2016-2017 Annual Report to the National Ocean Partnership Program (NOPP)

Title: Demonstrating an Effective Marine Biodiversity Observation Network in the Santa Barbara Channel

Agreement number: NNX14AR62A

PI: Robert Miller – Marine Science Institute, UCSB

Co-PIs: David Siegel (UCSB), Craig Carlson (UCSB), Kevin Lafferty (USGS), Andrew Rassweiler (UCSB), Daniel Reed (UCSB), Phaedon Kyriakidis (UCSB), Deborah Iglesias-Rodriguez (UCSB), B.S. Manjunath (UCSB), Milton Love (UCSB), Andrew Thompson (NOAA SWFSC), John Hildebrand (UCSD), Douglas McCauley (UCSB)

Date: July 29, 2017

Performance Period: October 1, 2016 to September 30, 2017

Project Goals and Objectives:

Although we spend millions annually monitoring marine resources, these efforts are uncoordinated and have major information gaps. We are developing a scalable and transferable demonstration Marine Biodiversity Observation Network (MBON) in the Santa Barbara Channel (SBC), one of the most monitored areas of the world. SBC MBON will network existing monitoring efforts and fill the major remaining information gaps. Our focus on SBC allows us to effectively cover the complete spectrum of biodiversity from ecosystems to microbes within a reasonable scope of funding. This is due to the profusion of existing biological monitoring and research programs by our partners including government agencies, universities and NGOs. Yet there are significant gaps in our knowledge of even this relatively well-studied area, such as microbial diversity. Our overall objective is to provide a complete picture of biodiversity in SBC using a transferable system that integrates and augments existing monitoring programs including the NSF-funded SBC LTER program, Channel Islands National Park, and NASA Plumes and Blooms. Broad goals are to:

A. Integrate biodiversity data to enable inferences about regional biodiversity. Synthesizing information relevant to biodiversity requires integrating highly heterogenous data collected at widely different temporal and spatial scales. We employ advanced techniques in spatial statistics for this synthesis and will provide multiple biodiversity-related data products, including holistic indices that will provide easily interpretable measures of ecosystem health.
B. Develop advanced methods in optical and acoustic imaging and genomics for monitoring biodiversity in partnership with ongoing monitoring and research programs to begin filling the gaps in our knowledge. A key element of our plan is a ‘pincer movement’ using two classes of methods that approach diversity observation from opposite directions: optical and acoustic imagery from the ecosystem scale down to the species level, and molecular biology from the genetic scale up through community level.

C. Implement a tradeoff framework that optimizes allocation of sampling effort. An effective marine BON will require targeted sampling to address key data gaps, while making best use of existing sampling efforts, thereby gaining a complete description of biodiversity while minimizing cost. Optimal decisions about data collection will require a framework for balancing costs and benefits of alternative sampling. Such a framework will be used to make recommendations for how resources should be allocated in a full-scale MBON as a function of the program’s goals and anticipated funding level.

In the Cooperative Agreement with NASA, the project agreed to fulfill the following more specific goals.

1. Develop a scalable and transferable demonstration marine biodiversity observation network (BON) in the Santa Barbara Channel.
2. Integrate and augment existing monitoring programs to provide geographically-integrated time-series metrics of biodiversity and ecosystem health, a transferable BON data management system, and a sampling cost-benefit optimization framework useful for designing a BON anywhere.
3. Disseminate products and results from the demonstration marine BON to a wide range of end users from scientists to school children.
4. Develop and implement advanced methods in optical and acoustic imaging and genomics for biodiversity monitoring.
5. Integrate this activity into the international Group on Earth Observations Biodiversity Observation Network (GEO BON).
6. Prepare for, participate in, and respond appropriately to reviews of the project.
7. Participate in scientific meetings or professional society meetings identified by U.S. Government program managers as pertinent to project goals.
8. Inform the relevant U.S. Government program managers of any results nearing publication and the release of final data products so that they might prepare for the announcement and any associated publicity and/or public outreach.
9. Data produced under this cooperative agreement will be available without restriction as to its disclosure, use or duplication except as otherwise negotiated by NASA and the University of California, Santa Barbara. The goal is full and
open exchange of data and metadata with minimum possible cost, delay and restriction.

Year 3 Progress:

Here we describe our approach and progress organized by the three broad goals above.

A. Integrate biodiversity data to enable inferences about regional biodiversity.

We have assembled data from 173 coastal sites monitored by 4 different programs, many of which have 30+ year data series. These datasets document abundance of more than 350 distinct taxa including fish, invertebrates and algae. To facilitate data synthesis, we have developed a taxonomic database to enable comparison of biodiversity data from different sources. The taxonomic database is scalable to accommodate the integration of additional data sets.

NPS Kelp Forest Monitoring (KFM)
- 37 sites across Channel Islands (San Miguel, Santa Rosa, Santa Cruz, Anacapa, Santa Barbara, and San Clemente Islands)
- Annual sampling from 1982 to present
- 118 distinct taxa
- Benthic density, benthic cover, fish density data, fish size-frequency, kelp tagging, subtidal temperature

Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO)
- 118 sites across Southern California Bight
- Varied sampling from 1999 to present
- 131 distinct taxa
- Benthic density, benthic cover, fish density, fish size-frequency

San Nicholas Island (SNI)
- 7 sites around San Nicholas Island
- Biannual sampling from 1980 to present
- 325 distinct taxa
- Benthic density, benthic cover, fish density, kelp tagging

Santa Barbara Channel Long Term Ecological Research (SBC LTER)
- 11 sites across north side of Santa Cruz Island
- Varied sampling from 1982 to present
- 307 distinct taxa
- Benthic cover, fish density, kelp tagging, subtidal temperature

Specific data that we have acquired and is in various stages of processing now includes approximately 50 datasets spanning taxa from microbes to whales as listed in Table 1.
**Table 1.** Dataset inventory for SBC MBON project. Datasets listed as anticipated are in various stages of processing.

<table>
<thead>
<tr>
<th>SBC MBON dataset_id</th>
<th>nickname</th>
<th>status</th>
<th>Pub_notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>x.6</td>
<td>gray whale sightings</td>
<td>anticipated</td>
<td>Smith</td>
</tr>
<tr>
<td>x.7</td>
<td>Plumes and Blooms profiles</td>
<td>anticipated</td>
<td>O’Brien to coordinate</td>
</tr>
<tr>
<td>x.8</td>
<td>Bisque - areal abundance</td>
<td>anticipated</td>
<td></td>
</tr>
<tr>
<td>pisco.19</td>
<td>Shallow subtidal - fish</td>
<td>cataloged</td>
<td>PISCO</td>
</tr>
<tr>
<td>knb-iter-sbc.61</td>
<td>sea otter sightings</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.74</td>
<td>landsat kelp biomass</td>
<td>anticipated</td>
<td></td>
</tr>
<tr>
<td>x.13</td>
<td>marine mammals - swfsc</td>
<td>anticipated</td>
<td></td>
</tr>
<tr>
<td>pisco.20</td>
<td>Shallow subtidal - invert, algae (cover)</td>
<td>cataloged</td>
<td>PISCO</td>
</tr>
<tr>
<td>pisco.21</td>
<td>Shallow subtidal - invert, algae (density)</td>
<td>cataloged</td>
<td>PISCO</td>
</tr>
<tr>
<td>x.22</td>
<td>Shallow subtidal - fish</td>
<td>anticipated</td>
<td>NPS KFMP</td>
</tr>
<tr>
<td>1001</td>
<td>kelp_forest_integrated_cover_upc_rpc</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>x.14</td>
<td>marine mammals - SIO</td>
<td>anticipated</td>
<td>Whale Acoustics labWhale Acoustics lab</td>
</tr>
<tr>
<td>x.15</td>
<td>deep reef epibenthos</td>
<td>anticipated</td>
<td>CINMS</td>
</tr>
<tr>
<td>1002</td>
<td>kelp_forest_integrated_fish</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.15</td>
<td>shallow subtidal invert, algae (cover)</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.24</td>
<td>shallow subtidal kelp - sbc iter</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.19</td>
<td>shallow subtidal invert, algae (density)</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.17</td>
<td>shallow subtidal fish - sbc iter</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>1003</td>
<td>kelp_forest_integrated_quad_swath</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>1004</td>
<td>combined_species_list</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.91</td>
<td>Beach invertebrates - SBC LTER</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>knb-iter-sbc.51</td>
<td>birds shore</td>
<td>cataloged</td>
<td></td>
</tr>
<tr>
<td>x.25</td>
<td>Rocky intertidal</td>
<td>anticipated</td>
<td>PISCO</td>
</tr>
<tr>
<td>1006</td>
<td>deep fish</td>
<td>anticipated</td>
<td>Kui</td>
</tr>
<tr>
<td>1005</td>
<td>calcofi ichthyoplankton</td>
<td>draft</td>
<td>Andrew Thompson</td>
</tr>
<tr>
<td>x.23</td>
<td>Shallow subtidal - invert, algae (cover)</td>
<td>anticipated</td>
<td>NPS KFMP</td>
</tr>
<tr>
<td>x.11</td>
<td>birds, pelagic and shore - calcofi</td>
<td>anticipated</td>
<td></td>
</tr>
<tr>
<td>x.10</td>
<td>microbial genomics</td>
<td>anticipated</td>
<td></td>
</tr>
<tr>
<td>x.18</td>
<td>deep soft bottom benthos - 2</td>
<td>anticipated</td>
<td>Love lab</td>
</tr>
<tr>
<td>x.17</td>
<td>deep soft bottom benthos - 1</td>
<td>anticipated</td>
<td>ccwrp</td>
</tr>
</tbody>
</table>
Integrated reef community data include EML metadata and have been submitted to a public data repository, which supports DOIs, revision control, and metadata aggregation by DataONE. Some datasets have already been updated with a second year’s data and example code for data transformations. Our model for data package design is the technical completeness of an “ESA data paper”, which are expected to ensure a high standard of usability, especially with respect to associated metadata.

Data citations for integrated reef datasets above (most recent revisions only) with DOIs:


This year we completed a transferable data integration flow that can be applied to other datasets and MBON projects. The basic steps of data integration consist of (1) Data cleaning, including adding missing information (e.g. sampling method), standardizing time/date format, and correcting any data entry errors; (2) Data standardization, including homogenizing column names across projects and checking species’ scientific names to meet the requirement of current convention; (3) Data integration, including extracting information from all projects (e.g. site locations, transect depths, and taxonomy), as well as combining datasets based on data categories (e.g. taxonomic categories); (4) Dataset production (sometimes including summary statistics) for SBCBON website and statistical analyses.

As part of this process we (a) customized a metadata database for SBC MBON based on a relational model, code and expertise used several sites in the NSF’s Long Term Ecological Research (LTER) Network, and (b) created reusable code in the R programming language, for EML-described data package production, based on open source tools (https://github.com/ropensci/EML/). Both the database and R code are reusable within the SBC MBON and beyond.
Additionally, we are in the process of combining biological data with physical data to allow analysis of the physical drivers of biodiversity. Because our goal is to use these results to develop regional scale inference about patterns of biodiversity, we focus on physical data for which we have continuous coverage within the region. These data include 1) bathymetry and substrate data derived from sonar surveys, 2) sea-surface temperature, chlorophyll and local kelp canopy derived from satellite imagery, 3) wave disturbance and ocean current data derived from regional models, and 4) data on the shape of the coastline calculated from maps. Physical data that have thus far been processed and compiled for predicting biological spatial-temporal variations include:

- Sea Surface Temperature
- Chlorophyll a
  [http://coastwatch.pfeg.noaa.gov/](http://coastwatch.pfeg.noaa.gov/)
- Ocean connectivity
  Siegel lab
- Wave disturbance
  SBC LTER
- Landsat-derived kelp biomass
  SBC LTER/Siegel lab
- Climate indices (MEI, PDO, NPGO)
  [http://www.esrl.noaa.gov/psd/](http://www.esrl.noaa.gov/psd/)

Data specialist Li Kui continues acquiring and integrating biodiversity data in ways that are sustainable into the future (clearly annotated code, versioning, full metadata). Dr. Kui is also an ecologist who is participating in data analysis efforts. MBON postdoc Thomas Lamy, who started in July 2015, is developing predictive and spatial models of marine biodiversity using physical and biological data derived from remote sensing. He has one manuscript in review and another two in advanced stages of preparation. In collaboration with the NOAA Biogeography Branch, we (Miller, Rassweiler, Reed) leveraged MBON to obtain NOAA funding for an additional postdoc in 2015. We recruited Dr. Rhiannon Rognstad for this position and she started this year. Rhiannon earned her Ph.D. in Biology at the University of South Carolina at Columbia, where she worked with Jerry Hilbish on marine population connectivity and how it shapes range boundaries and interacts with climate change. On SBC MBON she is developing species archetype models to relate physical and biological data. She already has a manuscript nearly ready for submission on this project, and is now working with Andrew Thompson on relating biological data from the CalCOFI program to remote sensing data.

Analyses thus far show that reef community variation in the SBC is mainly explained by a linear trend, from east to west, driven almost entirely by sea surface temperature. Communities also vary at broad scales (> 60km) within the SBC and this variation is explained primarily by differences in propagule delivery (e.g. site oceanographic destination strength average over all the model domain) and by differences in pelagic primary production (mean chlorophyll a). We find that communities do not, in general, vary significantly at scales smaller than ~30 km. Each island in the region has different assemblages and it is difficult to detect significant community variation (scales of ~30-60 km) within an island.
Surprisingly, wave disturbances (mean wave height and number of days waves exceed 2m) and kelp canopy biomass (mean long-term canopy biomass, inter-annual variability in kelp canopy biomass and extinction probability computed as the number of years a site had no canopy over the total number of years) turned out to be poor predictors of biodiversity. However, in a smaller high-resolution dataset (SBC LTER), biodiversity of fishes and sessile invertebrates was positively affected by kelp.

Statistical analyses have been done to test the temporal and spatial changes in deep-water fish communities. Three sites, Anacapa passage, Footprint, and Piggy Bank, for which we have time-series data, have been the main focus for these analyses. Preliminary results suggest that the fish communities are driven by habitat, especially depth and substrate. The fish species seem to have shifted in depth over the past 15 years and sizes of some fish species have decreased over time. The above results for shallow and deep benthic communities are being prepared for publication in 2017.

Structural equation modeling analysis of long-term kelp forest community data has shown how sea urchin grazing effects can be misinterpreted as kelp effects, because sea urchins can overgraze giant kelp, understory algae, and sessile invertebrates alike. Giant kelp structure had indirect effects on reef communities because it shades out understory algae that compete with sessile invertebrates. When released from competition, sessile species in turn increase the diversity of mobile predators. Giant kelp also directly increased predator diversity by providing habitat. These results are in review at Proc B.

**SBC BON Remote Sensing**

Remote sensing data is critical for scaling up local observations of biodiversity and for relating physical and ecological variables to marine biodiversity. The SBC mBON is focused upon three activities linking remote sensing to observations of marine biodiversity. These build on on-going NASA and NSF supported projects and span both subtidal and pelagic environments. Specifically, the remote sensing component of the SBC BON is focused on:

- The remote sensing of giant kelp populations and the application to these products to answer ecological and biodiversity questions
- The analysis of satellite ocean color observations using novel ocean color inversion approaches, and
- The analysis of planktonic biodiversity indices from the Plumes and Blooms (PnB) time-series study applied to SBC BON genomics observations.

 Accomplishments from all three of these components are reported in this annual report.
Both multispectral and hyperspectral remote sensing of giant kelp canopy fronds are used to help answer ecological and biodiversity questions for the SBC BON. We have assembled a Landsat giant kelp (*Macrocystis*) canopy biomass timeseries from 1984 – present, spanning its dominant range in the NE Pacific (San Francisco, CA to Punta Eugenia, Baja California Sur, Mexico). This dataset has recently been used to determine the geographical variability and non-linear response of the environmental controls of giant kelp biomass dynamics (Bell et al. 2015a), as well as a test of metapopulation theory, showing that well connected patches had high probabilities of colonization and lower probabilities of extinction (Castorani et al. 2015; 2017). Presently, this dataset is being used to identify large-scale biogeographic population clusters using seascape genetic approaches (Johansson et al. 2015). Multiple hyperspectral images of the SBC have been collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) as part of the HyspIRI Preparatory Airborne Campaign. Empirical relationships between laboratory reflectance of giant kelp blades and their physiological state (Chl:C) have been developed and applied to the hyperspectral images. These images have shown that the physiological state of the canopy is positively correlated to the depth of the reef where the kelp plants attach, and may have implications for measuring the productivity and age structure of kelp forests (Figure 1; Bell et al. 2015). Tom Bell, a postdoctoral scholar at UCLA and a former student in the Siegel lab, is leading this work.

The goal of the second project was to characterize what controls space-time variability of optical properties of the Santa Barbara Channel and, more specifically, to understand how changes in chlorophyll concentrations (*chl*) and particulate backscattering coefficients (*b*_bp) co-vary over time and space (Henderikx Freitas et al. 2017). Results from this project have recently been published in JGR-Oceans. We are also collaborating with Maria Kavanaugh (WHOI) in applying these data with her seascape characteriz

![Figure 1. Map of the distribution of (a) estimated Chl:C in the surface canopy (b) depth and (c) kelp pixel fraction of Santa Barbara area kelp forests on April 11th, 2013. Insets show magnified data of the Isla Vista kelp forest and reef. Data for a. and c. from AVIRIS hyperspectral images, and b. from NOAA National Geophysical Data Center coastal relief model (Divins & Metzger, 2009).](image)
ation approaches for mapping marine biodiversity indices.

Figure 2. Correlation coefficients between a) chlorophyll and particle backscatter and b) backscatter and significant wave heights from NDBC buoy 46054. This demonstrates the controls on patterns of water clarity in the region.

The last project uses data from the long-term Plumes and Blooms (PnB) time series study to assess and model ocean color changes in a complex coastal site and is supported by the NASA Ocean Biology and Biogeochemistry program. PnB conducts monthly day-long cruises at 7 stations crossing the Santa Barbara Channel (~45 km long transect). At each station, measurements of ocean color spectra, inherent optical properties, phytoplankton pigment, dissolved and particulate carbon, and macronutrient concentrations, and particle size spectra are measured. Samples are also collected for particulate DNA for both prokaryotic and eukaryotic (Emma Wear et al., work in progress). Our eventual goal in the SBC mBON is to statistically relate bio-optical information of
planktonic biodiversity from the PnB measurement suite (phytoplankton pigment concentrations, phytoplankton absorption spectra, etc.) with signatures from the SBC mBON genomic work. Emma Wear, a recent graduate from the Carlson laboratory is comparing her genomic indices of prokaryotic biodiversity with the PnB bio-optical data. Dylan Catlett, a student in Siegel lab on a NASA Earth System Science Fellowship, is continuing this work in collaboration with Profs. Siegel, Carlson and Iglesias-Rodriguez. Catlett has analyzed a 9 year data set of hyperspectral phytoplankton-specific absorption spectra (a_{ph}) and phytoplankton pigment concentrations from PnB (Catlett and Siegel, in review, JGR-Oceans). Absorption features in a_{ph} have been identified using a derivative analysis and these spectral signatures have been linked with individual phytoplankton pigment concentrations. Phytoplankton community structure in the SBC over this time period has been characterized via empirical orthogonal function (EOF) analysis and corroborated using conventional cluster-based analyses. Catlett finds excellent results (often > 80%) in deriving phytoplankton pigment concentrations and suites of pigment communities (determined via EOF analyses) from the a_{ph} spectral derivatives. This work is presently in review at the Journal of Geophysical Research – Oceans. These results provide a useful pathway for determining phytoplankton community structure from hyperspectral ocean color imagery, such as will be available from NASA’s Plankton, Aerosol, Cloud and ocean Ecosystem (PACE) mission. Moving forward, we plan to use compare the signatures in a_{ph} identified from the derivative analysis and the phytoplankton pigment community estimates with the eukaryotic genomic indices to provide a complete picture of planktonic biodiversity in the SBC.

B. Develop advanced methods in optical and acoustic imaging and genomics for monitoring biodiversity.

Optical imagery - field
We have used an SLR with 14mm lens mounted on a rigid frame (quadrapod) to image shallow benthic and kelp forest communities at 14 reef sites spanning the Santa Barbara Channel, including several Channel Islands National Park monitoring sites, as well as seven offshore oil platforms. In summer 2017 we photographed SBC LTER’s annual monitoring transects at the same time as diver surveys to examine how image analysis could be used to complement or replace time-consuming in situ surveys by humans. In deep water, we have obtained still imagery data collected from a SeaBED class autonomous underwater vehicle (AUV) that surveyed the benthos in the Santa Barbara Channel region. Transects were conducted on two seamounts known as the “Footprint” and “Piggy Bank.” Elizabeth Clarke, who leads the AUV program at NOAA NMFS NWFSC in Seattle WA, provided the data and will be collaborating with us as we move forward.

The SeaBED AUV is a hover-capable robotic vehicle that is able to work as close as 2 m off the seafloor while maintaining precise altitude and navigation control.
Its 2,000 m depth rating makes it an ideal tool for conducting surveys of reef, shelf and deep slope habitats. Its small footprint allows it to be operated from platforms ranging from global class oceanographic research ships to small vessels of opportunity. The SeaBed AUV can carry a wide variety of optical, acoustic, and oceanographic sensors for non-extractive surveys of the benthic communities in habitats that are too deep for divers and surface acoustics and too rugose for towed camera sleds and traditional bottom trawling. The area of each image is estimated from the measurement of the AUV altitude off the bottom and the specified camera field of view. This allows for density estimates of species abundance, biomass, and diversity.

The AUV data was collected during a “comparison cruise” in 2011. The objective of the cruise was to survey fish using the SeaBED AUV, the Dual Deepworker submersible, and an ROV coupled with fish detecting hydroacoustics and to compare the abundance, size distribution, biomass and diversity estimates from the different platforms to understand the capabilities of the different technologies and methods to assess West Coast groundfish populations in untrawlable areas.

The Piggy Bank is about 30 km² in area, ranging in depth from 275 to 900 meters; the Footprint Bank is about 10 km² in area, ranging in depth from 80 to 500 meters. The underwater visual surveys were planned to span from 400 m to the top of each seamount. The imagery includes stereo pairs that can be used to estimate size and biomass of organisms.

In September 2016, we conducted a new AUV survey in collaboration with Elizabeth Clarke, her colleagues affiliated with NWFSC, and CINMS to characterize deep coral and sponge communities in the Santa Barbara Channel region. The survey focused on an offshore feature named Calafia, which is composed of several asphalt “volcanoes,” the biology of which is little known. Density estimates of diversity, biomass, and species distributions from AUV imagery is now being analyzed and will allow the quantitative comparative analyses of habitats in this and other regions. Of particular interest are associations of fish communities with corals and other foundation species. The survey also provided imagery for expanding the training dataset for BisQue.

**Figure 3.** Locations (o) of the 25 survey blocks sampled in SBC using the SeaBed AUV.
This year, we have fully or partially annotated over 4000 images from the 25 survey blocks sampled by the AUV in Figure 3. Seven of the cells are located at the Footprint at depths ranging from 116 m near the top of the feature to 419 m, and four cells at the Piggy Bank ranging from 277 m at the top of this deeper feature to 455 m. We have identified at least 29 species among the 2752 fishes manually annotated in the AUV images. We have 43.5K data points of at least 60 invertebrate taxa and 15 unique types of sponge based on color and morphology located in the images. We are also annotating bacterial mats, algal drift, and anthropogenic debris (e.g., fishing gear, beer cans).

In the interest of developing BisQue as a tool to both characterize habitat in imagery and explore species diversity in relation to habitat variability, we are expanding the BisQue dataset by annotating physical substrate (e.g., mud, sand, cobble, boulder, bedrock) in images. Quantifying habitat heterogeneity in imagery by traditional methods such as overlaying a grid point matrix to estimate percent cover is time consuming and problematic, particularly when resolving small-scale variability. New techniques to automate the characterization of substrate in images will be explored. An emergent goal is to develop the automation of detecting and quantifying the overlap or proximity of benthic biota and habitat.

**Image processing**

During the first year we have performed extensive research on image analysis techniques targeted at underwater image classification on the operational level. Our main interest was targeted at finding techniques that could be trained automatically based only on provided training data and thus could be deployed automatically in our cloud-based image analysis and annotations system BisQue. Such techniques would have to overcome the complexity inherent in the multitude of existing image feature descriptions and classifications techniques. During this period we have explored several feature aggregation techniques that would allow automated classification based on a large number of computable feature descriptors and a distributed cloud based system. We have proposed and published in peer-reviewed conferences two novel techniques: 1) feature aggregation based on a CRF modeling of feature dependencies and 2) K-NN classification using dropout regularization.

During these studies we have used the extensive dataset of >2000 underwater images acquired during the shallow benthic field effort described above and manually annotated for percent coverage of sessile species in the BisQue system. Each image contained 100 annotated locations amounting to 200K data points with over 30 species. We have obtained very encouraging results on automated identification of the 11 most abundant classes (over 80% of data points) with classification accuracy of over 85%.
We further explored state of the art deep learning techniques using Convolutional Neural Networks on the same dataset and obtained comparable results. The major advantage of these techniques is that CNNs learn image features automatically (in convolutional layers) and operate directly on image pixels that can be efficiently accessed via the BisQue system. The major disadvantage is the computational complexity that requires use of the latest GPUs. We have used a very popular CNN library “Caffe” running on an nVidia K20 GPU and observed training times of 14 hours for the aforementioned dataset. This study, yet unpublished, will lead to the implementation of the automated classification module for the BisQue system, for which we are planning on utilizing an nVidia K40 GPU based server already acquired for another project.

**BisQue extensions**

In the past year we have integrated the Connoisseur model application component into the BisQue ecosystem for further development and potential user testing. Furthermore we containerized the deployments (utilizing the docker container ecosystem) which allows Connoisseur to be hosted on several types of hardware somewhat transparently by abstracting the access to the underlying GPU resources. We have tested these deployments on our labs cluster that includes one node with a single Nvidia K40 and another node with 4x Nvidia TitanX. We still have much integration work to do in order to allow scalable deployments of connoisseur across multiple GPU machines, but we believe that the current architecture will support a large range of existing problems without the need to scale in the near future.

**Acoustics**

Using the Channel Islands National Marine Sanctuary vessel *Shearwater*, Hildebrand's lab has continued to successfully service two high-frequency acoustic recording packages (HARPs) in the Santa Barbara Channel. Both instruments are part of a seafloor array of five high-frequency acoustic recording packages (HARPs) deployed to 1 km depth with a maximum horizontal aperture of 1 km that has been used to estimate the directionality of underwater sound radiated by current commercial ships. This array has returned good acoustic data for 2015 that has been analyzed to examine ship noise in SBC and the wider Southern California Bight. The underwater sound radiated by commercial ships is an unintended by-product of their operation and one of the most significant contributors to man-made noise at low frequencies in the ocean. As a ship of opportunity passed over the HARP array, the directions from the ship to each HARP along with the corresponding source levels were estimated for each ship location. Ships were tracked via satellites (Automatic Identification System—AIS). The directionality estimates of contemporary commercial ships exhibit significant stern-bow asymmetries among other quantitative characteristics. During 2015 the HARPS revealed an anomalously low presence of baleen whales in the Santa Barbara Channel, potentially associated with ocean conditions related to El Nino and “Warm Blob” conditions. His lab also continued to examine the potential impact of
anthropogenic sound on marine mammals including the use of explosive
deterrents by California fisheries, and the noise radiated by commercial shipping.
These results have important implications for managing ship traffic to minimize
impacts on marine mammals, particularly whales, in the SBC.

**Genomics**

Taxonomic identification of microorganisms has traditionally been challenging
because, in addition to limited morphological characteristics, less than 1% of the
microbial diversity has been cultured successfully. The introduction of genomic
approaches over the past two decades has allowed microbiologists to overcome
these limitations and assess microbial diversity in terrestrial and aquatic
ecosystems. The use of genetic markers with high-throughput Next Generation
(Next Gen) sequencing provides high-resolution taxonomy for phylogenetic
analyses as well as data for biogeographical distributions of marine microbial
plankton.

Identification of metazoans has traditionally been based on morphology, which
suffers from well-known limitations including phenotypic plasticity,
morphologically cryptic taxa and life stages and the need for taxonomic
expertise. In response, the use of DNA ‘barcodes’ has been increasingly adopted
for the study of biodiversity. There are several genetic markers that can be used
but the most common is mitochondrial cytochrome oxidase subunit I (COI), which
has been successfully used for the identification of fish, crustaceans, protists,
and many other organisms. We proposed to build on existing sampling efforts to
implement a sustained DNA collection program to be used for metabarcoding
and eDNA analyses.

**Prokaryotic and eukaryotic microplankton**

The BON supported Emma Wear as a graduate student researcher and she is
now a postdoc on the project, completing several manuscripts on her work.
Emma is a microbial oceanographer who is skilled in using molecular
approaches to reveal patterns in prokaryotic community structure and its link to
physical and biogeochemical patterns in the coastal ocean’s water column.
Emma has participated on monthly cruises across the Santa Barbara Channel as
part of the Plumes and Blooms times series program. This DNA sample archive
is being used to assess prokaryotic diversity as described below.

**Bacterial genomics:** We have refined the analysis of our existing time-series of
bacterial DNA samples, collected on the Plumes and Blooms cruise line
(http://www.oceancolor.ucsb.edu/plumes_and_blooms). This work also included
a comparative study of four 16S ribosomal RNA primers for bacteria and
archaea. We have two manuscripts in prep from this work, which will be

**Primer comparison study:** We designed this study to address apparent strong
biases in our time-series sequencing results, which were consistent with known
sequence mismatches and biases in the literature. By comparing four 16S rRNA primer sets against a common set of samples, we were able to assess both the biases inherent in the primer set we had selected, and to evaluate potential replacements.

This study ultimately indicated that the biases we had observed were not in fact due to our choice of primer set. We subsequently identified a bug in the bioinformatics program we were using to analyze our data (mothur, a common, community-supported program for 16S analysis); we reported that bug to the program developers and saw it resolved. Nonetheless, the primer comparison study will be a useful contribution to the field of marine microbial ecology, as it contrasts several commonly used primers against samples from a coastal aquatic system rather than the human- and terrestrial-focused standards commonly used in the literature for primer benchmarking.

We compared four primer sets targeting different regions of the 16S bacterial rRNA gene, all of which are currently in use in the community, through a series of tests. We first constructed a mock community of known composition using 22 cloned full-length 16S genes from bacteria and archaea from the Santa Barbara Channel. This allowed us to quantify which taxa the various primer sets over- and under-amplified relative to a standardized community (Figure 3). All of the primer sets had some degree of inaccuracy in replicating the mock community; by reporting which taxa each primer set is better or worse at sampling, we hope to facilitate selection of the primer set best suited for an individual system or study.

We further used each primer set to amplify 76 field samples of mixed bacterial and archaeal communities from the Santa Barbara Channel. We compared population dynamics of taxa of interest between the primer sets, testing the conclusions drawn from the mock community analysis under realistic conditions, and verifying that the majority of the biases observed with the mock communities do persist with real samples. We also analyzed population dynamics from 16S genes in 10 shotgun metagenomes, providing a primer-independent comparison for a subset of the field samples.

The field samples were also used to assess the effects of 16S primer selection on aspects of bacterioplankton community ecology. From this, we concluded that primer choice has minimal effects on community-level interpretations. All primer sets indicate similar community patterns of separation between surface and sub-euphotic zone communities and strong seasonality within the surface communities (Figure 5). Bottom-up physicochemical parameters such as chlorophyll a concentrations and chromophoric dissolved organic matter likewise show similar relationships with both community structure and populations of individual taxa between primer sets. This is important to verify for the broader community, because primer sets in use have changed rapidly over the last decade as advances in sequencing technology necessitate changes in amplicon length. These results allow us to say that the conclusions of older work in the literature can be validly compared with more recent sequencing work, even
across primer sets that are known to differ in their treatment of individual bacterioplankton populations.

**Time-series study:** Based on the findings of our primer comparison study, we were able to proceed in analyzing the time-series study of bacterioplankton 16S rRNA from the Santa Barbara Channel, which encompassed 578 samples from 43 cruises between 2010 and 2014. We focused on comparing variability in bacterioplankton community structure across spatial (horizontal and vertical) as well as temporal scales, as we often observe heterogeneity in parameters such as phytoplankton blooms across the SBC within a single Plumes and Blooms cruise. The largest phylogenetic distances between communities were observed over depth within cruises, largely driven by the higher relative abundance of archaea below the euphotic zone compared with the surface. Within depth horizons, we observed a strong annual cycle in community structure, which greatly exceeded the phylogenetic distance observed within single cruise transects.

Further analysis of the surface communities indicated that this annual cycle reflects a progression of distinct community types associated with particular physicochemical conditions. One community type was present when conditions were consistent with recent upwelling, two were associated with strong diatom-dominated phytoplankton blooms, and a fourth was observed following the upwelling-driven phytoplankton bloom season, when conditions indicated a moderate level of resources remained in the system. Two community types were present during the warm, stratified summer and fall months, one of which was present throughout the time-series and one of which was only observed starting in summer 2012, after which point the two communities alternated during stratified conditions and sometimes co-occurred within a cruise transect. A final, *Synechococcus*-dominated community type was observed only in August 2014. This is an intriguing observation, as this cruise took place as the Pacific Warm Anomaly was starting to impact the SBC, but we cannot draw strong conclusions as this portion of the time-series ended in September 2014 – we hope that the resumed time-series sampling will provide further insight into this phenomenon.
Figure 4. Results of the comparison between four 16S rRNA primer sets using a mock community of cloned 16S genes. The X-axis indicates the taxonomic identity of the cloned genes in the community, which were added in equal proportions. The Y-axis shows the log2-fold difference of the relative abundance of each cloned gene, relative to that expected, according to each primer set. When a cloned gene was not detected by a primer set, it is indicated here by a log2-fold difference of -10. Primer sets are named by the hypervariable region(s) of the 16S gene that they target.
Figure 5. Nonmetric multidimensional scaling ordination plots of bacterioplankton community structure of 76 field samples from the SBC, as assessed using four 16S rRNA primer sets. Samples are colored by month, with shape indicating sampling depth. Each primer set generated a similar V-shaped sample distribution, here oriented with the surface samples on the left on the deep-water samples on the right. Each primer set also detected similar patterns of seasonality in the surface samples. Surface samples from the upwelling months, in particular March, are most similar to the deep-water samples (the cluster of red circles in the notch of the V, which are near to the triangle-shaped samples from 75 m). Samples associated with strong phytoplankton blooms (May, the orange and yellow circles in the top left) and samples from warm, stratified summer and fall months (August through October, the blue circles at the bottom center) are both distant from the deep-water samples and from each other, regardless of primer set used to evaluate the communities.
Eukaryotic Plankton Diversity

Dr. Paul Matson is a postdoctoral scholar working with Iglesias-Rodriguez and Carlson on the Eukaryotic plankton diversity. Below is a brief description of plans and progress to date.

The goal of this work is to use next-generation sequencing technology (Illumina MiSeq) to quantify the diversity of eukaryote plankton in the Santa Barbara Channel across space and time. Eukaryotic plankton diversity and community composition will be assessed by Next Gen sequencing of 18S and 16S genes. This work is composed of three parts: (1) testing the taxonomic resolution and quantitative power of Next Gen sequencing approaches using a diverse mock community of phytoplankton of known identity; (2) applying these tools to an archival monthly time series of samples collected between 2011 and 2014; and (3) generate a new monthly time series beginning in 2016. In the past year, 24 unique phytoplankton cultures were obtained spanning a diverse range of taxonomic groups (diatoms, dinoflagellates, pelagophytes, haptophytes, raphidophytes, chlorophytes, prasinophytes, cryptophytes). Two sets of mock communities consisting of these phytoplankton strains will be used to independently assess DNA extraction and PCR amplification biases introduced in Illumina library preparation. To assess DNA extraction biases, we have assembled 4 different cell mixtures containing 7-10 phytoplankton strains of known cell abundances. To assess PCR amplification bias, each culture has been sampled individually, genomic DNA has been extracted, and we are currently cloning to obtain full length 18S amplicons from each strain. These amplicons will be combined in known proportions prior to amplification and sequencing. Assessing each source of bias independently will allow for more robust quantification of eukaryotic diversity in the archived field samples. In addition to the archived field samples, monthly sampling began again in March of 2016 and will continue for the next several years.

Ichthyoplankton

The main goal of the ichthyoplankton component of the SB MBON is to develop genetic tools to efficiently identify larval fishes collected through plankton sampling. Co-Investigator Andrew Thompson, an expert in larval fish identification, utilizes traditional morphology-based methods to identify larval fish species within plankton samples. After identification, the larvae are returned to the sample and then we test whether genetic tools can accurately identify the known species. Because most of the larval fishes in California can be identified to species based on morphology, we will be able to directly assess the efficacy of metabarcoding to assess ichthyoplankton biodiversity.

In 2015 and 2016 we collected plankton samples for this project with the help of several collaborators. First, in 2015 CalCOFI collected samples with pairovet nets from four quarterly cruises at 308 stations between San Francisco and San Diego specifically for SBC MBON. Second, in May 2016 members of the SBC MBON and the Channel Islands National Marine Sanctuary collected pairovet
samples in the Santa Barbara Channel at the Plumes and Blooms program stations. Third, in collaboration with the city and counties of Los Angeles ocean monitoring groups we collected ichthyoplankton samples from Santa Monica Bay in June 2016 and summer 2017.

We have made significant progress in morphologically identifying larvae. An SBC MBON technician has removed (i.e. sorted) larval fish from all samples, and Thompson has identified larvae to species from the January 2015 CalCOFI cruise and 2016 samples. SBC MBON postdoc Dovi Kacev, who is funded by the Southern California Coastal Water Research Project Authority (SCCWRP) and NOAA SWFSC, is now amplifying and sequencing DNA from individual fish larvae and mixed samples. Kacev began working on the project in January 2016 and will be analyzing and reporting on sequence data in year 4.

C. Implement a tradeoff framework that optimizes allocation of sampling effort.

A major goal of the project is to compare the effectiveness of different methods of biodiversity sampling, exploring how the optimal mix of sampling depends on the goals of monitoring and the resources available. We will also consider how the optimal mix will change as technology improves (e.g., lower cost and better accuracy of genetic and computational methods). Much of this work will not begin until biodiversity data derived from a range of sampling methods can be assimilated and analyzed in Year 4. However, we have started developing the modeling framework for these analyses, testing methods on simulated biodiversity datasets. Development of methods for analyzing tradeoffs will continue in year 4, and analysis of real data will begin as the data become available. Co-PI Andrew Rassweiler is leading this part of the project and will continue to do so in his new capacity as an Assistant Professor at Florida State University. Rassweiler is currently recruiting a data analyst to assist with this project component.

We are collaborating with the Sanctuaries MBON to compare data collection technologies for use in this model and general cross-validation. In July 2017 we conducted diver surveys with the SBC LTER program of 40 transects off Santa Barbara (two weeks of daily diving by three teams of 4 divers each). We photographed the transects for image analysis, and we collected triplicate eDNA samples on each transect. We are now shipping the eDNA samples to MBARI and will collaborate to compare these disparate datasets for purpose of evaluating their utility to future MBON efforts.

Data Management and Communication
On the SBC MBON web site, we have installed a data catalog linking to data products in the EDI repository and demonstration map application based on these datasets. We are working with Sanctuaries MBON and Dr. Ben Best on X-
MBON tools for data display and product generation. Dr. Best is housed in a SBC MBON office at UCSB and we work with him dynamically.

SBC MBON organized a cross-MBON DMAC workshop, 2016 June 7-9. The goals were to (a) establish preconditions for cross-MBON harmonization by comparing expected re-use scenarios with local processing and management systems, and (b) outline potential frameworks for important variables or vocabularies. Workshop products include:

1. Document describing anticipated data needs of potential data customers with high-level requirements, using NMS condition reports as an example,
2. Schematic of local processing streams and how these translate to a ‘pipeline’ of data for a customer, and
3. Initial list of dictionaries that have the potential to meet MBON needs for a structured measurement vocabulary.

The Axiom “research workspace” became a DataONE member node in 2017, which means that their ISO 19110 and 19115-2 metadata are now indexed alongside SBC MBON EML metadata (available through DataONE since 2016). SBC MBON personnel already employ mechanisms to query projects in DataONE, and has shared that knowledge with the MBON partners managing data through Axiom, to plan for MBON-wide queries.

Within the broader community, the SBC MBON is integrating its activities into the international Group on Earth Observations Biodiversity Observation Network (GEO BON) by joining GEOBON Working Group 8 (Data integration and interoperability, informatics and portals), and adhering to established guidelines for data management and dissemination. Further, DataONE leads a broad based group of scientists, data managers and knowledge modelers in computer science to develop standardized measurement descriptions in an ontological framework. We already participate in this project through our LTER collaboration, and SBC MBON datasets are planned to be use cases for the development of descriptions for primary measurements associated with the GEOBON Essential Biodiversity Variables (EBVs). Further work requires additional resources (pending).

Data manager O’Brien chairs the ESIP cluster “Sustainable Data Management” which investigates mechanisms to promote increased collaboration and coordination in the area of environmental data management to benefit both research networks and individual investigators.


We are coordinating SBC MBON R code design and sharing with the larger community of ecosystem science data managers, including LTER and DataONE. We participate in standard communication mechanisms, e.g., ESIP, which is compatible with IOOS and NASA practices.

In collaboration with LTER, we provided use case datasets for NOAA’s development of a mechanism to ingest EML-described datasets to the ERDDAP
system supported by NOAA (https://coastwatch.pfeg.noaa.gov/erddap/download/EDDTableFromEML.html, https://coastwatch.pfeg.noaa.gov/erddap/images/erddapTalk/eml_knb_erddap.pptx). The system was successfully tested with SBC MBON reef community datasets and the SCCOOS ERDDAP. A formal process for transfer awaits further discussion and resources (pending).

A priori, the EML-described data packages created by SBC MBON are nearly complete Darwin Core Archives (DwC-A), the required format for OBIS and GBIF data. With input from the DwC-A designers, we are determining the additional content or code required for our EML-described data packages to be directly ingested by OBIS or GBIF, which would significantly expedite the process compared to current channels.

Outreach and education

SBC BON participated in the 2017 NGSS Local Phenomena Summer Institute funded by a NOAA B-Wet Grant and supported by the Channel Islands National Marine Sanctuary, UCSB Marine Science Institute Oceans to Classrooms, and the Santa Barbara Unified School District. Miller presented the keynote address, an overview of the Biodiversity Observation Network concept and how it links to ocean science in California. He used MBON science activities to illustrate the value of hypothesis testing and rigorous use of the scientific method in science and education. MBON staff Sarah Sampson and Devin Spencer conducted a half-day workshop for the participating teachers on using the online image analysis system BisQue for education.

We (Emmett Duffy, Miller, Muller-Karger) organized a session at the 2016 ASLO Aquatic Sciences meeting on Ocean Observing and Data Management, titled MBON Voyage: Integrating marine biodiversity into ocean observing systems.


Miller and Reed participated in the Channel Islands National Marine Sanctuary Condition Report meeting in June 2016, and SBC MBON collated and contributed biodiversity data to inform the report. We are now finalizing those data for the publication of the CR in 2017.

We (Siegel, Muller-Karger, Schimel) organized an NCEAS workshop "Prospects and Priorities for Satellite Monitoring of Global Marine Biodiversity" working group with Dave Siegel, David Schimel, Frank Davis, and Ryan Pavlick, that was held
in June 2016. This year we have contributed to a manuscript based on the results of that workshop, which is in review at Ecological Applications. The paper gives recommendations for future coastal remote sensing needs.

Other project presentations are listed below.

**Presentations:**


Dugan, J.E., R. Miller, K. Emory, D. Hubbard, 2017, Patterns and processes affecting the transport, retention and fate of trophic subsidies to sandy beach ecosystems. Poster presented at 2017 Annual Meeting of the Ecological Society of America Meeting. Portland, Oregon, USA. August 11, 2017

John A. Hildebrand, Using sound to study Baleen Whale Populations, SOMMEMA XXXV, La Paz, BCS, Mexico, May 2-5, (2016).


Rognstad, R., A. Rassweiler, R. Miller, D. Reed, 2017, Detecting and forecasting climate effects on patterns of biodiversity and productivity in West Coast Sanctuaries. NOAA NCOS Ecological Forecasting project meeting, Santa Barbara CA, July 28, 2017.


**Publications:**


