. Smaller voxels increased computation time but did not improve predictions. 3D computations are done sparsely, meaning on a list of occupied voxels instead of as matrix operations on every location whether it is occupied or not. Sparse operations greatly improve speed and efficiency operations on data in which most matrix locations are empty, which is the case with our reef surface; there is no useful information above or below the surface, so a 3D cube containing that surface is mostly empty space. SCN is fast, simple, and provides the additional benefit of operating on discretized locations, just as the 2 and 2.5D image segmentation neural networks we used in this study do. For 2.5D segmentation we used ESANet (Seichter et al., 2021), while 2D segmentation was done with ESANet with its depth channel components disabled (in this configuration it only considers RGB color information, as opposed to color and depth). ESANet uses a ResNet-34 (He et al., 2015) encoder with an additional "context module," similar to those found in the popular DeepLab architectures (Chen et al., 2018), for detecting large-scale patterns. These architectures were chosen because they perform well on publicly-available benchmarks and are among the more commonly used.

All networks were trained with batch size 8 for 512 epochs with the Adam optimizer (Kingma and Ba, 2015) and the cross

TABLE 2 | Highest automated segmentation accuracy of each neural network dimensionality as determined by validation IoU score.

	Porites IoU (2013–2019 validation set) [%]	*Porites* IoU (2013–2019 test set) [%]	*Porites* IoU (2020) [%]
2D	82.8	82.0	57.4
2.5D (RGBD)	83.1	81.4	45.0
3D	81.4	82.3	70.4

entropy loss function, using similar color, scale, and deformation augmentations. During training, the networks were evaluated with the 2013–2019 validation set and saved at regular intervals (every 10 epochs), after which the best performing of the saved versions were selected and then tested on the 2013–2019 and 2020 test sets. The results of these tests are shown in **Table 2**.

The difference between 2.5D and 3D neural networks is subtle. 2.5D neural networks operate in flat, two-dimensional windows placed over an image; using, say, a 5 × 5 window of pixels, the network searches for patterns in each of the RGB color channels, as well as the Z depth channel, and combinations thereof. In a 3D neural network, the analogous window would be 5 × 5 × 5. Patterns of color are sought in different 3D shapes, just as in 2D neural networks patterns of color are sought in 2D shapes. The primary shortcoming of the 2.5D approach is that each pixel can only contain one value for each of R, G, B, and Z—if two surfaces share the same horizontal location in XY, only the one with the higher Z value will be stored in an RGBD image.

2.4 Error Metrics

For training we used the cross-entropy loss function. To evaluate predicted segmentations, we use the intersection over union (IoU) error metric. For each class (meaning possible labels, here just 1) *Porites* and 2) everything else), IoU is defined as the intersection of the prediction and ground truth—the true positive count—divided by the union of the prediction and ground truth—the sum of the true positive, false positive, and false negative counts: TP/(TP + FP + FN).

In contrast, accuracy is defined as (TP + TN)/(TP + TN + FP + FN). Because prediction is relatively easy for most of the ~95% of our plots that is not *Porites*, TN will be a very large number relative to TP for the *Porites* class. If the automated segmentation network identified 5% of the plot as *Porites* but it was entirely the wrong 5%, it would still be 90% accurate, because only 10% of the points are mislabeled (5% that should have been labeled *Porites* but weren't and 5% that should not have but were). This means that segmentation accuracy is a less useful metric for performance evaluation than IoU; IoU better reflects our goal of tracking the growing edges of individual corals of specific genera that often comprise only a small fraction of the total area of the plot.

3 RESULTS AND DISCUSSION

3.1 Consistency Amongst Labelers

All four authors that segmented the same 3 m × 3 m subsection of the 2020 plot agreed on 76% of points that were labeled *Porites* by at least one author, meaning that, of the points labeled *Porites* by at least one segmenter, 76% of them were labeled *Porites* by all four. These segmentations are illustrated in **Figure 1**, while the disagreed-upon points are isolated in **Figure 2**. This bulk disagreement metric is not ideal however, as it is likely to be influenced by the number of segmenters (more people will disagree more often than fewer people). We created a more stable alternate accuracy metric by first creating a "ground truth" from majority consensus labeled points (those labeled *Porites* by at least three of us), and then calculated IoU error against that ground truth for each labeler. Results are shown in **Table 1**. The mean IoU of all labelers against majority consensus ground truth was 92%, with the non-expert scoring the lowest at 88%.

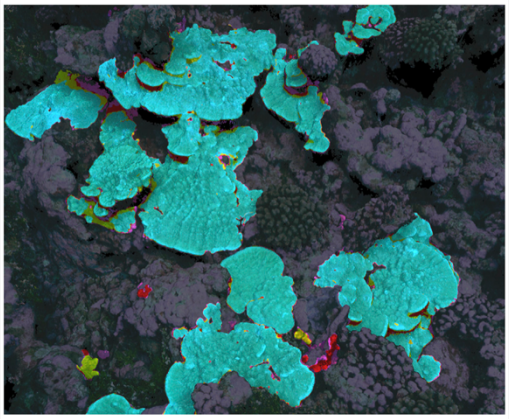


FIGURE 1 | 3D model showing agreement and disagreement amongst manual segmenters. Dark background is consensus background, while blue highlights show agreed-upon *Porites* by all four segmenters. Points highlighted yellow were labeled *Porites* by three segmenters, red by two, and purple by only one.

3.2 Automated Segmentation

Prediction IoU scores for the top-performing 2D, 2.5D, and 3D neural networks as determined by validation IoU score are reported numerically in **Table 2**, while a visual example of 3D results is shown in **Figure 3**, corresponding to the same area of the 2020 model that was used for the consistency experiment in

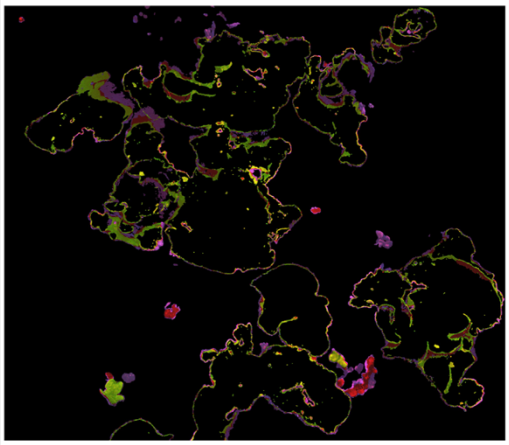


FIGURE 2 | 3D model showing manual segmenter *Porites* disagreement only. Points highlighted yellow were labeled *Porites* by three of four segmenters, red by two, and purple by only one.

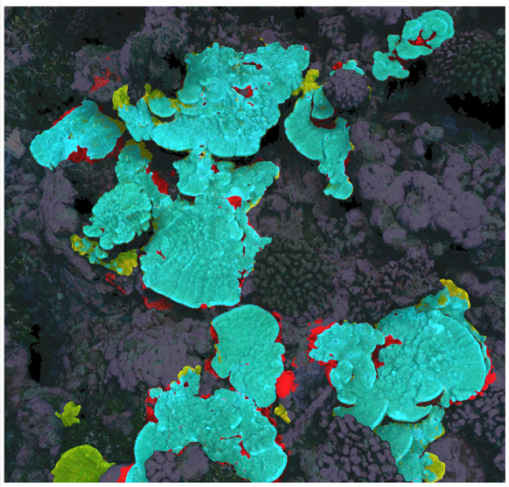


FIGURE 3 | Example of 3D neural network prediction (correct predictions are highlighted blue, false positive predictions are red, and false negative are yellow), corresponding to the same area as **Figures 1**, **2**.

Section 3.1. 2D, 2.5D, and 3D approaches performed similarly (~80% IoU) on the 2013–2019 validation and test sets, but performance on the 2020 plot held out of the training dataset varied: the 3D network achieved the highest 2020 score (70.4%), while the 2D network scored second highest (57.4%), and the 2.5D network scored the lowest (45.0%).

TABLE 3 | Automated segmentation accuracy of the 3D network on the 2013-2019 test set broken down by year.

Year	*Porites* IoU (test set) [%]
2013	78.5
2014	83.1
2015	84.6
2016	85.6
2017	77.2
2018	85.3
2019	84.9

Prediction IoU scores on the 2013–2019 test set of the most successful 3D network (as determined by validation IoU) are broken down by year in **Table 3**. Scores were relatively stable, ranging from 77.2% to 85.6%. Scores on the 2015, 2016, and 2017 models traced by experts were 84.6%, 85.6%, and 77.2% respectively, while scores on non-expert-traced models were 78.5%, 83.1%, 85.3%, and 84.9%.

The consistency of performance across media dimensionality is likely explained by a combination of factors. First and foremost, our 2D and 2.5D orthoprojections are derived directly from the 3D model, which is constructed from top-down survey imagery. As such, there is little in the 3D models that is not in the orthoprojections because images were not captured from side-on angles. While the extra dimensions offer additional opportunities for pattern detection, if there is little difference in significant information between the three input dimensionalities, then there are few resulting additional patterns for the higher-dimensional networks to detect. A second important factor is that these 2D, 2.5D, and 3D neural networks are algorithmically quite similar: they all depend on cascaded multi-scale convolutional pattern detectors. Fundamentally, neural networks are optimization problems just like curve-fitting, and as in that simpler scenario there are often a number of similarly-performing algorithmic alternatives but none that are near-perfect without overfitting.

A third factor to consider is labeling errors and subjective decisions, in both the training set and the ground truth used to evaluate performance—it is possible that the accuracy/consistency of our *Porites* training and testing dataset is only ~80%. If there are errors in training data, a neural network might fail to learn useful information or learn incorrect information, while errors in the validation and test sets introduce uncertainty into performance metrics. As can be seen in **Figures 1**–**3**, much of both segmenter disagreement and neural network error concerned exact placement of coral colony boundaries. If training data is inconsistent in those areas, the neural network may not be able to learn how to be precise and consistent when delineating boundaries, and if the validation/test data is inconsistent in those areas it may be difficult to tell from performance metrics like IoU if the network is right or wrong.

The comparison between contributing segmenters (**Section 3.1**, **Table 1**) suggested individual segmenters score ~90% by IoU

when compared to a more reliable consensus segmentation. However, that experiment was only conducted on one small 3 m × 3 m patch of reef where segmenters knew they would be compared to one another. It is possible the segmenters were fatigued or less careful when segmenting entire 10 m × 10 m models and made more mistakes, especially as four of the seven models were segmented by the least-experienced contributor (HR). If the 2013, 2014, 2018, and 2019 segmentations produced by the least-experienced author (HR) were indeed more flawed, we might expect the neural network IoU prediction scores on those years to be lower, but the per-year results in **Table 3** show that automated performance was relatively consistent across years, and the lowest score (77.2%) occurred on a model traced by one of the experts.

The better performance of the 3D network on the 2020 plot, none of which was in the training set, could be because, while color varies due to e.g. camera settings and lighting conditions, 3D structure is more consistent across samples. If this explanation were correct, we would expect the 2.5D network to outperform 2D on the 2020 dataset because 2.5D also contains information about physical structure, but that was not the case (**Table 2**). It is possible that the 3D networks are also structurally smaller and simpler than 2D or 2.5D networks because the latter are more memory efficient and developmentally mature. More complex networks often perform better, but sometimes become more sensitive to small deviations in data distribution (generalize more poorly).

Generally speaking, these results should not be expected to consistently extrapolate to any coral dataset from anywhere on the globe. This is especially true of performance on the 2020 dataset. Neural networks' accuracy depend on a number of factors, such as: the number of genera they are asked to predict and the phenotypic plasticity of those genera, variation in lighting conditions during photo surveys, including turbidity and depth-dependent color changes, specifics of the camera sensor and settings including resolution, white balance, focal length, manual or autofocus settings, as well as the image acquisition pattern, vagaries of the 3D reconstruction process, and more. As shown by the relatively good and consistent performance of all three dimensionalities on the 2013–2019 validation and test sets, neural networks can learn to account for these variations if they are present in the training dataset—if training and testing sets strongly resemble one another. The variable and unreliable performance on the 2020 dataset demonstrates a commonly-understood shortcoming of neural networks: they often don't perform well on samples from outside the dataset on which they were trained. IoU scores on the 2020 dataset were dramatically lower, even though the 2020 dataset was from the same geographic site on Palmyra as 2013–2019 data (i.e., the corals were mostly the same, excepting a year's worth of growth and death) and all datasets were collected with similar equipment and procedures. Many strategies have been developed on how to limit performance loss on unseen data, generally referred to as domain adaptation or domain generalization, but they are beyond the scope of this report.

Coral researchers hoping to expedite segmentation with neural networks should strive for consistency, comprehensiveness, and quality control in field image collection methodology and manual segmentation of training and validation data, but the results in this report suggest automated performance loss is likely in applications concerning unsegmented new survey sites or resurveys of existing sites.

For most ecological applications, the neural network prediction accuracy achieved in this study is not adequate, so researchers must manually correct the automated predictions before they can be used in analysis. Such human-in-the-loop workflows (Pavoni et al., 2021) are more realistic at this time than fully automated ones, and they have the added advantage of a human expert verifying the segmentations as they are created. Relying on automated predictions carries the risk of the accuracy of those predictions changing without researcher detection, resulting in mislabeled data and erroneous conclusions. However, this manual editing and verification process is time-consuming: in our experience, the time required to segment an entire plot *de novo* and the time required to edit predictions is often similar. Generally, the most time-consuming aspect of manual segmentation is tracing complex boundaries. Unfortunately, as illustrated in **Figure 3**, that is the area neural network predictions tend to be least reliable, so boundaries often must be segmented manually whether one is editing imperfect predictions or segmenting *de novo*. Further, **Figures 1**, **2** show that boundaries are also the place manual segmenters are most likely to disagree. The boundaries of coral segmentations can be viewed as zones of uncertainty, where neither humans nor neural networks are reliable. Researchers intending to use segmentations of image products to study mm-scale growth at the boundaries of individual coral colonies are advised to rigorously engage this topic.

3.3 Considerations of Data Dimensionality

While it is a well-vetted analytical approach, in many applications working in 2D has significant shortcomings—largely arising from a 2D model's inherent limitations in fully representing a 3D environment. Sites exhibiting greater geometric complexity—with ledges and overhangs, or with branching coral morphologies (e.g. *Acropora*)—cannot be fully represented in either 2D or 2.5D. Aspects of 3D structure unaffected by occlusion—colony height, say, or benthic slope or rugosity—are representable in 2.5D but not in 2D. Another more subtle issue arises concerning the direction of orthoprojection: the pattern of occlusions, as well as the relative density at which different parts of a site are (re) sampled (for gridded models such as DEMs) both depend on the chosen projection direction. Furthermore, 2D measurements—distances, projected areas, angles, (non-detrended) rugosity estimates—are all affected by the choice of projection. Selecting the projection direction systematically and consistently is especially important for making quantitative comparisons (e.g. growth rates) across time series. Orthoprojecting along the direction of gravity is often a reasonable standard—however, many reefs contain habitats of interest that feature steep slopes, or stark vertical structures, especially on shallow reef flats and forereefs, both of which are critical habitats for study. In such cases, orthoprojecting along the

direction of gravity makes little sense as the projected area of organisms in these habitats will represent a fraction of their true footprint in 3D. Measuring the exact vertical direction underwater can also be difficult and imprecise, which introduces projection error—particularly undesirable when trying to accurately measure a few millimeters of growth at colony edges.

However, working in 3D presents its own set of challenges, making greater demands than 2D on every stage of the workflow, from acquisition to analysis to dissemination. Full 3D datasets are much larger than 2D—orthoprojections range from 0.5–2 GB per time point, with depths per pixel adding another ~1 GB, while the 3D models in this study range from 10 GB to 75 GB per time point—and require significantly more computational power to work with. This technical challenge is addressed by the Viscore platform, allowing multi-year multi-site projects (easily totaling several TB of data) to be inspected and annotated interactively on commodity laptops.

Even with data-handling impediments removed, working in 3D is generally more difficult than in 2D. Reconstructing complete 3D models (without holes) requires more photographs, from more view-angles, than are required for 2D. In this work, we photograph primarily with the camera pointed straight down: the resulting models are incomplete underneath overhangs or on vertical surfaces where sight lines are occluded. The peripheries of images, especially from wide-angle lenses, capture much of the information necessary to build accurate digital surrogates, but they nevertheless do not capture the entirety of the 3D structure of the reef.

This incompleteness of 3D coverage complicates analysis. Measuring colony size and growth is one example: while top-down projected areas can be reliably measured in 2D (subject to the 2D caveats discussed above), making similar measurements in 3D (surface area and volume) requires more care. Measuring total 3D surface area is relatively straightforward for well-defined surfaces, but models resulting from incomplete coverage (e.g., only top-down views) require a degree of assumption or interpolation in areas of low coverage and reconstruction point density. Further, estimates of volume can be particularly problematic, as for many coral taxa they require assumptions of where the coral skeleton ends and the underlying substrate begins—information often not captured by top-down imagery—though changes in volume can be more reliably estimated. Scale and model resolution also have a significant impact on 3D measures: coral often has much more 3D surface area than 2D (particularly branching, corymbose, and foliose forms), due to structures ranging from m- or cm-sale (e.g., branches) to sub-mm corallite walls, ridges, and other delicate features. Accurately capturing 3D structure at such small scales is difficult. In an ideal world the coral research community would work exclusively with complete sub-mm resolution 3D models, but that is not always practical.

Those coral researchers pursuing applications where the drawbacks of 2D are relatively minimal (e.g., where the morphology of the taxa of interest is not overly complex, the bathymetry is relatively flat, the most useful information is contained in color rather than 3D morphology) will likely prefer 2D due to speed and simplicity. On the other hand, those more concerned with geometric fidelity in complex environments will prefer to operate with 3D pointclouds if the necessary resources can be mustered.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

HR did the neural network programming and testing. SS, EA, NP, and CE collected image products in the field, created 3D pointclouds from them, and oversaw manual segmentation. VP provided orthoprojections of the pointclouds as well as database and software support. SS and FK played advisory roles.

ACKNOWLEDGMENTS

Thank you to The Nature Conservancy and the Palmyra Atoll Research Consortium for logistical support and the United States Fish Wildlife Service for special use permit # 12533-13025 and access to the refuge. This work was in part made possible in part with the support of the Gordon and Betty Moore Foundation, Reef of Tomorrow Initiative and The Beyster Family Foundation.

REFERENCES

Alonso, I., Cambra, A., Munoz, A., Treibitz, T., and Murillo, A. C. (2017). "Coral-Segmentation: Training Dense Labeling Models with Sparse Ground Truth," in Proceeding of the 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, 22-29 Oct. 2017 (IEEE), 2874–2882. doi:10.1109/ICCVW.2017.339

Alonso, I., and Murillo, A. C. (2018). "Semantic Segmentation from Sparse Labeling Using Multi-Level Superpixels," in Proceeding of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1-5 Oct. 2018 (IEEE), 5785–5792. doi:10.1109/IROS.2018.8594185

Alonso, I., Yuval, M., Eyal, G., Treibitz, T., and Murillo, A. C. (2019). CoralSeg: Learning Coral Segmentation from Sparse Annotations. *J. Field Robot.* 36, 1456–1477. doi:10.1002/rob.21915

Bachhofner, S., Loghin, A.-M., Otepka, J., Pfeifer, N., Hornacek, M., Siposova, A., et al. (2020). Generalized Sparse Convolutional Neural Networks for Semantic Segmentation of Point Clouds Derived from Tri-stereo Satellite Imagery. *Remote Sens.* 12, 1289. doi:10.3390/rs12081289

Beijbom, O., Edmunds, P. J., Kline, D. I., Mitchell, B. G., and Kriegman, D. (2012). "Automated Annotation of Coral Reef Survey Images," in Proceeding of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, 1170–1177. doi:10.1109/CVPR.2012.6247798

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.

ArXiv180202611 Cs. Available at: http://arxiv.org/abs/1802.02611 (Accessed July 2, 2021).doi:10.1007/978-3-030-01234-2_49

Edwards, C. B., Eynaud, Y., Williams, G. J., Pedersen, N. E., Zgliczynski, B. J., Gleason, A. C. R., et al. (2017). Large-area Imaging Reveals Biologically Driven Non-random Spatial Patterns of Corals at a Remote Reef. *Coral Reefs* 36, 1291–1305. doi:10.1007/s00338-017-1624-3

Ferrari, R., Figueira, W. F., Pratchett, M. S., Boube, T., Adam, A., Kobelkowsky-Vidrio, T., et al. (2017). 3D Photogrammetry Quantifies Growth and External Erosion of Individual Coral Colonies and Skeletons. *Sci. Rep.* 7, 1–9. doi:10.1038/s41598-017-16408-z

Fox, M. D., Carter, A. L., Edwards, C. B., Takeshita, Y., Johnson, M. D., Petrovic, V., et al. (2019). Limited Coral Mortality Following Acute Thermal Stress and Widespread Bleaching on Palmyra Atoll, Central Pacific. *Coral Reefs* 38, 701–712. doi:10.1007/s00338-019-01796-7

Gracias, N., and Santos-Victor, J. (2000). Underwater Video Mosaics as Visual Navigation Maps. *Comput. Vis. Image Underst.* 79, 66–91. doi:10.1006/cviu.2000.0848

Graham, B., Engelcke, M., and Maaten, L. v. d. (2018). "3D Semantic Segmentation with Submanifold Sparse Convolutional Networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9224–9232. doi:10.1109/CVPR.2018.00961

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. ArXiv151203385 Cs. Available at: http://arxiv.org/abs/1512.03385 (Accessed April 9, 2022).

Kingma, D. P., and Ba, J. (2015). "Adam: A Method for Stochastic Optimization," in *3rd International Conference on Learning Representations, ICLR 2015*. Editors Y. Bengio and Y. LeCun San Diego, CA, USA. Conference Track Proceedings. Available at: http://arxiv.org/abs/1412.6980 (Accessed April 5, 2022).

Kodera, S. M., Edwards, C. B., Petrovic, V., Pedersen, N. E., Eynaud, Y., and Sandin, S. A. (2020). Quantifying Life History Demographics of the Scleractinian Coral Genus Pocillopora at Palmyra Atoll. *Coral Reefs* 39, 1091–1105. doi:10.1007/s00338-020-01940-8

Mohammadi, H., Samadzadegan, F., and Reinartz, P. (2019). 2D/3D Information Fusion for Building Extraction from High-Resolution Satellite Stereo Images Using Kernel Graph Cuts. *Int. J. Remote Sens.* 40, 5835–5860. doi:10.1080/01431161.2019.1584417

Pavoni, G., Corsini, M., Ponchio, F., Muntoni, A., and Cignoni, P. (2021). TagLab: A Human-Centric AI System for Interactive Semantic Segmentation. ArXiv211212702 Cs. Available at: http://arxiv.org/abs/2112.12702 (Accessed January 28, 2022).

Petrovic, V., Vanoni, D. J., Richter, A. M., Levy, T. E., and Kuester, F. (2014). Visualizing High Resolution Three-Dimensional and Two-Dimensional Data of Cultural Heritage Sites. *Mediterr. Archaeol. Archaeom.* 14, 93–100.

Pfeifer, M., Disney, M., Quaife, T., and Marchant, R. (2012). Terrestrial Ecosystems from Space: a Review of Earth Observation Products for Macroecology Applications. *Glob. Ecol. Biogeogr.* 21, 603–624. doi:10.1111/j.1466-8238.2011.00712.x

Pizarro, O., Eustice, R. M., and Singh, H. (2009). Large Area 3-D Reconstructions from Underwater Optical Surveys. *IEEE J. Ocean. Eng.* 34, 150–169. doi:10.1109/JOE.2009.2016071

Sandin, S. A., Edwards, C. B., Pedersen, N. E., Petrovic, V., Pavoni, G., Alcantar, E., et al. (2020). "Considering the Rates of Growth in Two Taxa of Coral across Pacific Islands," in *Advances in Marine Biology Population Dynamics of the Reef Crisis*. Editor B. M. Riegl (Academic Press), 167–191. doi:10.1016/bs.amb.2020.08.006

Saralioglu, E., and Gungor, O. (2020). Semantic Segmentation of Land Cover from High Resolution Multispectral Satellite Images by Spectral-Spatial Convolutional Neural Network. *Geocarto Int.* 37, 657–677. doi:10.1080/10106049.2020.1734871

Seichter, D., Köhler, M., Lewandowski, B., Wengefeld, T., and Gross, H.-M. (2021). Efficient RGB-D Semantic Segmentation for Indoor Scene Analysis. ArXiv201106961 Cs. Available at: http://arxiv.org/abs/2011.06961 (Accessed April 9, 2022).doi:10.1109/icra48506.2021.9561675

Smith, J. E., Brainard, R., Carter, A., Grillo, S., Edwards, C., Harris, J., et al. (2016). Re-evaluating the Health of Coral Reef Communities: Baselines and Evidence for Human Impacts across the Central Pacific. *Proc. R. Soc. B* 283, 20151985. doi:10.1098/rspb.2015.1985

Song, A., and Choi, J. (2020). Fully Convolutional Networks with Multiscale 3D Filters and Transfer Learning for Change Detection in High Spatial Resolution Satellite Images. *Remote Sens.* 12, 799. doi:10.3390/rs12050799

Weinberg, S. (1981). A Comparison of Coral Reef Survey Methods. *Bijdr. Tot Dierkd.* 51, 199–218. doi:10.1111/j.1746-1561.1981.tb02226.x

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., and Reynolds, J. M. (2012). 'Structure-from-Motion' Photogrammetry: A Low-Cost, Effective Tool for Geoscience Applications. *Geomorphology* 179, 300–314. doi:10.1016/j.geomorph.2012.08.021

Yuval, M., Alonso, I., Eyal, G., Tchernov, D., Loya, Y., Murillo, A. C., et al. (2021). Repeatable Semantic Reef-Mapping through Photogrammetry and Label-Augmentation. *Remote Sens.* 13, 659. doi:10.3390/rs13040659

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

CHAPTER 3 ACKNOWLEDGEMENTS

Chapter 3, in full, is a reprint of material as it appears in *Frontiers in Robotics and AI*, Runyan, H.; Petrovic, V.; Edwards, C. B.; Pedersen, N.; Alcantar, E.; Kuester, F.; Sandin, S. A; 2022. The dissertation author was the primary investigator and author of this paper.

CHAPTER 4

AN EVALUATION OF THE NEURAL NETWORK TOOL CORALNET FOR LARGE-SCALE CORAL RESEARCH

Hugh Runyan, Clinton B. Edwards, Nicole Pedersen, Vid Petrovic, Katie Lubarsky, Nathaniel Holloway, Falko Kuester, and Stuart A. Sandin

ABSTRACT

In recent years, coral reef researchers have increasingly adopted image-based surveys as a means to collect information from field sites. Previous methods, based on *in situ* observations, recordings, and interpretations often capture less information from the field site than these much more comprehensive recently-developed practices. However, image-based methods, for many research applications, require not just the capturing of images, but also some form of annotation or interpretation of what the images contain. A popular annotation method is point sampling: randomly selecting a subset of image pixels to taxonomically classify, rather classifying all of them. This enables researchers to use relatively simple and quick computer user interfaces to create survey-image-based measurements for analysis. However, this process is quite slow, so there is growing interesting in the research community to harness rapidly-improving neural network capabilities to further expedite this point sampling process. The most popular neural-network-based tool among coral researchers is CoralNet, which trains neural networks for users without requiring expertise on machine learning techniques or the ability to use programming languages. In this study, the user experience of CoralNet was evaluated, with the hope of providing information to researchers who may be considering investing their time and effort in adopting it in lieu of their existing entirely-manual point sampling methodology. In the first of two experiments, CoralNet's performance was evaluated as a function of the number of expert-created point sampled images provided for neural network training. In the second experiment, experts used CoralNet-generated AI predictions to investigate their effect on the time require to point sample images. When using the large and specific sets of possible taxonomic classifications preferred by coral researchers, the first experiment suggests that overall neural network accuracy rises quickly with a small number of point sampled training images, and then improves much more slowly as training

samples increases. However, the accuracy of relatively rare taxonomies, which encompassed the vast majority of possible classification labels, rose slower than overall accuracy, and was still increasing after 5000 point sampled images were uploaded. Accuracy was significantly higher for more generic functional group classifications than specific genus or species level labels. These results suggest that, when using a small set of broader and more generic classifications, relatively high accuracy can be achieved with relatively small training sets. However, much larger training sets are required to achieve similar same neural network accuracy on larger and more specific sets of possible point labels (if similar accuracy is achievable with modern neural network technology at all), as many of these classifications occur only a small number of times even in seemingly-large collections of point sampled images. The second experiment showed that, even with imperfect CoralNet AI predictions, experts using those predictions to point sample images were able to do so 36% faster on a per-point basis, representing a significant improvement in the time required to create expert-verified point sampled image datasets.

4.1 INTRODUCTION

Coral reef researchers increasingly prefer image-based surveys over methods depending on divers' *in situ* observations, recordings, and interpretations. Image surveys capture much more information from the field site, enabling variety and flexibility in later analysis. However, these newer and more technologically-complicated methods introduce additional complications; diver recordings can often be analyzed directly, while images require significant processing/annotation back in the lab. Depending on the research application, they may need to be annotated, color-corrected, or arranged by the geometric relationship from which they were captured. Ocean science labs rarely have the resources to manually process hundreds or thousands of images from a single site, let alone many sites. Unmanned underwater vehicles will soon provide access to an even greater volume of imagery, further exacerbating the big data problem.

The rapidly-advancing computer vision capabilities of neural networks (NNs) offer hope that future semi- or fully-automated processing pipelines will be able to keep up with the rate images can be collected in the field (Krizhevsky et al., 2012b). The most widely-used NN tool in the coral science community is CoralNet, hosted at coralnet.ucsd.edu (Beijbom et al., 2012). ReefCloud, developed by the Australian Institute of Marine Science, offers similar functionality, but is less used: a Google Scholar search for CoralNet returns 618 results, while ReefCloud returns 29. CoralNet's primary functionality is to support a method called point sampling: a random subset of image pixels is chosen, and then the organisms those points/pixels correspond to are taxonomically identified. CoralNet is free to use and cloud-based, so researchers from around the world can upload their own images, define taxonomic label sets, randomly create sample points, and use the built-in graphical user interface (UI) to classify the points, or the points can be classified by neural networks which are automated trained by CoralNet. Annotations created in

other software can be uploaded easily. Using these images, their point sample locations, and the points' expert-assigned labels, CoralNet trains unique NNs for each user and dataset. Users can train individualized models relatively quickly and with a relatively small number of training samples, because CoralNet uses a selection of all annotations uploaded by all users to train a foundational model that provides a good starting point (this process is known as transfer learning).

Despite its apparent utility, CoralNet is not widely used by the global coral research community. Researchers consider a wide variety of factors when evaluating a new technology, including:

1. Time. Researchers are busy with short- and medium-term deadlines, so it can be difficult to find time to learn new methods that might expedite or improve work in the long-term but cost time now with little or no immediate payoff.
2. Money. Cameras, high-speed computers, and large volumes of digital storage are expensive, especially for researchers in less wealthy areas. Available money is often tied to specific and short-term demands, leaving little to experiment with new methods. CoralNet alleviates some storage and computational processing demands.
3. Difficulty/frustration level. Learning new technology is often frustrating. This is especially true if the new technology is difficult or unintuitive to use, as software so often is. Using high-end cameras properly can also be tricky, especially with the added complication of a waterproof housing and the low, strange ambient light underwater.
4. Perceived value. Researchers generally operate on something approximating a cost/benefit approach: if they are confident the data made available by some new technology or methodology will surpass current capabilities, they are likely to invest the time and money to experiment with the new technology and potentially incorporate it into their repertoire.

If the technology asks substantial time and money but is unlikely to lead to much by way of novel data, researchers are unlikely to pursue it.

In this study, two experiments were performed to test the utility of CoralNet. The CoralNet UI was used to point sample thousands of images from field expeditions to the Maldives, Micronesia, and the Caribbean. This report aims to provide useful information to potential CoralNet users by answering A) how accurate CoralNet predictions are as a function of training set size, and B) how much time (if any) is saved by incorporating AI predictions in CoralNet's UI.

4.2 METHODS

The imagery for this report was originally collected for a variety of purposes over many years, so each of the three region datasets (Maldives, Micronesia, and Caribbean) is comprised of multiple field expedition datasets from different times and places. Regional datasets were sorted into two groups: the first was previously point sampled by coral experts and used to test and train CoralNet NNs, while the second had not been point sampled before this experiment and was used to test whether CoralNet AI assistance expedites point sampling. See Table 1 for a complete list of the test/train datasets, and Table 2 for a list of the datasets used to test the time impact of CoralNet-assisted point classification. Time impact data is not yet available.

The dataset for each region was comprised of 6000 images. Once collected, images were point sampled with the free software ImageJ (Schneider et al., 2012). Using a stratified random sampling strategy, between 5 and 50 points were randomly placed on pixels in each image, and then the organism containing those pixels was taxonomically classified. Information (image filenames, point locations, associated labels) was then compiled in Python, saved in a CSV format compatible with CoralNet, and then CSVs and corresponding images were uploaded to CoralNet.

Supplemental Information Tables 4.S1, 4.S2, and 4.S3 show the sets of possible taxonomic labels for each dataset (hereafter referred to as labels, classifications, classes, or categories), as well as the number of points in each dataset that were assigned that taxonomic label.

The procedure to upload data to CoralNet is straightforward and well-documented on the CoralNet website. A full technical description of CoralNet can be found in (Beijbom et al., 2012; Chen et al., 2021). Briefly: the first step is to create a CoralNet "source", which is an image dataset. There is a simple UI with a button researchers can press to upload images. Label sets can be uploaded as a CSV, or created on the website with a UI point and click tool. Label sets are flexible—researchers can select specific (genus or species) or generic (morphology, functional group) labels. There are also upload buttons for pre-existing point locations and associated classifications in CSV or other formats. Creating a compatible CSV requires only four columns, with the headings "Name", "Row", "Column", and "Label". These columns are populated by image filenames, vertical point indices, horizontal point indices, and point taxonomic classifications respectively. Labels must be converted to shortcodes provided by CoralNet when researchers determine a label set. CoralNet will begin training AI models once point annotations are uploaded or manually generated. It does so by cutting out a pixel square around each point and using the image patches with their associated classifications to train an image classification NN.

Once a CoralNet NN model has been trained with point sampled images, it generates automated predictions on uploaded images that remain un-point-sampled/classified. In a convenient UI, users can view the point locations and associated top automated predictions, keep the top prediction, or override them on a point-by-point basis.

The only parameter researchers must determine themselves is known as the confidence threshold. When NNs make classification predictions, they do so by independently predicting a score for each possible taxonomic classification (possible labels are referred to in the ML literature as "classes"). Low scores indicate the NN is confident a class label is not correct, high scores indicate the NN is confident a class label is correct, and scores in the middle suggest uncertainty. The final predicted classification is the class label with the highest score. By looking at the individual scores for each class, it is possible to estimate how confident the NN was: if the highest-scoring class scored very highly, and much higher than other classes, the prediction was confident, while if the highest-scoring class had a low score or a score only slightly above other classes, the NN was not confident.

CoralNet allows users to pick a number between 0 and 100 that limits automated predictions to only those that were confident. If a confidence threshold of 0 is selected, then all points will be automatically labeled, because all predictions have a confidence score above 0, while if 100 is selected, no points will be automatically labeled because NNs very rarely give 100% confident scores. CoralNet provides an easy-to-use UI tool to help users select a desirable confidence threshold value: an interactive plot shows, for a chosen confidence threshold between 0 and 100, what percentage of points will be labeled (i.e., what % of point classification predictions are above a certain confidence score) and how accurate those predictions are expected to be. Users can decide to use a low threshold, which results in more points being automatically labeled but with a relatively low expected accuracy, or a high threshold, which results in more accurate but fewer predictions.

To calculate accuracy, CoralNet reserves a small percentage of uploaded images during training (i.e., they are not used in training), so they can be used for evaluation. However, these

evaluation accuracy metrics do not necessarily reflect what a user will experience when using CoralNet to predict classifications on new datasets—if a user is generating predictions for images that are significantly different than those used to train the AI model (different lighting, geographic location, camera settings, etc), they are likely to be less accurate than predictions on evaluation images that came from the same dataset used for training. As users label new images, the NN will continually retrain, so it will become more accurate on new datasets in time.

To evaluate CoralNet's utility for coral researchers, two experiments were conducted. Experiment 1 investigated how NN accuracy changed as the number of training images/points increased incrementally, in order to help researchers understand how many images need to be labeled manually in order to generate a desired NN accuracy. Images for the three dataset locations were uploaded incrementally: 50 images were upload first and the NN was trained and evaluated, then 50 more for a total of 100, then enough to achieve totals of 200, 500, 1000, 2000, and 5000 training images. Accuracy was tested in two different ways: internally-generated CoralNet evaluation accuracy is reported (images used for this are hereafter referred to as the validation set), and a separate dataset of 1000 images from each of the three locations was reserved for additional testing (hereafter referred to as the test set). This is important because the evaluation set is a subset of the uploaded training data, so its size and constituent images changed as the number of uploaded training images increased.

Experiment 2 investigated how much time, if any, is saved by using automated CoralNet predictions. Confidence intervals were selected for each region that resulted in 75% of points being labeled automatically. Labelers were encouraged to review automatic predictions as well as classify the remaining points. Estimated accuracy for these predictions varied by site: 83% for the Caribbean, 89% for Maldives, and 85% for Micronesia. One coral expert per region point sampled

200 images without automated assistance and recorded the time required, and then repeated the process on 200 more images with CoralNet NN assistance.

4.3 RESULTS AND DISCUSSION

4.3.1 EXPERIMENT 1

To evaluate how CoralNet accuracy was affected by training dataset size, we uploaded our datasets sequentially – first just 50 images, then 100, 200, 500, 1000, 2000, and 5000. At each stage, average accuracy of the entire NN was recorded on the evaluation and test datasets, as well as prediction accuracy on individual classes. Evaluation set accuracy was provided directly by the CoralNet UI—no programming is necessary—but test set accuracy was determined with Python. For this investigation, we used the standard definition of prediction accuracy: the number of correctly-predicted points divided by the total number of validation/test points.

Selected results of Experiment 1 are depicted in Figures 1 through 6. Complete results for the validation sets are shown in confusion matrices in Tables 4.S4, 4.S5, and 4.S6. Figures 1-6 show accuracy on both the validation set (internal to CoralNet) and the test set (same 1000 images for all tests), and they show these accuracies in terms of both the number of uploaded images and the number of uploaded points. This distinction between images and points is important: some taxonomic classes are much more common than others, so, for a given number of uploaded images, the training data available for each class varies widely (see Figure 7 for an illustration, and Tables 4.S1, 4.S2, and 4.S3).

Figures 1 and 2 show accuracy of CoralNet predictions on the Maldives dataset by uploaded images (a, c) and points (b, d). Subfigures a) and c) have linear x-axes, while subfigures b) and d) have logarithmic x-axes. Figures 3 and 4 show results from the Micronesia

dataset organized in the same way, and Figures 5 and 6 show results from the Caribbean dataset. All plots show overall accuracy on all classes, as well as accuracy for some common hard corals. Accuracy for all classes can be found in Tables 4.S4, 4.S5. Rows show the expert-assigned label, while columns show CoralNet predictions. For any given cell, the value shown is the percentage of points assigned that row's label by an expert that were predicted to be the column's label by CoralNet. The percentage of correct predictions (accuracy) can be found along the diagonal. The data illustrated in Figures 1-6 show a few important characteristics and trends:

1) Incidence rates of taxonomic classes vary dramatically, and most make up only a tiny fraction of the total points (see Figure 7, which shows the distribution of label frequency). A full list of classes and their incidence rates in training, validation, and test datasets can be found in the Supplemental Material. Accuracy of fully-trained NN predictions on all classes can also be found in the Supplemental Material.
2) CoralNet performance is inconsistent with small datasets, as the NNs have few (and not necessarily representative) training examples to draw from, and the validation datasets are also small (and not necessarily representative). The test datasets do not vary in size as training data is incrementally increased, but validation datasets do.
3) As the number of training images increases, linear x-axis Figures suggest performance saturates quickly.
4) However, logarithmic Figures indicate otherwise. Overall performance continues to improve even with hundreds of thousands of training points.
5) Individual hard coral classes show a similar pattern – they appear to generally saturate relatively quickly on linear axes, but logarithmic axes show they generally continue to improve as the number of training points increases. Figures 1-4 indicate saturation may

eventually occur beyond 1000 training points per class, but Figures 5 and 6 do not mirror this trend.

6) It is difficult to determine from these data if class prediction accuracy will, with increasing training data, saturate at similar levels, or if some classes (likely those most visually distinct, to both computer and human expert) will indefinitely remain more accurate than others.

It is important to note that coral NN performance varies widely with the "difficulty" of the assigned task (Runyan et al., 2022). All the experiments in this report were conducted with very large label sets (100+ classes), many of which are visually quite similar, making them prone to error in both NN prediction and ground truth training/validation/test sets prepared by experts. This is particularly true of algal genera/species (see Supplemental Material). NNs tasked with much simpler distinctions (e.g., coral v algae v. inorganic/other, or a single genus v. everything else) achieve equivalent or higher accuracy with much less training data, as the task requires making far fewer distinctions. This is illustrated in Figure 8, which shows a CoralNet validation accuracy plot—performance on functional groups (hard coral, algae, invertebrates, etc) is significantly better than with the full, detailed label set. Figure 8 also shows how accuracy increases as the confidence threshold increases (i.e., as the less-confident predictions are removed from evaluation)

Additional "difficulty" is introduced by variation in image collection and expert ground truth point sampling—camera equipment, lighting conditions, white balancing/color correcting, taxonomic expertise and opinions, and much more vary by individual operator, lab, and field expedition. A NN trained and tested on data widely-varying in these characteristics will generally

require more training data to achieve the same accuracy as a NN trained/tested with narrow, consistent data—if equivalent performance can be achieved at all (Runyan et al., 2022).

Finally, it is important to emphasize that accuracy and consistency of expert ground truth point labeling play a major role in NN predictions, both in achieved accuracy and ability to measure it. Learning from imperfect labeling is possible, but generally leads to less accurate predictions, as the patterns NNs learn to make distinctions with during training are not consistent (Song et al., 2022). Additionally, imperfect ground truth affects accuracy determined by evaluation/test datasets: a NN that is 100% accurate evaluated with labels that are 90% accurate will be reported as 90% accurate.

No two coral experts will produce 100% identical classifications for a large number of points. Studies internally and published (Beijbom et al., 2015) show disagreement rates of 5-25+%, again depending on the difficulty/specificity of the labeling task. No completely reliable ground truth is available for classifications used in this study, as tissue samples or other ~100% reliable taxonomic identification methods were not part of the image survey procedures—classifications considered correct were produced by fallible experts.

4.3.2 EXPERIMENT 2

To evaluate how much time can be saved by using CoralNet's AI predictions, we selected three participants for this study from within our lab. They are all coral experts with advanced degrees and specific familiarity with their assigned regions: Katie Lubarsky labeled images from the Maldives, Clinton B. Edwards labeled images from Micronesia, and Nathanial Hanna Holloway labeled images from the Caribbean. Each participant was given a regional CoralNet source trained with the 5000 image datasets also used for Experiment 1. Each participant used the

CoralNet annotation tool to label 50 points each on 200 images with all AI assistance functions turned off, recorded how much time it took, and then repeated the procedure on 200 different images, this time with AI predictions assistance. The 200 images per region used for this experiment were not part of the NN training datasets.

To perform the component of this experiment incorporating CoralNet AI assistance, we adjusted the confidence threshold of CoralNet for each region/source to the level at which 75% of points were labeled automatically. This means that the 75% for which AI predictions were most confident were labeled automatically, while the other 25% were not labeled by CoralNet. The experts were tasked with reviewing these automated predictions, and also labeling the remaining 25% of points. For all points, the experts were shown CoralNet's prediction of which taxonomic categories were most likely correct.

Results for Experiment 2 are shown in Table 3. Without CoralNet AI assistance, the three experts required an average of 4.61 seconds to assign a taxonomic category to a single point in an image, while introducing CoralNet AI predictions reduced the time required to 2.97 seconds—a reduction in time of 36%. As labeling sample points on images of coral reefs is slow and very demanding on the labelers, a 36% reduction of the time required to do so is very significant. Additionally, Experiment 1 showed that neural network accuracy was still increasing for most genera when the 5000 image training datasets were exhausted—this result implies that, as more training data becomes available, the CoralNet AI predictions will improve, likely further increasing how much time using them saves, as fewer will need to be corrected and suggestions to the experts will be more reliable.

4.3.3 USING CORALNET PREDICTIONS WITHOUT EXPERT VERIFICATION

These results suggest an important question: Why can't we let neural networks annotate image survey products without editing or verification? After all, editing predictions takes orders of magnitude more time than generating them, so NN tools that do not require post-prediction editing are vastly more useful than those that merely reduce expert editing time.

This question has been investigated in the context of CoralNet—experiments have shown that CoralNet can produce accurate percent cover data without expert editing and verification of its predictions. If researchers only need percent cover data, CoralNet can be trained to provide it (Williams et al., 2019). However, Williams et al. (2019) only investigated bulk coral cover (the percentage of the benthos covered by all coral taxonomies), as well as a few common coral genera. Our results from Experiment 1 show that CoralNet is much less reliable when predicting rarer genera/species, at least with the datasets we had available. Rare taxonomic categories present a difficult problem: because they are so rare, it is likely that any single rare taxonomy is the correct label for only a small number of points, even in large datasets. Incorrect predictions of only a few points will therefore dramatically alter the relative area assigned to a given rare taxonomy. For example, if CoralNet predicts 66% of the benthos to be occupied by all coral, and the correct coral cover percentage was 67%, it can be considered quite accurate for that task. However, if a rare taxonomy composing 0.05% of total area cover is predicted to cover 0.1%, the total error of 0.1% is a big problem, as it represents a 100% relative error for that single taxonomy. Therefore, CoralNet predictions might be reliable for bulk metrics like total coral cover and therefore useful without expert verification for some applications, but verification is required for investigating all but the most common genera.

If coral cover metrics for all corals or specific genera/species are generated at a site, changes in time can be investigated by revisiting the site, conducting another image survey, and

point sampling those new images. However, studying trends in time often requires a number of time points for statistical confidence, and long time series of coral reef image surveys are uncommon, both because the image survey methodology is relatively new and because point sampling the resulting images is very time-consuming. Therefore, point sample labels need to be as accurate as possible, in order to confidently reveal trends as quickly as possible and examine dynamics between individual time points. Identifying trends with data containing the error/uncertainty associated with unverified AI predictions would require longer time series to identify trends, and also reduce the significance of changes between any two time points. Additionally, performance of the NN may change across time—a problem not identifiable unless some predictions are verified.

Tracking coral at a single site across a few points in time might require very accurate measurements for maximal utility, but what about tracking the growth of all coral or a specific genus across a large area such as a region or the entire world? Tens of thousands, if not many more, images would be available for such an experiment, providing much more statistical information than short time series. If NN predictions were generated on those thousands+ images, it is conceivable that trends could still be identified even if the predictions were only 70-90% accurate. However, this is not true if individual predictions are biased by an unknown degree, rather than exhibiting consistent variation around a mean. To determine if the NN predictions are biased, one has to test how the NN behaves at different sites, with different lighting conditions, and more. That testing requires expert-generated ground truths.

We find editing predictions to high accuracy provides the most useful form of data for investigating coral ecology. In the future, larger training datasets might make it possible to train a NN that is relatively unbiased on most possible inputs, or at least its bias can be characterized with

reasonable confidence. Such a neural network could be used to investigate trends across many 3D models, even if it is not perfectly accurate on a per-coral basis, because the bias in its predictions can be accounted for.

TABLES

Table 4.1: Location and date of dataset collection (experiment 1).

Region	Site	Year(s) collected
Maldives	Maldives	2020, 2021
Micronesia	Ant	2014
	Pohnpei	2018
	Palau	2017
Caribbean	Curacao	2015
	Montserrat	2015
	Windward Islands	2016
	Antigua	2017

Table 4.2: Location and date of test dataset collection (experiment 2).

Region	Site	Year(s) collected	Expert labeler
Maldives	Maldives	2022	Katherine Lubarsky
Micronesia	Palau	2022	Clinton Edwards
Caribbean	Windward Islands	2018	Nathaniel Holloway

Table 4.3: Time required to label a point with and without CoralNet predictions.

Time without CoralNet (seconds per point)	Time with CoralNet (seconds per point)	Percent difference (%)
4.61	2.97	36%

FIGURES

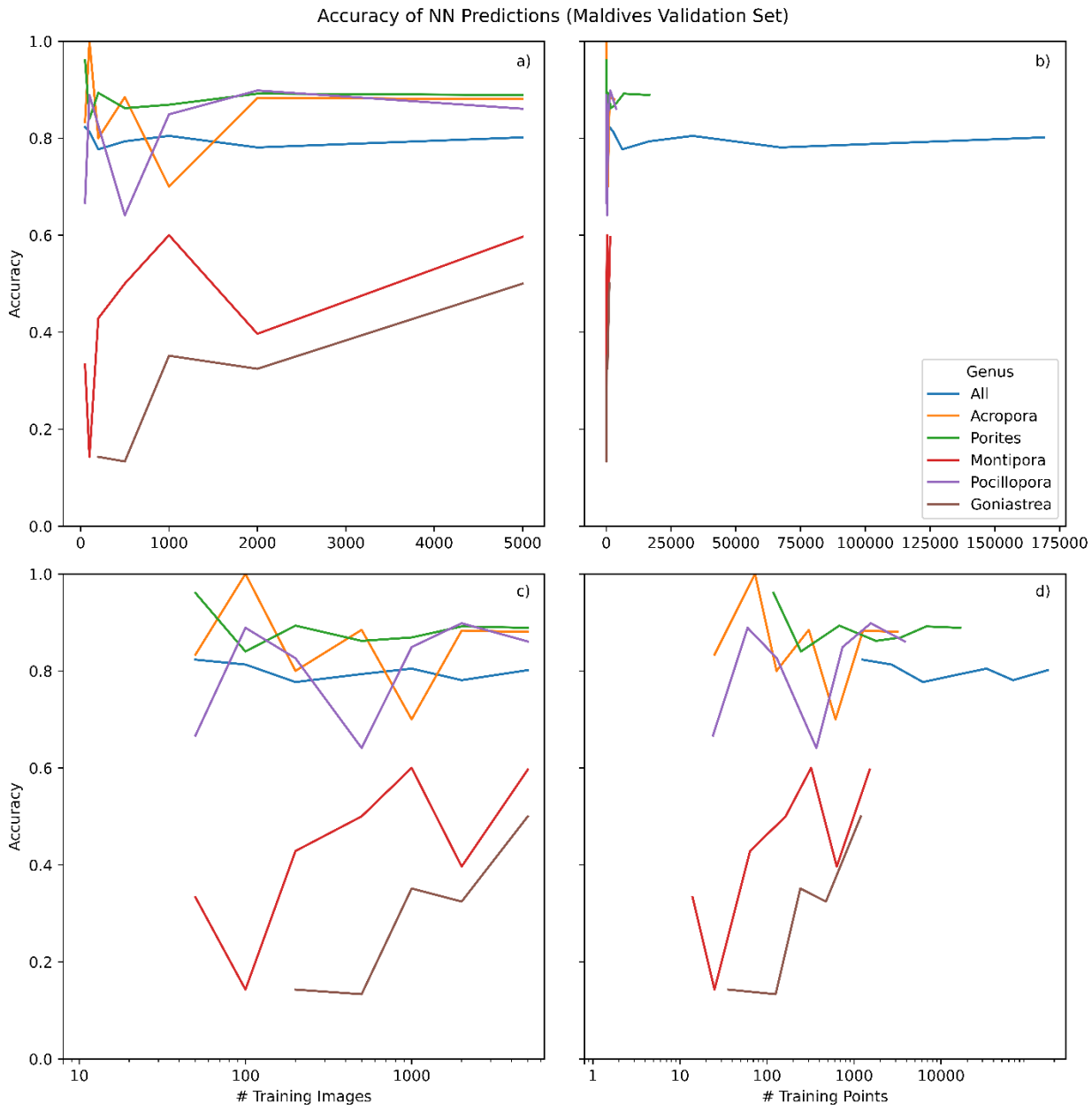


Figure 4.1: Accuracy of NN predictions (Maldives validation set), by number of training images (a, c) and number of training points (b, d).

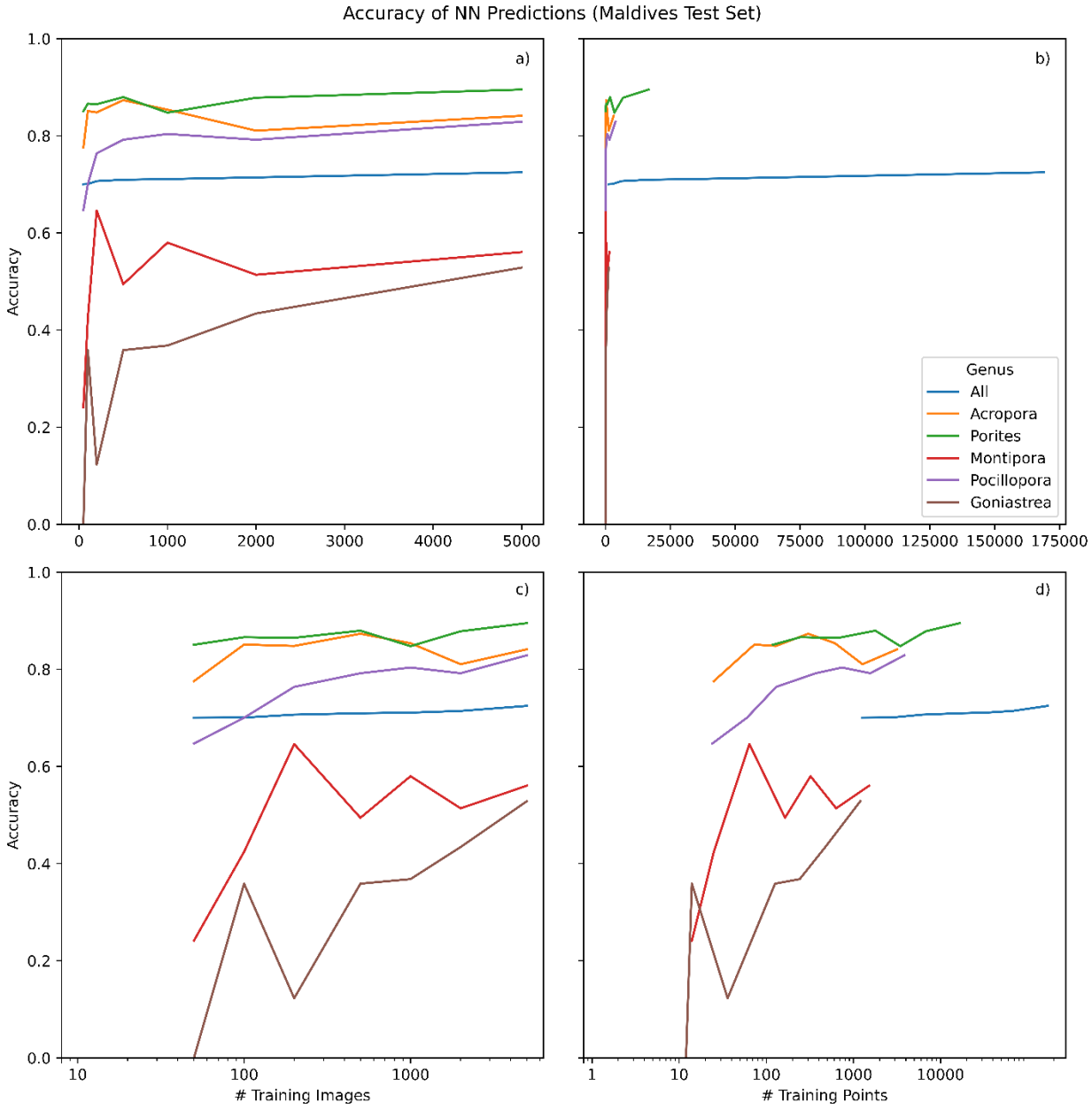


Figure 4.2: Accuracy of NN predictions (Maldives test set), by number of training images (a, c) and number of training points (b, d).

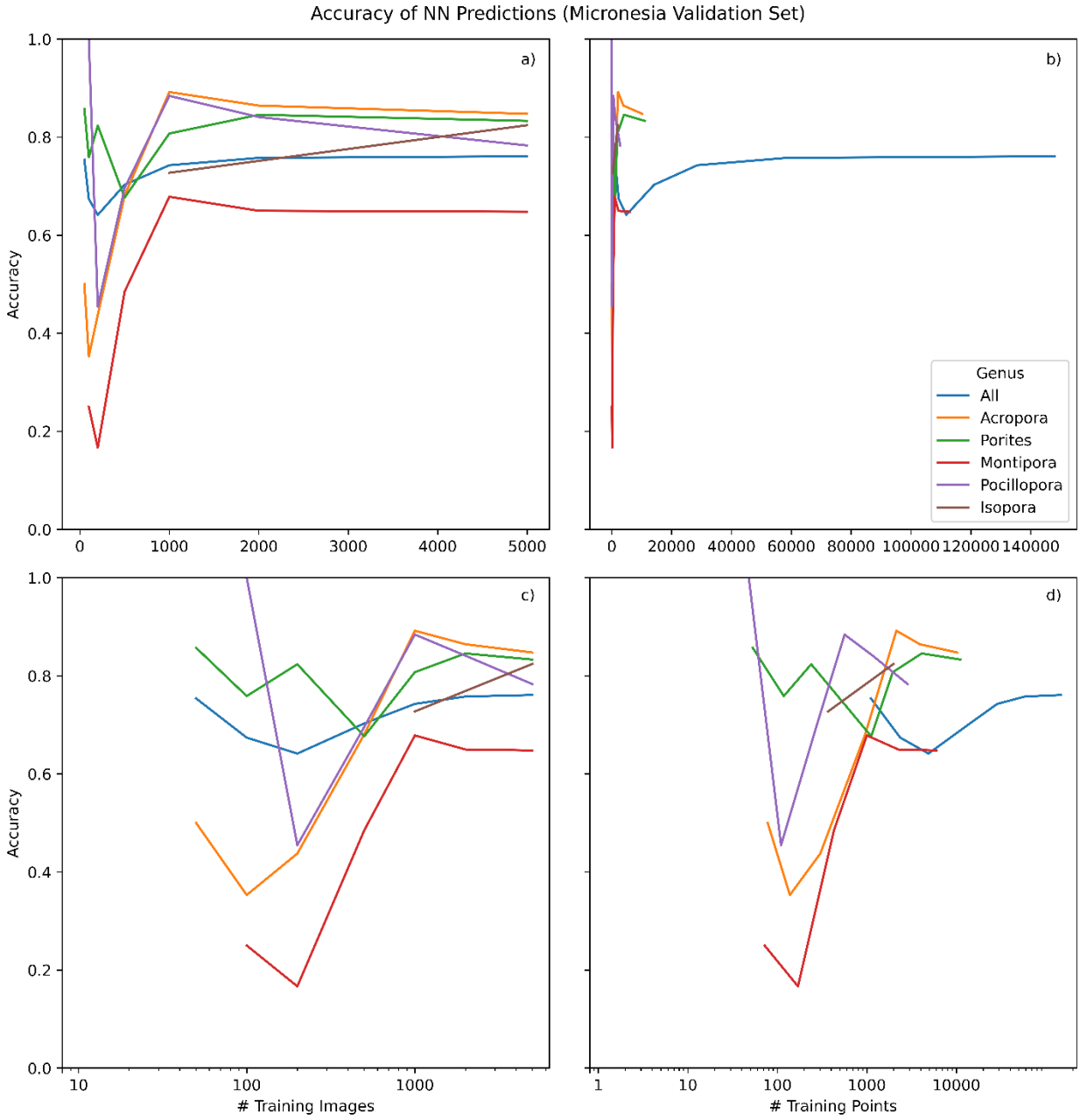


Figure 4.3: Accuracy of NN predictions (Micronesia validation set), by number of training images (a, c) and number of training points (b, d).

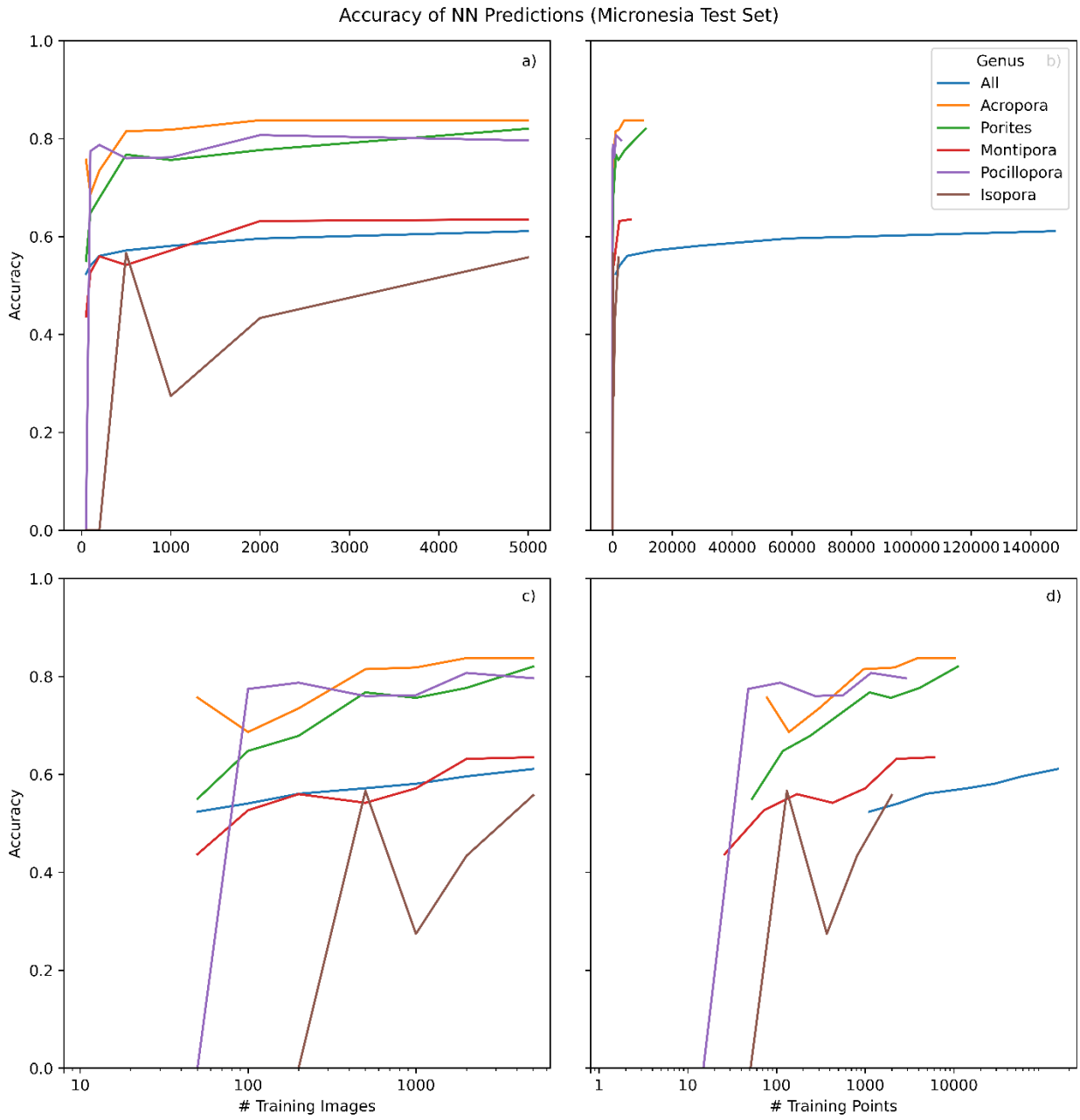


Figure 4.4: Accuracy of NN predictions (Micronesia test set), by number of training images (a, c) and number of training points (b, d).

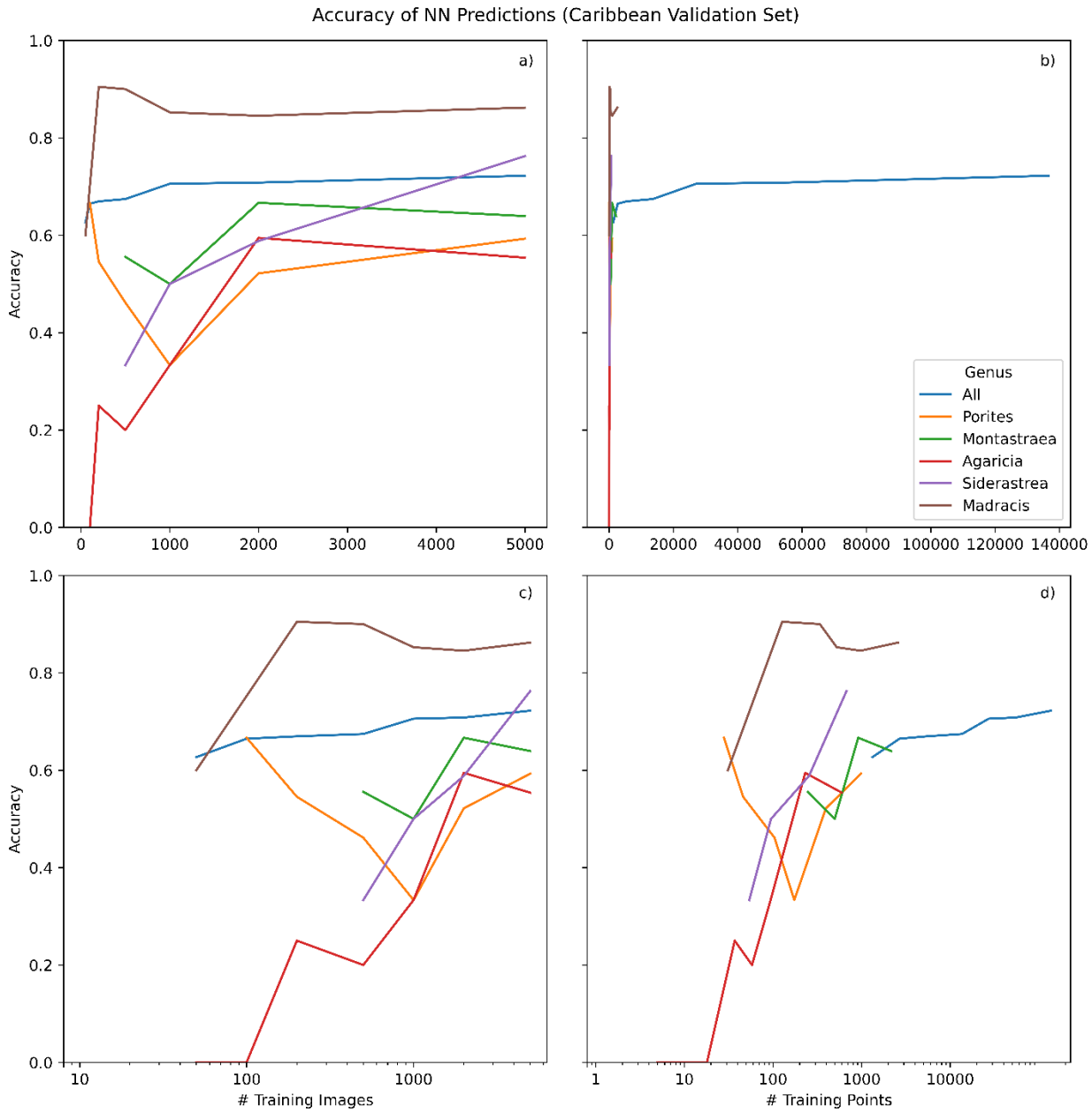


Figure 4.5: Accuracy of NN predictions (Caribbean validation set), by number of training images (a, c) and number of training points (b, d).

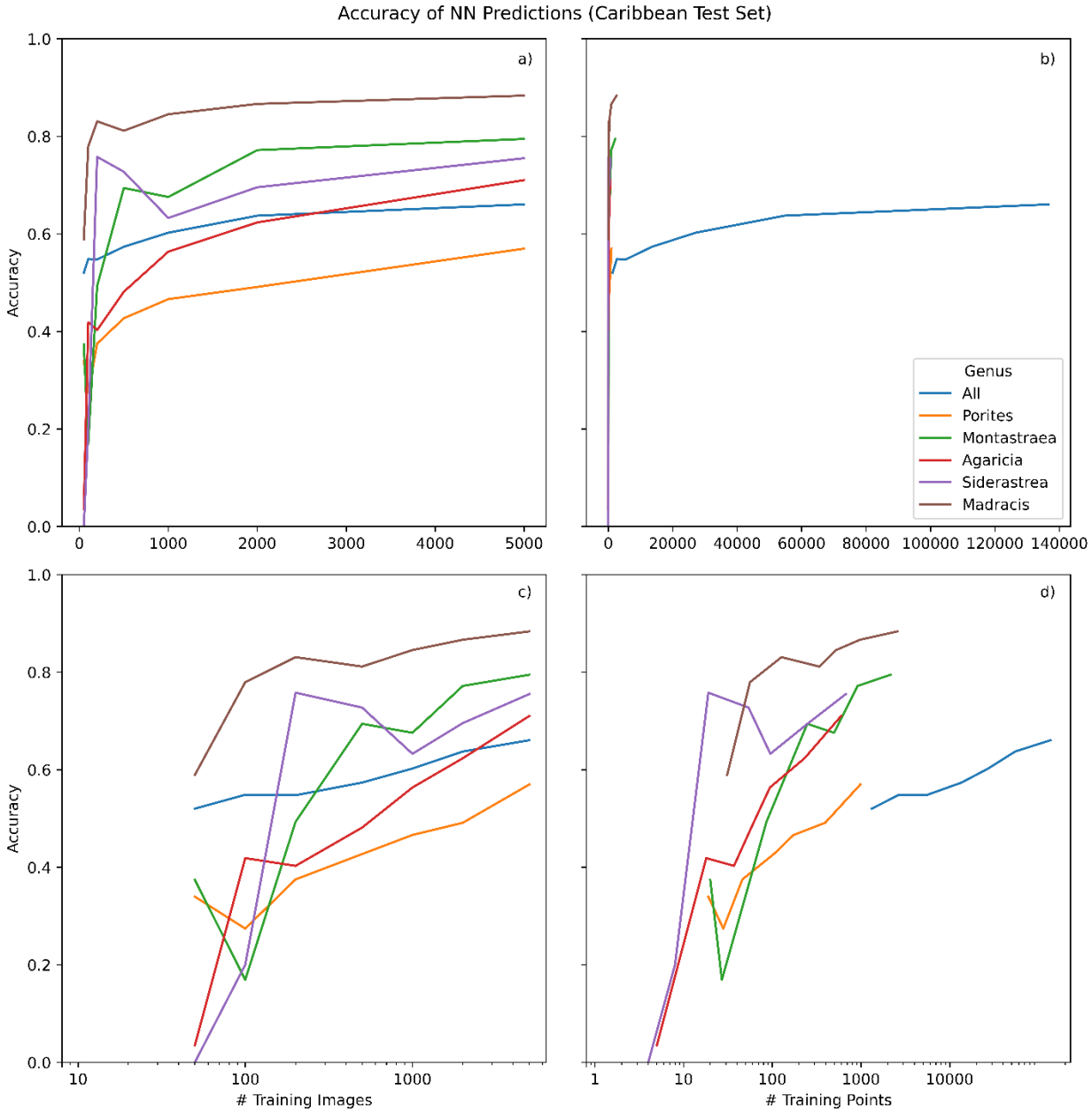


Figure 4.6: Accuracy of NN predictions (Caribbean test set), by number of training images (a, c) and number of training points (b, d).

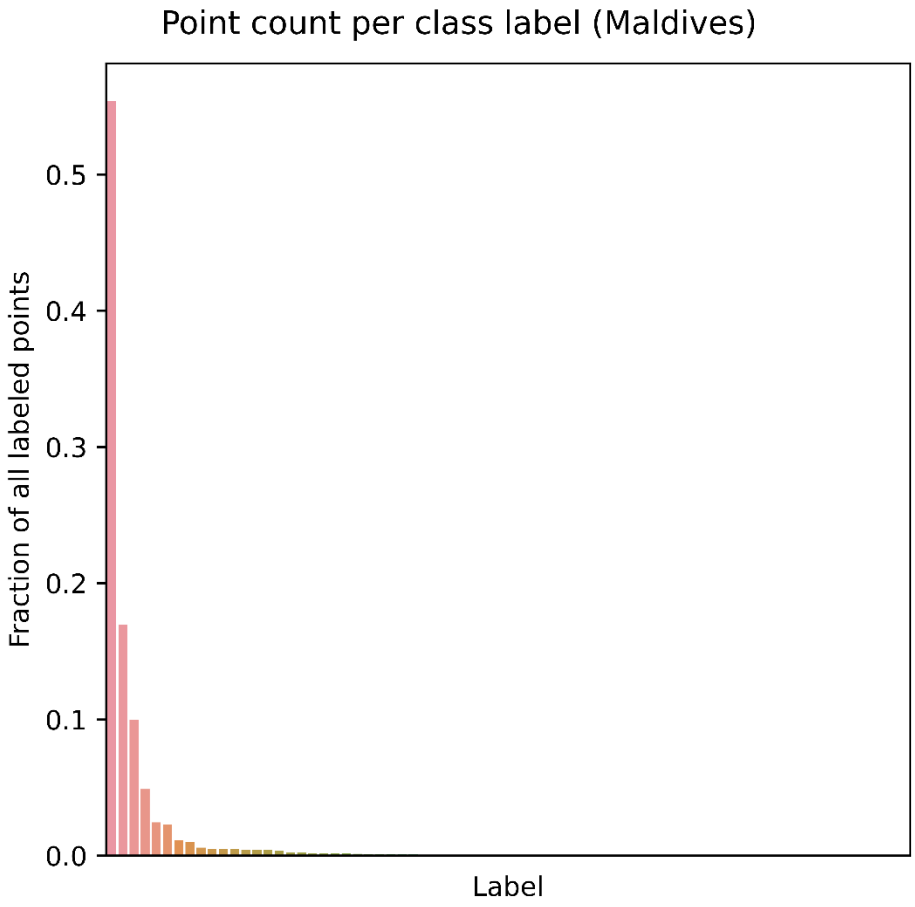


Figure 4.7: Bar plot showing fraction of total points by label for the Maldives dataset. The first two columns are Turf and CCA respectively. Full data for this chart can be found in Supplemental Information Table 4.S1.

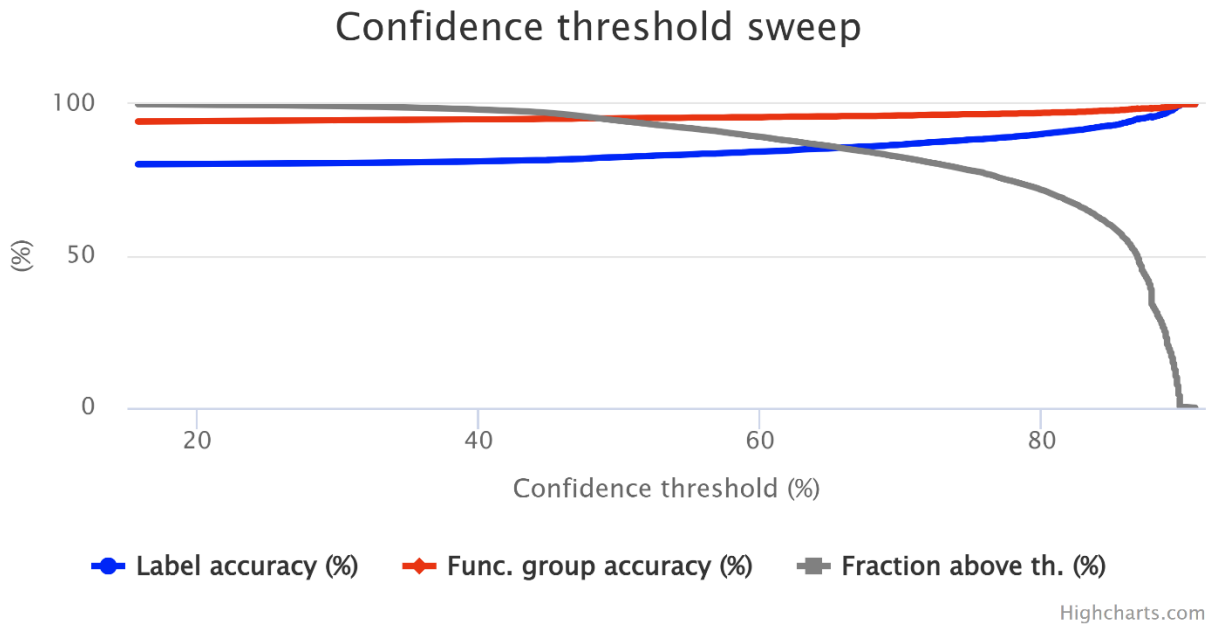


Figure 4.8: Screenshot of CoralNet website showing accuracy by total label set and the more generic functional group label set.

REFERENCES

Beijbom, O., Edmunds, P. J., Kline, D. I., Mitchell, B. G., and Kriegman, D. (2012). Automated annotation of coral reef survey images. in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 1170–1177. doi: 10.1109/CVPR.2012.6247798.

Beijbom, O., Edmunds, P. J., Roelfsema, C., Smith, J., Kline, D. I., Neal, B. P., Dunlap, V. M., Fan, T., Tan, C., Chan, S., Treibitz, T., Gamst, A., Mitchell, G., Kriegman, D. (2015). Towards Automated Annotation of Benthic Survey Images: Variability of Human Experts and Operational Modes of Automation. *PLOS ONE* 10, e0130312. doi: 10.1371/journal.pone.0130312.

Q. Chen, O. Beijbom, S. Chan, J. Bouwmeester and D. Kriegman, "A New Deep Learning Engine for CoralNet," 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2021, pp. 3686-3695, doi: 10.1109/ICCVW54120.2021.00412.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, eds. F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc.), 1097–1105. Available at: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

Runyan, H., Petrovic, V., Edwards, C., Pedersen, N., Alcantar, E., Kuester, F., Sandin, S. A. (2022). Automated 2D, 2.5D, and 3D Segmentation of Coral Reef Pointclouds and Orthoprojections. *Frontiers in Robotics and AI* 9. doi: 10.3389/frobt.2022.884317.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9, 671–675. doi: 10.1038/nmeth.2089.

Song, H., Kim, M., Park, D., Shin, Y., and Lee, J.-G. (2022). Learning From Noisy Labels With Deep Neural Networks: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, 1–19. doi: 10.1109/TNNLS.2022.3152527.

Supplemental Material

Table 4.S1: Complete label set from Maldives experiment and point label incidence number

Label	N
Turf	66820
CCA	20475
Porites	11999
Peyssonnelia	5877
Acropora	2944
Pocillopora	2713
Montipora	1360
Sand	1216
Favites	689
Halimeda	602
Leptastrea	582
Pavona	562
Zoanthid	517
Leather Coral	493
Goniastrea	483
Astreopora	420
Psammocora	294
Sponge	261
Black Fungus Disease (BFD)	222
Tunicate	194
Cyanobacteria	187
Galaxea	175
Hydnophora	149
Platygyra	139
Heliopora	133
Dipsastraea	128
Lobophora variegata	119
Fungia	97
Cyphastrea	90
Echinopora	68
Tydemania spp.	66
Astrea	58
Other hard coral	56
Symphyllia	46
Goniopora	31
Echinophyllia	25
Hydrozoa	21
Acanthastrea	20

Label	N
Turbinaria (algae)	19
Leptoseris	19
Other invertebrate	16
Encrusting green algae	16
Stylophora	15
Coscinaraea	15
Limestone stable substrate	15
Gorgoniidae	15
Pachyseris	14
Lobophyllia	14
Millepora	12
Pachyclavularia	11
Anemone	9
Tridacna: Giant clam	8
Chlorodesmis	8
Leptoria	7
Merulina	7
Green fleshy algae	7
Gardineroseris	6
Encrusting brown algae	6
Diploastrea	4
Other Algae	4
Avrainvillea	4
Dictyota	4
Padina	4
Caulerpa	4
Green diplosoma	4
Herpolitha	3
Sandolitha	3
Red fleshy algae	3
Seriatopora	3
Palmophyllum crassum	3
Halomitra	3
Porites superfusa	2
	2

Table 4.S2: Complete label set from Micronesia experiment and point label incidence number

Label	N
CCA	44040
Halimeda	37510
Turf	27114
Porites	13226
Acropora	12671
Peyssonnelia	8502
Montipora	7419
Pocillopora	3529
Heliopora	2782
Isopora	2586
Pavona	1924
Black crust	1881
Lobophora variegata	1633
Dipsastraea	1619
Microdictyon	1545
Pachyclavularia	1396
Rhipilia spp.	1154
Sand	1005
Millepora	916
Echinopora	896
Caulerpa	881
Sponge	825
Platygyra	761
Red fleshy algae	552
Merulina	537
Hydnophora	509
Favites	496
Porites superfusa	484
Leather Coral	478
Leptastrea	475
Stylophora	452
Goniastrea	350
Psammocora	264
Galaxea	229
Fungia	217
Cyanobacteria	214
Seriatopora	204
Pachyseris	200
Astreopora	186

Label	N
Cyphastrea	145
Diploastrea	130
Goniopora	130
Zoanthid	124
Tunicate	112
Leptoseris	107
Green diplosoma	82
Astrea	82
Leptoria	71
Tridacna: Giant clam	66
Gardineroseris	62
Anemone	42
Turbinaria (algae)	40
Dictyosphaeria cavernosa	39
Symphyllia	35
Echinophyllia	31
Lobophyllia	30
Physogyra	27
Gorgoniidae	27
Dictyota	26
Avrainvillea	25
Corallimorpharia	24
Dictyosphaeria versluysii	22
Mycedium	13
Green fleshy algae	12
Coscinaraea	12
Hydrozoa	11
Tubastrea	10
Limestone stable substrate	10
Lace Coral	9
Ctenactis	9
Acanthastrea	8
Stylocoeniella	8
Other	7
Padina	6

Table 4.S3: Complete label set from Caribbean experiment and point label incidence number

Label	N
Turf	145519
Sand	110274
Cyanobacteria	30554
Dictyota	28253
CCA	10200
Seagrass	8328
Gorgoniidae	7656
Madracis mirabilis	6887
Sponge	6713
Encrusting sponge	5166
Lobophora variegata	3443
Peyssonnelia	3441
Montastraea cavernosa	2684
Porites astreoides	2561
Plexaura	2487
Millepora	2395
Halimeda copiosa	2370
Pseudopterogorgia	2182
Siderastrea	2026
Sargassum hystrix	1756
Diploria strigosa	1613
Agaricia	1608
Montastraea faveolata	1599
Montastraea annularis	1551
Pseudoplexaura	1340
Branching Gorgonian	1295
Colpophylia natans	1195
Palythoa	1169
Millepora complanata	953
Amphiroa	952
Eunicea	832
Montastraea franksi	828
Meandrina meandrites	552
Erythropodium	459
Sargassum	421
Sargassum fluitans	381
Porites	378
Calcium Carbonate	356
Padina boergesenii	331

Label	N
Madracis decactis	273
Other invertebrate	225
Diploria labyrinthiformis	222
Chondrilla	197
Stephanocoenia intersepta	193
Diploria clivosa	179
Green fleshy algae	178
Acropora palmata	156
Muricea	143
Stable substrate	137
Turbinaria (algae)	135
Clionid	120
Zoanthid	102
Dendrogyra cylindrus	96
Red fleshy algae	88
Dichocoenia stokesi	86
Acropora cervicornis	84
Antipatharia	75
Eusmilia fastigiata	71
Udotea	59
Other	48
Agaricia lamarcki	47
Galaxaura	44
Hydrozoa	44
Muriceopsis	43
Agaricia humilis	37
Siderastrea radians	28
Wrangelia penicillata	27
Briareum	25
Tridemnum	25
Madracis	21
Plexaurella	18
Montastraea	17
Diploria	15
Stypopodium	15
Padina jamaicensis	12
Neomeris	10
Dictyopteris	9
Rhipocephalus phoenix	9

Label	N
Ventricaria ventricosa	8
Dipsastraea	8
Isophyllia rigida	8
Branching algae	7
Dictyosphaeria cavernosa	7
Liagora	7
Madracis senaria	6
Isophyllia sinuosa	6
Stylaster roseus	4
Schizothrix	4
Colpophyllia breviserialis	4
Scolymia lacera	4
Tubastrea	4
Madracis pharensis	4
Mycetophyllia ferox	4
Millepora squarrosa	4
Caulerpa serrulata	4
Agaricia grahamae	3
Orbicella	2
Scolymia cubensis	2
Porites branneri	2
Mussa angulosa	2
Acropora prolifera	2
Porites divaricata	2

Table 4.S4: Confusion matrix of Maldives validation set. Numbers are percentages. Rows show the expert-assigned label, while columns show CoralNet predictions. The percentage of correct predictions can be found along the diagonal.

	Turf	CCA	Porites	Peyssonnelia	Pocillopora	Sand	Acropora	Lobophora	Montipora	Halimeda	Echinopora	Goniastrea	Sick CCA	Brown Macroalgae	Favites	Leptastrea	Pavona	Sponge	Astreopora	Psammocora	Zoanthid	Heliopora	Leather Coral	Platygyra	Galaxea	Tunicate	Cyanobacteria	Fungia	Black Fungus Disease	Tydemania	Astrea	Cyphastrea	Hydnophora	Dipsastraea	Millepora	Leptoseris	Symphyllia	Favia	Other Hard Coral	Green Macroalgae	Gardineroseris	Goniopora	Turbinaria (algae)	Other invertebrate	Caulerpa	Stylophora	Hydrozoa	Coscinaraea	Echinophyllia	Alveopora	Other
Turf	93	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CCA	33	63	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porites	7	1	88	1	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Peyssonnelia	37	11	1	46	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pocillopora	8	1	2	0	87	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sand	46	0	1	0	0	52	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acropora	6	2	1	0	5	0	84	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Lobophora	44	4	0	14	0	0	0	37	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Montipora	15	2	8	4	2	0	1	0	66	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Halimeda	17	3	0	2	0	0	1	0	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Echinopora	6	3	2	0	0	0	0	0	0	0	89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Goniastrea	17	3	8	1	0	0	0	0	0	0	0	57	0	0	2	2	2	0	0	2	1	0	0	1	0	0	0	1	0	0	0	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sick CCA	47	31	3	2	0	0	0	1	0	1	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Brown Macroalgae	55	5	1	0	0	0	0	0	0	0	0	0	0	33	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Favites	18	5	3	1	1	0	0	0	0	0	1	2	0	0	50	1	2	1	1	1	1	0	0	1	0	0	0	0	0	0	0	5	1	0	1	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0
Leptastrea	18	1	11	0	1	0	0	0	0	0	0	4	0	0	1	58	2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pavona	23	7	5	0	0	0	0	0	1	0	0	2	0	0	0	0	59	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sponge	33	12	6	11	0	0	0	0	1	0	0	1	0	0	0	0	1	33	0	0	0	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Astreopora	17	2	3	0	1	0	3	0	10	0	0	0	0	0	0	1	0	0	55	0	1	0	0	0	0	0	0	0	0	0	0	2	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Psammocora	19	2	10	0	2	0	0	0	1	0	0	2	0	0	0	2	4	2	0	52	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Zoanthid	19	1	1	1	0	0	0	0	1	0	0	3	0	0	1	1	0	0	0	1	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Heliopora	24	2	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	30	3	0	0	0	0	6	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
Leather Coral	14	0	20	0	5	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	58	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Platygyra	10	4	0	0	0	0	0	0	0	0	0	12	0	0	8	2	2	0	0	0	0	0	0	47	0	4	0	0	0	0	0	4	0	2	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0
Galaxea	25	5	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tunicate	39	11	0	3	3	0	0	0	3	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cyanobacteria	50	22	3	22	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fungia	9	9	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Black Fungus Disease	76	10	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tydemania	50	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Astrea	13	0	0	0	0	0	0	0	0	4	13	0	0	0	9	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	26	0	0	13	0	0	0	17	0	0	0	0	0	0	0	0	0	0	0	0
Cyphastrea	41	9	0	0	9	0	5	0	9	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	9	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hydnophora	14	5	0	0	5	0	0	0	10	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dipsastraea	27	0	7	7	0	0	0	0	0	0	0	13	0	0	20	0	0	0	0	0	7	0	0	0	0	0	0	7	0	0	0	7	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Millepora	7	33	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Leptoseris	0	29	14	0	0	0	0	0	0	0	43	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Symphyllia	15	0	0	0	38	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0
Favia	0	18	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0	0	0	0
Other Hard Coral	60	0	20	10	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Green Macroalgae	11	0	11	0	0	0	0	0	0	11	0	0	0	0	0	0	0	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0
Gardineroseris	0	0	0	0	0	0	0	0	0	0	0	67	0	0	0	11	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Goniopora	38	0	25	0	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0
Turbinaria (algae)	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Other invertebrate	20	80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Caulerpa	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stylophora	0	0	33	0	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hydrozoa	33	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0	0
Coscinaraea	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Echinophyllia	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4.S5: Confusion matrix of Micronesia validation set. Numbers are percentages. Rows show the expert-assigned label, while columns show CoralNet predictions. The percentage of correct predictions can be found along the diagonal.

	CCA	Halimeda	Turf	Porites	Acropora	Peyssonnelia	Montipora	Sand	Calcified Algae	Porites rus	Pocillopora	Isopora	Microdictyon	Heliopora	Pavona	Cyanobacteria	Dipsastraea	Black crust	Lobophora	Pachyclavularia	Rhipilia	Sponge	Echinopora	Platygyra	Caulerpa	Millepora	Merulina	Stylophora	Galaxea	Hydnophora	Pachyseris	Favites	Padina	Lobophora	Red algae	Goniastrea	Diploastrea	Dictyota	Goniopora	Porites superfusa	Leptastrea	Soft Coral	Fungia	Cyphastrea	Leather Coral	Psammocora	Seriatopora	Tunicate	Other hard coral	Astreopora	Other
CCA	81	3	10	0	1	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Halimeda	5	89	3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Turf	15	5	69	1	0	2	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porites	4	1	3	82	1	0	3	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Acropora	6	3	2	0	85	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Peyssonnelia	20	7	16	0	1	51	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Montipora	9	2	6	7	1	1	64	0	0	1	1	1	0	1	1	0	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sand	3	1	22	1	0	0	0	73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Calcified Algae	9	6	22	1	0	0	0	0	57	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porites rus	3	2	11	5	0	0	1	2	3	72	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pocillopora	5	1	3	0	5	0	1	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Isopora	4	1	5	1	5	0	4	0	0	0	2	76	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
Microdictyon	5	1	23	1	0	0	1	0	0	0	0	0	68	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Heliopora	3	3	6	8	0	1	1	0	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pavona	16	2	7	24	2	1	2	0	0	0	0	0	0	0	40	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Cyanobacteria	11	8	47	2	0	2	0	0	13	1	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dipsastraea	2	0	5	3	0	1	2	0	0	0	0	0	0	0	0	0	78	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Black crust	20	13	7	0	1	7	1	0	0	0	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lobophora	15	8	13	1	0	15	0	0	0	0	0	0	0	0	0	0	0	4	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pachyclavularia	2	1	9	3	2	0	2	0	0	0	2	0	0	0	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Rhipilia	5	7	16	3	0	2	1	0	0	0	0	0	0	1	0	0	0	0	0	0	62	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Sponge	14	5	14	6	0	2	6	0	5	1	1	0	0	1	0	2	1	0	0	0	1	41	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Echinopora	2	1	7	5	2	0	6	0	0	2	0	2	0	5	0	0	5	0	0	0	0	0	60	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	1
Platygyra	4	0	5	5	0	0	0	0	2	0	0	0	0	0	1	0	1	0	0	0	0	0	0	75	0	0	1	0	2	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Caulerpa	4	13	8	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	72	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Millepora	2	0	11	4	0	0	8	0	1	0	0	0	4	1	1	0	0	0	0	0	0	1	1	0	0	2	63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
Merulina	1	0	8	4	0	1	15	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	1	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
Stylophora	4	4	0	11	5	0	0	0	0	0	2	5	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
Galaxea	5	1	8	1	6	0	0	0	1	0	0	0	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	71	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
Hydnophora	9	0	10	7	7	0	1	0	3	0	0	1	0	6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pachyseris	0	0	3	3	0	6	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	82	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Favites	17	0	0	2	0	3	7	0	2	0	3	0	0	0	3	0	10	0	0	0	0	0	0	0	0	2	0	0	2	0	0	36	0	0	0	2	2	0	0	0	7	0	0	0	0	2	0	0	0	2
Padina	5	2	25	7	0	0	0	7	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	46	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Lobophora	0	2	12	0	0	0	2	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	65	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Red algae	0	25	43	0	0	7	0	0	0	0	0	0	0	2	0	4	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Goniastrea	2	0	5	7	0	0	4	0	2	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	64	0	0	0	0	4	0	0	0	0	4	0	0	0	2	0
Diploastrea	0	0	4	4	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	0	4	0	0	2	0	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0
Dictyota	0	6	43	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0	0	0	0	0	0	0	0	0	0
Goniopora	12	0	14	2	2	0	0	0	6	4	0	0	0	0	0	0	0	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	2	0	0	2	0	0
Porites superfusa	33	2	0	4	2	4	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	49	0	0	0	0	0	0	0	0	0	0
Leptastrea	10	0	7	0	0	5	12	0	0	0	0	0	0	0	17	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	0	0	0	0	0	0	0	0	2
Soft Coral	6	6	9	18	3	0	3	0	0	6	0	0	0	0	0	0	0	0	0	3	0	3	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	3	0	0	38	0	0	0	0	0	0	0	0	0
Fungia	3	0	13	3	0	3	3	0	0	3	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	0	0	0	0	0	0	0	3
Cyphastrea	12	0	24	8	4	0	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	4	0	0	0	4	0
Leather Coral	0	0	8	12	0	0	4	0	0	0	4	0	0	0	0	0	0	0	0	8	0	0	0	0	0	8	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	0	0	0	0	0	4
Psammocora	9	0	14	9	0	5	14	0	0	5	0	0	0	5	5	0	18	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	9	0	0	0	0	0
Seriatopora	0	15	5	0	15	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	0	0	0	10
Tunicate	16	0	37	11	0	0	0	0	0	0	0	0	0	5	0	0	0	11	0	0	5	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0
Other hard coral	11	17	33	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0
Astreopora	0	0	6	0	41	0	6	0	0	0	6	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	12	12
Other	13	3	24	5	0	0	3	0	4	4	0	0	0	4	4	0	8	0	0	0	0	0	0	5	0	0	0	0	0	4	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	3	0	0	12

Table 4.S6: Confusion matrix of Caribbean validation set. Numbers are percentages. Rows show the expert-assigned label, while columns show CoralNet predictions. The percentage of correct predictions can be found along the diagonal.

	Turf	Sand	Cyanobacteria	Dictyota	CCA	Seagrass	Gorgoniidae	Madracis mirabilis	Sponge	Encrusting sponge	Lobophora	Peyssonnelia	Montastraea	Halimeda	Porites	Plexaura	Siderastrea	Millepora	Orbicella faveolata	Agaricia	Pseudopterogorgia	Orbicella annularis	Pseudodiploria strigosa	Colpophyllia	Amphiroa	Gorgonian	Pseudoplexaura	Sargassum	Palythoa	Millepora complanata	Eunicea	Orbicella franksi	Sargassum	Pterogorgia	Meandrina	Erythropodium	Madracis decactis	Porites porites	Padina	Calcium Carbonate	Acropora	Stephanocoenia	Green Fleshy Algae	Other invertebrate	Red Fleshy Algae	Turbinaria (algae)	Zoanthid	Diploria	Pseudodiploria clivosa	Muricea	Other
Turf	82	8	4	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sand	13	85	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cyanobacteria	36	6	51	1	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Dictyota	41	1	1	55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CCA	29	5	8	1	53	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Seagrass	5	4	0	1	0	90	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gorgoniidae	17	5	1	1	0	1	68	0	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Madracis mirabilis	7	1	2	0	1	0	1	86	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sponge	24	5	4	1	1	0	2	0	49	7	1	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Encrusting sponge	40	4	8	2	5	0	1	0	13	26	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Lobophora	43	1	5	3	0	1	0	1	1	0	41	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Peyssonnelia	33	1	26	0	8	0	0	0	1	2	0	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Montastraea	12	2	0	1	0	0	1	0	1	0	0	0	75	0	0	0	1	1	2	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Halimeda	36	1	4	4	2	0	0	1	0	0	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porites	25	2	1	0	0	0	1	0	1	1	0	0	0	0	62	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Plexaura	11	1	1	1	1	0	1	0	1	0	0	0	0	0	0	60	0	0	0	0	2	0	0	0	0	0	13	0	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
Siderastrea	9	3	1	0	0	0	2	0	1	1	0	0	2	0	2	0	78	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Millepora	19	4	2	0	1	0	1	2	5	1	0	0	0	0	0	0	0	57	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Orbicella faveolata	7	0	0	0	0	0	1	1	0	0	0	0	1	0	1	0	1	0	79	1	0	4	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Agaricia	22	3	1	0	0	0	3	0	1	0	0	0	1	0	0	0	0	0	0	58	0	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
Pseudopterogorgia	9	4	1	3	0	1	18	0	1	0	0	0	0	0	0	5	0	0	0	0	50	0	1	0	0	0	2	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Orbicella annularis	13	2	1	1	0	0	2	1	1	0	0	0	4	0	1	0	0	0	23	0	0	50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	1
Pseudodiploria strigosa	7	1	0	1	0	0	2	0	1	0	0	0	3	0	0	0	0	0	0	0	1	0	75	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	2	0	0	0	0	0	0
Colpophyllia	11	1	0	0	0	0	4	0	0	0	2	0	2	0	0	1	0	1	0	1	0	0	11	65	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
Amphiroa	48	0	4	4	1	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	39	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gorgonian	14	3	2	1	1	0	2	0	3	0	0	0	1	0	0	0	3	0	0	1	1	0	1	0	0	66	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pseudoplexaura	10	1	1	0	1	0	0	0	1	0	0	0	0	0	0	14	1	0	0	0	1	0	0	0	0	0	1	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sargassum	30	1	7	4	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Palythoa	7	1	0	0	2	0	0	0	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	88	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Millepora complanata	20	2	0	0	2	0	1	2	2	1	1	0	0	0	0	0	0	27	0	1	1	0	1	0	0	0	0	0	0	0	40	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Eunicea	24	3	2	1	0	0	25	0	3	2	0	0	2	0	0	8	0	0	0	0	3	0	2	0	1	0	6	0	0	0	0	12	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	1	0
Orbicella franksi	19	1	0	0	0	0	0	0	0	2	0	0	1	0	2	1	5	1	22	1	0	1	0	0	0	0	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sargassum	49	1	3	4	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	13	0	0	0	0	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pterogorgia	30	13	0	3	0	2	8	0	0	0	0	0	0	0	0	0	0	2	0	0	10	0	0	0	0	2	0	0	0	0	0	0	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
Meandrina	9	5	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	2	0	2	0	5	11	0	0	0	0	0	0	0	0	0	0	0	58	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Erythropodium	43	4	2	2	4	0	11	0	8	2	0	0	0	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	15	0	0	0	0	0	2	0	0	0	0	0	2	0
Madracis decactis	44	12	0	7	0	0	2	7	0	0	0	0	0	0	2	0	0	2	5	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0
Porites porites	27	0	7	0	2	0	0	2	0	2	0	0	0	0	0	2	0	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0	0	0
Padina	53	0	0	8	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	6	0	0	0	0
Calcium Carbonate	40	31	3	0	0	0	0	3	0	0	0	0	3	0	0	3	0	0	0	0	3	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	9	0	0	0	0	0	0	0
Acropora	18	0	14	0	0	0	50	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0
Stephanocoenia	20	12	0	0	0	0	4	0	0	0	0	0	4	0	20	0	4	0	8	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
Green Fleshy Algae	62	0	0	19	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0
Other invertebrate	33	0	5	0	0	0	10	5	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	10	10	0	0	0	0	0	0	0	14	0	0	0	0	5
Red Fleshy Algae	20	0	55	10	10	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Turbinaria (algae)	29	0	0	0	0	0	0	0	0	0	12	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	47	0	0	0	0
Zoanthid	25	0	6	0	0	0	6	0	0	0	0	0	6	0	0	0	6	0	6	0	0	13	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0
Diploria	23	8	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	31	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0
Pseudodiploria clivosa	15	0	0	0	0	0	8	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	23	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0
Muricea	30	10	0	10	0	0	10	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Other	28	22	6	3	0	0	0	3	0	0	0	0	6	0	0	0	0	6	0	3	0	0	6	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9

CHAPTER 4 ACKNOWLEDGEMENTS

Chapter 4 contains material being prepared for publication coauthored with Petrovic, V.; Edwards, C. B.; Pedersen, N.; Lubarsky, K.; Holloway, N.; Kuester, F.; Sandin, S. A. The dissertation author was the primary researcher and author of this material.

This research was made possible through efforts of the 100 Island Challenge, and the various organizations that have supported this project, including the Scripps family and the Moore Family Foundation. Data were collected under permit and in collaboration with partner agencies in the Maldives (including the Maldives Marine Research Institute), the Federated States of Micronesia (including the Conservation Society of Pohnpei and OneReef Micronesia), and the territories of St. Maarten, St. Eustatius, and Saba (including the Dutch Caribbean Nature Alliance).

CONCLUSIONS

In this dissertation, I investigated and discussed means to improve the collection of oceanographic measurements. In Chapter 1 I showed that Junge-type power law models do a poor job of capturing oceanographic particle size distributions, and in Chapters 2, 3, and 4 I presented a range of novel information to the coral science community on field instrumentation and the processing and analysis of 2D and 3D image products with machine-learning-based tools.

In Chapter 2, my coauthors and I showed that power law models do not adequately parametrize oceanic particle size distributions, based on a large dataset of high-resolution particle size distributions collected by Dariusz Stramski, Rick Reynolds, and others over decades in the Arctic Ocean. When tested with this large dataset, power law models did a poor job predicting measured particle size fractions. This is an important result, because power law models are often used to parametrize particle size distributions in oceanography. Particle size directly physically affects light, so the particle size distribution affects the signals of optical instruments. We also showed that models based on CDF percentile diameters predicted particle size fractions well. If relationships between these percentile diameters and optical measurements can be found, it may be possible to measure some particle size characteristics from space, and separate the components of satellite optical signals that are due to particle size from those influenced by other seawater constituents and characteristics.

In Chapter 3 I discussed methods for large-area underwater image surveys of coral reefs. We proposed reporting standards for large-area-image-survey-derived 3D models, and explained how to acquire clear images with substantial overlap over the entire field site. Calculating accurate 3D models requires many overlapping clear views of each point on the bottom of the ocean. The goal of this study is to increase the number of coral reef labs around the world using large-area

image surveys, and to provide some useful resources on how to do the image surveys well. More labs doing large-area image surveys means more experiments on how to get useful coral science information out of 3D models, more potential users for AI tools, and more sharable high-quality training and testing databases.

Chapters 3 and 4 evaluated AI methods for expediting processing of the huge volumes of image-derived data the surveys produce. In Chapter 3, we compared the performance of 2D, 2.5D, and 3D neural networks to segment 3D models of coral reefs and 2D/2.5D orthoprojections thereof. Neural networks segmented 2D orthoprojections and 3D models of field sites equally well. We also discussed many possible reasons to prefer a 2D workflow over 3D or vice versa, but concluded that the performance of machine learning models should not be a determining factor at this point. Desired data outputs and available image data product annotation software are more important factors when deciding if working in 3D is worth the extra effort. This study was the first to be published on using 3D neural networks to segment 3D point clouds of coral reefs; it shows that 3D segmentation works, and can become a useful tool for coral researchers with further development.

In Chapter 4, we evaluated the publicly-available neural-network-driven software CoralNet. We used very large datasets (5000 images per region, 5-50 points per image), but found that, when working with large sets of possible taxonomic categories, there will be few training points for most taxonomies because a few commonly-occurring taxonomies account for most of the training points. This imbalance led to relatively good overall CoralNet prediction accuracy, because accuracy scores are dominated by common classes, but unreliable performance on most taxonomic categories, as most of them are relatively rare. We also presented evidence that

CoralNet, despite its imperfect accuracy with our training and testing datasets, reduced the time required to point sample an image by 36%.

These four investigations occurred at different scales in different ecosystems, but they work towards the same goal: improving the ability of researchers to collect valuable data about the environments they study by using optical instruments. Chapter 1 reversed the long-held belief that oceanographic particle size distributions are well represented by a Junge-type power law model. Taken together, Chapters 2, 3, and 4 represent a significant contribution to the coral research community's understanding of large-area image surveys and machine-learning-assisted annotation of the resulting image products.

DIRECTIONS OF FUTURE ML CORAL RESEARCH

In Chapter 3, we compared the performance of neural networks trained on 2D orthoprojections and 3D pointclouds. To expand on this study, one could investigate the addition of a time dimension, resulting in a neural network designed to input 4D information, or 3D but with a time dimension instead of the third spatial dimension. Neural networks designed to segment videos perform an analogous task—interpreting successive 2D frames as the contents advance in time. The Sandin and Smith labs have been collecting image surveys of sites on Palmyra annually since 2013, so reasonably long time series of coral reef 3D models and 2D orthoprojections are available.

Time dimension information can be used in two different ways. First, if an expert-verified ground truth is known for one point in time, it can be projected forward. Logically, it's reasonable to suspect that a neural network, when predicting a segmentation of a coral reef site, would perform better if shown an expert-verified ground truth of the site from the previous year. The task is much

simpler that way, as few corals grow from nothing or move in short to intermediate periods of time, so most taxonomic identities could be inferred from the previous year. It's likely that this method would be particularly advantageous for segmentation of time series, as an accurate expert-verified ground truth of the site in the first year would assist automated predictions of the following years.

Time dimension information can be useful even if an expert-verified ground truth is not available for any of the points in time. A neural network can be given additional information by providing it with 3D models or 2D orthoprojections of a site across many points in time, which might improve performance. This arrangement is analogous to providing a bird identification neural network with multiple frames from a video as a bird flew by—a neural network that can consult multiple frames at once has more information than one viewing a solitary frame. For our coral 3D models we take images from the same top-down perspective every year, so the information benefit of consulting multiple years' models at once might be limited, but I think it is worth investigating.

A different but also potentially valuable direction of research is unsupervised learning. It may be possible to improve domain generalization and reduce training data volume demands by pre-training a neural network on the entire unlabeled database of thousands of coral reef 3D models. As these models add up to something like a petabyte of pointcloud information, an unsupervised learning operation would require substantial database engineering to support, and might take a long time or require a very high number of GPUs. Unfortunately, many of the Sandin lab 3D models are stored on UCSD library servers. The library servers are set up as an archive, so they do not support rapid transfer of large volumes of data, which are necessary for unsupervised learning from a large database.

ACCURACY LEVEL REQUIRED FOR BIOLOGICAL SIGNIFICANCE

It is important to ask the question: When can we let neural networks annotate image survey products without editing or verification? Editing predictions takes orders of magnitude more time than generating them, so it would be a huge benefit if we did not have to do it.

This question has been investigated in the context of CoralNet—experiments have shown that CoralNet can produce accurate percent cover data without expert editing and verification of its predictions. If researchers only need total coral percent cover data, CoralNet can be trained to provide it. More information can be found in "Leveraging Automated Image Analysis Tools to Transform Our Capacity to Assess Status and Trends of Coral Reefs" by Williams et al., published in 2019. I discussed this issue in more detail in Section 4.3.3 of Chapter 4.

What about segmentation? Measuring growth of individual corals that amounts to millimeters or centimeters a year requires precise measurements, so the imprecision of unedited neural network predictions that are only ~80% accurate is undesirable. However, as I discussed in Chapter 3, the expert-generated ground truth used to generate that accuracy metric is not 100% reliable. Different experts will produce different ground truth segmentations. Therefore, AI-generated predictions do not have to be 100% accurate to be useful, they just have to be as accurate as an average expert. However, accuracy and consistency of coral experts has been poorly studied and remains difficult to quantify, so future experiments are needed to create benchmarks against which to compare automated predictions.

In the meantime, we can then ask the question: can unedited neural network predictions be useful even if they are less accurate than experts? In Chapter 3, my coauthors and I found that experts were about 90% accurate relative to a consensus of them, while neural networks were

about 80% accurate, so predictions are still likely to be less accurate than experts in coral applications. Unfortunately, we find that editing automated segmentations to the standard of an expert is strongly preferable to using them unedited, for some complex reasons.

In the context of tracking the growth of an individual coral through time, it is important to consider that only a few time steps are generally available, as collecting time series of coral reefs can be quite hard. Segmentations of corals need to be precise in order to confidently reveal trends as quickly as possible and examine dynamics between individual time points.

Tracking a single coral across a few points in time might require very precise measurements for meaningful results, but what about tracking the growth a specific genus across a large area? Tens or hundreds of 3D models would be available for such an experiment, providing much more statistical information than short time series of individual corals. If predicted segmentations of that genus were generated on hundreds of 3D models, it is conceivable that trends could still be identified even if the predictions were only 80% accurate. However, this is not true if individual predictions are biased by an unknown degree, rather than exhibiting consistent variation around a mean. While the segmentation neural networks we trained were approximately unbiased on the data we tested them on, it is likely they would become biased when making predictions on 3D models that vary significantly from the training data, due to e.g. different geography or water clarity. For such predictions to be meaningful measurements, the neural network's bias on each unique 3D model would have to be evaluated by segmenting some of each model manually, or the predictions would need to be edited and verified by experts. We find editing predictions to high accuracy provides the most useful form of segmentation for investigating coral ecology. In the future, larger training datasets might make it possible to train a segmentation neural network that is relatively unbiased on most possible inputs, or at least its bias can be characterized

with reasonable confidence. Such a neural network could be used to investigate trends across many 3D models, even if it is not perfectly accurate on a per-coral basis, because the bias in its predictions can be accounted for.

GENERAL THOUGHTS AND CONCLUSIONS

The projects in all four chapters suffer from much-smaller-than-ideal datasets, limited by both expensive/laborious processing and the time/resources required to get instruments to field sites. My first chapter drew from a few hundred measurements in the Arctic, collected over approximately a decade. This Arctic dataset is the largest of its kind, and yet it is quite small compared to an ideal dataset, which would cover the entirety of the Arctic at a variety of depths at many points in time. Chapter 3 used a time series from a single site on the island of Palmyra, because—to be best of my knowledge—that is the largest existing dataset of traced 3D high-resolution coral reef pointclouds. The datasets from the Maldives, Micronesia, and the Caribbean used in Chapter 4 were comprised of tens of thousands of point sampled images; however these images contained only tens, hundreds, or thousands of points assigned to coral and algae genera/species scientists wish to study, resulting in NN accuracies that were still rising when the last of our training data was exhausted. This suggests more training data would result in more accurate, more useful AI.

Initiatives to work at globally-comprehensive scale are already underway. The study discussed in Chapter 1 was based on data only from the Arctic, but Dariusz Stramski and Rick Reynolds continue to pursue PSD, POC, and other algorithms globally, enabled by their dogged commitment to acquiring field data, and decades of investment by NASA and other American and international science supporting agencies and institutions. Chapter 3 was based on data concerning

one genus at one site on one island, but the Sandin and Kuester labs too are aspiring globally with admirable initiatives like their 100 Island Challenge, which entails visiting a dozen sites on 100 islands twice in five years. We are making progress, but there is a long way to go, especially at the federal/institutional level above the many scattered, independent research groups.

In a nutshell, we need way, way more data. An optical satellite algorithm that is accurate in all oceans, at any time of the day or year, requires a similarly comprehensive database of measurements of seawater optical and particulate characteristic. The same is true of AI for coral—without globally comprehensive training and testing databases, we will not be able to build globally comprehensive AI tools. Without a comprehensive database, we will not even be able to test how an AI tool behaves in all situations. How can we expect anyone to trust it if we cannot even test it?

To expand the collection of measurements, I see three primary areas of focus. The first is obvious—go get more! Funding agencies should support scientists, who have the experience necessary to collect good field data, to get out in the field with better and more instruments manned by larger teams of employees, graduate students, and postdocs. Further emphasis should be given to standardization of field collection and data processing/storage procedures to simplify pooling of datasets from disparate research groups.

A second area of emphasis is technology: instruments, platforms, and software to process and analyze instrument data. New instruments (e.g., NASAs hyperspectral PACE mission), combined with new platforms (e.g., unmanned vehicles), should be developed to expand what we are able to measure in terms of accuracy and breadth in parameters measured, spatial coverage, and chronological repetition.

User-friendly, fast, minimally-infuriating software is necessary to help scientists evaluate this ever-growing volume of data. NASA is a leading example of supporting broad and detailed availability of its data products and underlying algorithms. Vid Petrovic/Falko Kuester's Viscore 3D pointcloud software and Oscar Beijbom/David Kriegman's cloud-based CoralNet point sampling tool are exemplary, but they and similar initiatives require an order of magnitude more financial support and dozens of engineers to reach/satisfy a global user base, many of whom have limited familiarity with math and programming. Unfortunately, many innovative software tools are distributed primarily by code made publicly available at websites like GitHub—this is a great way for programmers to share, but is not very accessible to many oceanographers and biologists.

The third avenue I see for progress is probably the most important—staffing. I began this dissertation, and this conclusion, by decrying the scale of our datasets relative to the magnitude of the world's oceans. However, the original theme of my dissertation was not that we do not have enough sensors or field measurements, but that we can already collect more than we are able to process and analyze. This not true of *in situ* particle size measurements (we have relatively few of them and processing procedures are well-established), but it IS true of the satellite imagery they aim to help interpret—if satellites collect images that we lack algorithms to extract information from, they are not being maximally processed, analyzed, or utilized. The situation in the cases of coral and satellite optical data collection is the same: we can take pictures, but we cannot access all the information they contain.

To address this shortcoming, we need more people. Coral scientists need more people annotating field imagery, and satellite oceanographers need more people on more ships making measurements in the field. We will not be able to maximally extract information from optical

sensors until we have robust algorithms, models, or neural networks to interpret measured optical signals.