OCS Study
BOEM 2019-063

DOI Partnership: Distinguishing between Human and Natural Causes of Changes in Nearshore Ecosystems Using Long-term Data from DOI Monitoring Programs

US Department of the Interior
Bureau of Ocean Energy Management
Pacific OCS Region

BOEM
Bureau of Ocean Energy Management

OCS Study
BOEM 2019-063

DOI Partnership: Distinguishing between Human and Natural Causes of Changes in Nearshore Ecosystems Using Long-term Data from DOI Monitoring Programs

June 2020 (Revised)

Authors:
Daniel C. Reed and Andrew Rassweiler

Prepared under Cooperative Agreement M11AC00012
By
Marine Science Institute
University of California, Santa Barbara
Santa Barbara, CA 93106

US Department of the Interior
Bureau of Ocean Energy Management
Pacific OCS Region

BOEM
Bureau of Ocean Energy Management

DISCLAIMER

Study collaboration and funding were provided by the US Department of the Interior, Bureau of Ocean Energy Management (BOEM), Environmental Studies Program, Washington, DC, under Agreement Number M11AC00012. This report has been technically reviewed by BOEM, and it has been approved for publication. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the US Government, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

REPORT AVAILABILITY

To download a PDF file of this report, go to the US Department of the Interior, Bureau of Ocean Energy Management Data and Information Systems webpage (https://www.boem.gov/Environmental-Studies-EnvData/), click on the link for the Environmental Studies Program Information System (ESPIS), and search on 2019-063.

CITATION

Reed DC, Rassweiler A. 2020. DOI Partnership: Distinguishing between human and natural causes of changes in nearshore ecosystems using long-term data from DOI monitoring programs. Camarillo (CA): US Department of the Interior, Bureau of Ocean Energy Management. OCS Study BOEM 2019-063. 117 p. [Originally published September 2019. Revised June 2020.]

ABOUT THE COVER

ACKNOWLEDGMENTS

This research relied on data collected by the National Park Service, the United States Geological Survey, the Partnership for Interdisciplinary Studies of Coastal Oceans, and the Santa Barbara Coastal Long Term Ecological Research project. We thank these programs and all the staff and volunteers involved in collecting and assembling the data. We also think the Santa Barbara Channel - Marine Biodiversity Observing Network project for assistance with data synthesis. We thank the many scientists who participated in some aspect of this work including Donna Schroeder, Jim Estes, David Kushner, Mike Kenner, Josh Sprague, Dan Okamoto, Mark Novak, Tim Tinker, Jenn Caselle, Ally Dubel, Li Kui, Shannon Harrer, Jacob Staines, and Katie Davis.

Contents

List of Abbreviations and Acronyms

ASCII	American Standard Code for Information Interchange
BACI	Before-After-Control-Impact
BOEM	Bureau of Ocean Energy Management
BON	Biodiversity Observation Network
BOT	Bottom
CDIP	Coastal Data Information Program
CAN	Canopy
CINP	Channel Islands National Park
CNMD	Canopy-Midwater
dbMEM	distance-based Moran's Eigenvector Maps
DOI	US Department of the Interior
ESP	Environmental Studies Program
ESPIS	Environmental Studies Program Information System
FUL	Full
KFM	Kelp Forest Monitoring program
LTER	Long Term Ecological Research program
OCS	Outer Continental Shelf
PCNM	Principal Coordinates analysis of Neighborhood Matrices
PDF	Probability Density Function
PISCO	Partnership for Interdisciplinary Studies of Coastal Oceans
RDA	Redundancy Analysis
RDFC	Random Diver Fish Counts
ROMS	Regional Ocean Modelling System
RPC	Random Point Contact
SCB	Southern California Bight
SNI	San Nicolas Island monitoring program
SST	Sea Surface Temperature
UPC	Uniform Point Contact
USFWS	US Fish and Wildlife Service
USGS	United States Geological Survey

1 Executive Summary

Monitoring and predicting the potential impacts of outer continental shelf (OCS) energy production on nearshore ecosystems requires an ability to distinguish between changes caused by natural processes and those caused by human activities. The ability to distinguish such changes in turn requires long-term, spatially extensive data to describe natural patterns of temporal and spatial variation in species abundances and the environmental factors that influence them. This is particularly true for giant kelp forests, which are highly productive and diverse ecosystems in temperate regions that fluctuate greatly in space and time. These systems are highly valued for the milieu of goods and services they provide to society and there is general interest in minimizing anthropogenic activities that adversely affect them. The purpose of this project was to partner with agencies in the Department of the Interior (DOI) to document, integrate and analyze data produced from long-term kelp forest monitoring programs to improve our understanding of the causes and consequences of change in these iconic ecosystems.

The primary objectives in this collaborative partnership were fourfold: (1) Work with two Department of the Interior agencies to assimilate, document and published their long-term (30 + years) data sets pertaining to kelp forest community structure at the northern Channel Islands and San Nicolas Island; (2) Expand the spatial scope of these data sets by integrating them with other long-term kelp forest monitoring programs in the region and with appropriately temporally and spatially scaled environmental data to produce a data resource with unrivaled temporal, spatial and taxonomic scope; (3) Analyze integrated data sets across multiple spatial and temporal scales to ascertain patterns of variation in population and community dynamics and to identify key environmental and anthropogenic factors that drive them; (4) Use the fully integrated data sets to collaborate with BOEM partners and other BOEM-funded programs on issues relevant to BOEM's mission

The value of this project to BOEM lies in its ability to assist managers in detecting and evaluating possible impacts from offshore energy activities, and in developing options to mitigate these impacts. In addition, identification of patterns in these data sets will aid in predicting potential ecosystem impacts due to climate change and advancing adaptive management, both of which are goals central to DOI stewardship responsibilities and trust resources.

One lesson of this project is that the time and expense necessary to make data visible, easy to use and easy to combine with other data should not be underestimated, but that such efforts can have very positive results on how widely that data is accessed and used. When long term ecological monitoring continues for decades, there is a danger either that too little metadata is assembled, preventing interpretation of the data, or that so much metadata is assembled that the volume of information becomes a barrier to understanding what was done. The process of publishing data sets forces metadata to be constructed in a standard format and ensures that an appropriate level of detail is maintained. When combining data sets, we note that decisions about taxonomic resolution can present obstacles to data synthesis, especially when organisms are separated into life stages or size classes. If a new monitoring project is proposed, it is worth investing time up front to ensure that the newly collected taxa groups and size classes will be compatible with existing data sets.

We show that the combined data sets we produce have considerable potential for detecting impacts in this region. Although there are statistical challenges to detecting localized impacts, Before-After-Control-Impact analyses or similar approaches can reliably detect impacts of moderate size and severity particularly if several species are similarly impacted and thus included in the analysis. However, we also reveal that the spatial and temporal dynamics of this marine system can result in extremely high false positive rates if temporal autocorrelation is not properly accounted for. Because adjusting for such

temporal effects requires long-term monitoring data, this emphasizes the continued value of DOI monitoring for impact analysis in the region.

Our analyses of the drivers of spatial variation within the region indicate that moderate scale variation (65 km+) is relatively predictable based on available environmental covariates. This suggests that detection of moderate or large sized impacts such as oil spills could be further improved by considering such covariates in analyses. Unfortunately, local scale impacts such as cable burial occur on a background of high small-scale natural variability. This small-scale variability was not well predicted by any of the covariates we examined, and would rarely be co-located with existing monitoring sites. These results emphasize the challenge of reliably estimating small impacts without direct collection of ecological data before and after an anthropogenic disturbance.

2 Project Description

2.1 Introduction

A primary goal of environmental impact assessments is to quantify trends in species responses following a perturbation. Evaluation or predictions of the effects of anthropogenic perturbations typically are made based upon the observed or hypothesized magnitude and direction of these trends. To be useful for impact assessment analyses, however, trends observed to occur at local spatial scales (such as near an offshore energy facility) must be placed within a larger regional context. In the case of predicted trends, these changes should reflect not only the hypothesized effects due to the perturbation, but also those effects likely to occur as a consequence of unrelated, natural changes in regional environmental conditions. Consequently, studies examining long-term temporal and large-scale temporal and spatial gradients in species composition and abundance are needed to distinguish anthropogenic effects from trends arising from natural occurring phenomena. Such temporally and spatially extensive studies provide both a regional context within which observed trends can be placed, and an opportunity to explore responses of species to both natural environmental and human-induced perturbations.

Existing offshore oil and gas activities and future renewable energy projects can directly and adversely affect nearshore ecosystems. As mandated in the Outer Continental Shelf Lands Act (43 USC §1346(b)), part of BOEM's mission is to conduct additional studies, subsequent to leasing and developing any area or region, to provide time-series and data trend information for the purpose of identifying any significant changes in the quality and productivity of the marine environment. Ocean energy development activities that affect the nearshore may include: (1) direct alteration of habitat through the installation, maintenance, and/or removal of foundations, platforms, pipelines, cables, anchors, and other structures; (2) release of contaminants into the marine environment by oil spills and discharges; (3) decreased water quality via sediment disturbance during anchoring, dredging, etc.; and (4) onshore activities that result in erosion or spillage into the nearshore environment. Thus, in order to predict the potential environmental impacts of human activities, there is a need to describe and understand the natural dynamics of coastal ecosystems. This need is particularly pressing for State natural resource managers, because, by definition, nearshore ecosystems reside in State waters.

The purpose of this project was to document, integrate and analyze data produced from long-term monitoring programs overseen by DOI and other entities to improve our understanding of the causes and consequences of change in giant kelp forests, highly productive and diverse nearshore ecosystems that occur along the Pacific coast of the US and other temperate regions of the world. The value of this project to BOEM lies in its ability to assist managers in detecting and evaluating possible impacts from offshore

energy activities, and in developing options to mitigate these impacts. In addition, identification of patterns in these data sets will aid in predicting potential ecosystem impacts due to climate change and advancing adaptive management, both of which are goals central to DOI stewardship responsibilities and trust resources.

Our primary objectives in this collaborative partnership were fourfold: (1) Work with two Department of the Interior agencies to assimilate, document and published their long-term (30 + years) data sets pertaining to kelp forest community structure at the northern Channel Islands and San Nicolas Island; (2) Expand the spatial scope of these data sets by integrating them with other long-term kelp forest monitoring programs in the region and with appropriately temporally and spatially scaled environmental data to produce a data resource with unrivaled temporal, spatial and taxonomic scope; (3) Analyze integrated data sets across multiple spatial and temporal scales to ascertain patterns of variation in population and community dynamics and to identify key environmental and anthropogenic factors that drive them; (4) Use the fully integrated data sets to collaborate with BOEM partners and other BOEM-funded programs on issues relevant to BOEM's mission.

2.2 Data acquisition, documentation and publication of DOI data sets

The focus of this element of the project was on working with two DOI agencies, the National Park Service (NPS) and the United States Geological Survey (USGS) to assemble, document and publish their long-term kelp forest monitoring data sets to facilitate their use by academic, governmental and privately funded scientists.

The mission of the NPS is to manage diverse park resources unimpaired for the enjoyment of future generations. To help meet this challenge, the NPS developed the Natural Resources Inventory and Monitoring Program to provide a broad-based understanding of the status and trends of park resources as a basis for making decisions and working with other agencies and the public. The NPS initiated the Kelp Forest Monitoring (KFM) program in 1981 to determine the health of kelp forests in the Channel Islands National Park, California. The KFM program seeks to establish the normal limits of variation of population parameters, to provide early warnings of abnormal conditions, and to identify possible agents of abnormal change. Permanent transects are used to monitor 70 species of algae, fish and invertebrates or groups of taxa at 33 sites within the National Park. Sixteen original sites were established between 1982 and 1986, one site was established by a commercial fisherman with funding from Santa Barbara County in 2001, and an additional 16 sites were established in 2005 as part of a fine scale study of marine protected areas that were established by the State of California in 2003.

In order to understand the ecological consequences of translocating a southern sea otter population to San Nicolas Island, the United States Fish and Wildlife Service (USFWS) in cooperation with the University of California, Santa Cruz established the San Nicolas Island monitoring (SNI) program. The program has monitored six shallow subtidal sites around the island since 1980, and a seventh site since 1986. Surveys collect information on the abundance and size of algae, macroinvertebrates and reef fishes in permanent plots. The sites have been sampled twice a year in spring and fall with few exceptions. The SNI program is managed and directed by the USGS Western Ecological Research Center stationed at the University of California, Santa Cruz.

This portion of the project entailed: (1) Working with NPS and USGS scientists to acquire and assimilate all the data from their monitoring programs; (2) Employing data quality control and assurance measures to improve the accuracy and usability of the data; (3) Developing fully documented metadata in

accordance with standards endorsed by national environmental data repositories; and (4) Publishing final data and metadata products in the peer-reviewed scientific literature.

The KFM and SNI data collections consist of multiple data tables that reflect different methods for sampling different types of organisms. For each collection we produced fully documented metadata for each data table and published the two data collections (KFM and SNI) as independent data papers in the journal *Ecology*, one of the most highly cited ecological journals in the world. These two papers (Kushner et al. 2014 and Kenner et al. 2014) are included as Appendix A and B of this report. The published products are readily accessible to scientists and the public at large and are actively being used by the science community to provide insight into the patterns and causes of change in kelp forest ecosystems within the Southern California Bight, and to provide a general framework for understanding the processes controlling the dynamics of giant kelp forest communities throughout their broader geographic distribution.

The publication of these data sets as data papers in *Ecology* has resulted in increased visibility and use of the data. For example, since publication (2013-2019) the KFM data has been cited in published scientific work 23 times, at a rate of 3.8 citations per year (Google Scholar search on 9/9/19). The SNI data has been cited 11 times since the data was published, a rage of 1.8 publications per year (Google Scholar search on 9/9/19). In each case, these citation rates are much higher than prior to publication. We are aware of the KFM data being cited in publications 24 times between 1982 and 2013 (a rate of 0.8 per year; See Appendix A section V.C for a list of these publications). The SNI data was cited in publications 11 times in that period (a rate of 0.4 per year; see Appendix B section V.F for a list of these publications). For both data sets, the citation rate after we published the data was almost 5 times higher than the rate in the prior years. Some of this may reflect the natural increase in value and visibility of these data sets as they increase in temporal length, but this is still an impressive increase.

2.3 Integrating DOI data sets with data from different monitoring programs to facilitate analyses across a broad range of temporal, spatial and taxonomic scales

A primary objective of the DOI partnership project was to increase the utility of large data sets and the programmatic ability of BOEM to assess changes in nearshore communities and species at risk in California through the development and curation of novel data products. The first step in this endeavor was to formalize and publish data from two long-term DOI kelp forest monitoring programs as described in Section 2.2. Once this step was completed we sought to improve the spatial resolution and the regional scope of these data sets by integrating them with an additional 11 reefs monitored by the Santa Barbara Coastal Long-Term Ecological Research Program (LTER) and 25 reefs monitored by the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO). The Santa Barbara Coastal LTER (http://sbc.lternet.edu/) is an interdisciplinary research program focused on developing a predictive understanding of the processes that structure kelp forest ecosystems under conditions of changing climate and human use. It is a member of the U. S. National Science Foundation's LTER Network charged with providing the scientific community, policy makers, and society with the knowledge and predictive understanding necessary to conserve, protect, and manage the nation's ecosystems, their biodiversity, and the services they provide. The LTER Network was founded on the principal that long-term research is needed to help unravel the principles and processes of ecological science, which frequently involves long-lived species, legacy influences, and rare events. PISCO (http://www.piscoweb.org/) is an academic consortium led by scientists from four core campuses: Oregon State University; Stanford University's Hopkins Marine Station; University of California, Santa Cruz; and University of California, Santa Barbara. Its mission is to conduct long-term research to advance understanding of the coastal ocean

within the California Current Large Marine Ecosystem and inform management and policy. Giant kelp forests and the rocky intertidal serve as the focal ecosystems for their research.

Finding a common data structure and comparable reporting metrics is essential for integrating data sets into a common framework. This is often challenging when integrating data from different sources as methods typically differ as do the data structures. Such is the case for the KFM, SNI, LTER and PISCO data as the sampling methods, frequency and timing of data collection and taxonomic resolution were individualized to meet the mission, goals and operating budgets of the respective monitoring programs. Integrating data from KFM, SNI, LTER and PISCO required us to develop a common data framework in which all data sets overlapped, and allowed users to consider any combination of data sources in an analysis. An additional problem with past efforts to join data sets is that the combined data are only valid to the point in time that the data were joined. Therefore, we developed programming scripts that make it possible to automatically update the combined data set once the source data have been updated.

A key challenge to producing data sets was to determine what biological variables have been measured in comparable ways by each program. In particular, the taxonomic resolution and the range of species counted varied widely from program to program. Our focus here was to produce data sets that are ready for analysis; thus, it is important that each species (or taxonomic group) included be counted using comparable methods by all data sources that are being synthesized. As a consequence, the taxonomic list for each integrated data set can only include species that are shared across all source data. Each data set included lacks some taxa (or taxonomic resolution) contained in the other data sources; therefore, the greater the number of sources included, the smaller the resulting usable taxa. We therefore include five versions of the biological data here (Appendix C), a core data set including all four monitoring programs, and four additional data sets - representing all three-way combinations of these sources. Although we do not include all pairwise combinations in this report, we developed computer code (Also included in appendix C) which will produce any pairwise combination except for the LTER-SNI combination (this pair is unlikely to be used in isolation because of the geographic layout of the sites, with LTER in the north, SNI in the south and PISCO and KFM in between).

Each version of the data set includes three biological data tables, a “fish” table representing counts of fish and the area sampled, a “quadswath” table representing counts of algae and invertebrates and the area sampled, and a “benthic” table containing the percent cover of macrophytes and sessile invertebrates. We also present diversity data for each of the three types of data tables that include values for richness, Shannon, and Simpson’s diversity indices calculated from the included taxa. Throughout, data are presented from 2000 onward, as both PISCO and LTER data do not start until the early 2000s.

To supplement these biological data, we developed environmental data sets that have been processed to match the spatial and temporal scales of the biological data. The environmental data sets include information about sea surface temperature, wave climate, biomass of giant kelp, and sea surface chlorophyll in the area immediately surrounding a sampling site. Environmental data are averaged over two temporal scales; the six months prior to a sampling date, and the 12 months prior to a sampling date. Environmental data sets and metadata are included in Appendix C.

In addition to supplying analysis-ready data sets, we provide the computer codes used to produce them (Appendix C). At the top of each of these scripts are user-friendly parameters that allow the user running the script to select which of the four source data sets are included in the output (using a zero to exclude the data and a 1 to include it). The fish and invertebrate processing scripts also include a parameter that allows the user to set the balance between keeping more observations and obtaining better taxonomic resolution. These codes curate the data, and draw on a large integrated data table produced by the Santa Barbara Channel Marine Biodiversity Observing Network (http://sbc.marinebon.org/).

2.4 Patterns of spatial and temporal variability in kelp forest communities and environmental and anthropogenic factors that drive them

We used the integrated data set compiled from KFM, SNI, LTER and PISCO data and publically available environmental data to explore patterns of spatial and temporal variability in giant kelp forest communities and to identify environmental and anthropogenic factors that may drive dynamics. We focused our efforts on two areas. First we conducted numerical analyses using observational surveys for a broad spectrum of nearshore marine taxa to determine how nearshore species vary in their value for estimating and detecting impacts. Second we used redundancy analysis (RDA) to independently model the spatial structure of fishes, invertebrates and macroalgae to determine the spatial scales over which different environmental factors affect their species composition and abundance. We briefly summarize these activities below and present detailed descriptions of these efforts in Appendices D and E.

2.4.1 Detecting human impacts in a highly variable ecosystem using long-term monitoring data

Evaluating the effects of unexpected human impacts is a major challenge for applied ecology. When ecological monitoring data are available, a Before-After-Control-Impact (BACI) analysis is often applied, which can control for natural spatial and temporal variation to better isolate an impact and help to distinguish it from other sources of variation. The virtues of BACI-type designs are well established, but the degree to which the approach may be compromised by patchy population distributions and dynamics has not been systematically explored in nearshore marine systems. Here we quantified the potential for BACI analyses of long term monitoring data to detect anthropogenic impacts with the goal of better understanding its reliability when applied to particular species (or sets of species). We started with a spatially and temporally extensive monitoring data set from reefs in southern California and simulated impacts of different spatial-scales and severities. We found the BACI approach had substantial potential to detect local reductions in the population sizes of 28 species of fish, invertebrate and macroalgae, as long as temporal autocorrelation was accounted for in the statistical model (neglecting this temporal structure led to very high false positive rates). However, the power to detect impacts varied widely from species to species and false positives exceeded target levels in some cases because of spatial synchrony in dynamics (false positives rates were lower if impact sites were scattered across the region, rather than clustered,). Impacts were most likely to be detected in species with relatively stable, widely distributed populations, but the same factors increased false positive rates. Finally we found that combing even a small number of impacted species in our analyses greatly improved the method's success rates, but that the choice of species to be combined was crucial. These results provide guidance to long-term monitoring projects by suggesting characteristics of species that would serve as good indicators of an impact, and by setting expectations for what sizes and severities of impacts are likely to be detectable. A paper detailing this work is in preparation for publication and is included in its entirety as Appendix D.

2.4.2 Scale-specific drivers of kelp forest communities

Identifying spatial scales of variation in natural communities and the processes driving them is critical for obtaining a predictive understanding of biodiversity. This is particularly true for highly productive and diverse marine communities associated with shallow subtidal reefs distributed across fragmented landscapes. Such is the case for kelp forests in southern CA and other regions of the world. Using the integrated data sets described above, we combined long-term subtidal community surveys from 86 sites with detailed environmental data to determine what structures assemblages of fishes, invertebrates and algae at multiple spatial scales. We identified the spatial scales of variation in species composition using a hierarchical analysis based on eigenfunctions, and assessed how sea surface temperature (SST), water column chlorophyll, giant kelp biomass, wave exposure and oceanographic connectivity contributed to

community variation at each scale. Spatial effects occurring at multiple scales explained 60% of the variation in fish assemblages and 52% of the variation in the assemblages of invertebrates and algae. Most variation occurred over broad spatial scales (> 200 km) consistent with spatial heterogeneity in SST and oceanographic connectivity, while this latter also explained community variation at medium scales (65-200 km). Small scale (1-65 km) community variation was substantial, but not linked to any of the measured drivers. Conclusions were consistent for both reef fishes and benthic invertebrates and algae, despite sharp differences in their adult mobility. Our results demonstrate the scale dependence of environmental drivers on kelp forest communities, showing that most species were strongly sorted along oceanographic conditions over various spatial scales. Such spatial effects must be integrated into models assessing the response of marine ecosystems to climate change. A peer-reviewed paper detailing this work by Lamy et al. (2018) was published in the journal *Oecologia* and is included in its entirety as Appendix E.

Taken together, our results emphasize the value of multisite marine monitoring projects for detecting impacts and suggest ways that better understanding of the drivers of variability in community structure might improve impact assessment. Our analysis using existing data emphasizes that even relatively simple statistical approaches such as BACI analyses can detect impacts of moderate spatial scale such as oil spills (section 2.4.1), particularly if the disturbances impact multiple species simultaneously. Natural spatial variation occurring synchronously on scales at or above 65 km can complicate such analysis, but we find that much of this variation can be predicted based on environmental factors operating at those scales such as sea surface temperatures and waves (2.4.2). This suggests the potential value of more sophisticated impact analyses that incorporate physical covariates to account for naturally-driven changes in community structure. However, we also document high levels of small-scale community variation which were not reliably predicted by any of the drivers we examined. This small-scale variation presents a real challenge for estimating the effects of more localized disturbances such as cable burial. Our BACI analyses indicate that such small impacts would rarely be resolved with existing data because they would be unlikely to occur near one of the monitoring sites, and the high degree of local variability we observe means that we have limited ability to predict the community that would have occurred at a sites in the absence of human activity. If the location of small-scale impacts can be predicted in advance, for example with cable as with cable burial, then collecting pre-impact data on ecological communities at those locations may be the only way to obtain reliable estimates of those impacts.

2.5 Integration with other BOEM-funded studies

2.5.1 Estimates of wave energy and its effects

A key integrative activity has been to combine the biological and environmental data synthesized here with wave energy estimates. The development and validation of wave energy estimates was done under a separate BOEM award, *Predicting the Consequences of Wave Energy Absorption from Marine Renewable Energy Facilities on Nearshore Ecosystems* (PC-13-05). This project evaluated the spatial and temporal variation in wave height for the study system in the broader context of the Southern California Bight. This helped identify two wave seasons as well as the distribution of exposed and protected sites in the region. We deployed, retrieved and processed data from a new, low-cost pressure sensor at 32 sites around the Channel Islands where long-term kelp forest monitoring occurs (a combination of KFM and SNI sites). We then compared hourly data from the sensors with hourly hindcasts from the Coastal Data Information Program (CDIP; https://cdip.ucsd.edu/). This made it possible to generate simple statistical corrections to CDIP hindcasts to help improve their accuracy. With this correction, we generated hourly hind-casts for these 32 sites back to 2000.

These high frequency wave data were then summarized at an annual timescale to better match the timing of biological sampling. This allowed us to extrapolate the results of Lafferty and Morton (2018) to product hindcasts for an additional 56 sites monitored by PISCO and LTER, bringing our total number of sites to 88. At each of these sites we produced hindcasted estimates of mean wave height, energy, and orbital velocity calculated at 6- and 12-month intervals before biotic sampling. The hind-cast wave data are publicly available on the USGS ScienceBase-Catalog (Lafferty et al. 2018) and can be used to address questions about how waves affect nearshore communities in the Southern California Bight.

Lastly, we integrated the wave hindcast data with the biological and environmental data sets described in Section 2.3 to explore the degree to which wave height and other abiotic factors drive species dynamics at the 88 sites. This integrated data set was the primary data source for our investigations on scale-specific drivers of kelp forest communities (Lamy et al. 2018; see Section 2.4.2).

2.5.2 Marine Biodiversity Observation Network

We collaborated extensively with the study *A Demonstration Marine Biodiversity Observation Network (BON) for Ecosystem Monitoring* (PC-15-05) in their efforts to integrate data from diverse sources across the region into a common data structure. Our work in preparing the KFM and SNI data sets for publication (Section 2.2) paved the path for the BON to produce several integrated data sets of nearshore biological data from the region (Miller et al. 2016a, b, Miller et al. 2017a, b). We also collaborated with the BON to add the wave data from Lafferty et al. (2018) to the integrated biological data set in this report and into a separate BOEM-funded integrated data table produced by the Santa Barbara Channel Marine BON.

A limitation of these BON-generated data sets for researchers seeking to use them is that the data are kept in a form as close as possible to the original, which creates substantial hurdles to using data from multiple sources. Data users must either limit themselves to a single data source within the integrated data set, or understand the details of how the data collection methods differ and make their own decisions about how to combine the data. In cases where the data sources use different methods/rules to aggregate species, life stages or size classes into a taxonomic group, the user must consider these differences and decide whether the data are comparable and if so how the data set should be aggregated. In many cases a species may be sampled by multiple methods within a single data set and the user must decide which method is more reliable and comparable with data from the other sources being combined. Scaled across hundreds of species and four different sampling programs, each with its own complex history, this decision-making process represents months of work.

As described in Section 2.3, we created curated data sets and associated data processing scripts to assist the user in overcoming the above hurdles associated with using data from multiple sources, which enables them to address their questions more immediately. This is a particularly important product for BOEM and other DOI agencies, as a primary goal of this project was to ensure that relevant nearshore biological data are ready for immediate analysis to support decision-making or impact assessment related to offshore energy production.

3 References

Lafferty KD, Morton DN. 2018. Site table and bias corrections for Coastal Data Information Program (CDIP) hind casts at the California Channel Islands: U.S. Geological Survey data release, https://doi.org/10.5066/P9FV1X7H.

Lafferty KD, Morton DN, Gotschalk CC, Henderikx F, Rassweiler A, Washburn L. 2018. Hourly wave height and period hindcasts at 32 sites throughout the Channel Islands National Park and San Nicolas Island from 2000-2017: U.S. Geological Survey data release, https://doi.org/10.5066/P91KH2Q7.

Miller R, Rassweiler A, Reed D, Lafferty K, Kui L, O'Brien M. 2016. Santa Barbara Channel Marine BON: Integrated quad and swath cover. Environmental Data Initiative. https://doi.org/10.6073/pasta/51d2db26e90d4b8687db81fb40bc58c4. data set accessed 8/19/2019.

Miller R, Rassweiler A, Reed D, Lafferty K, Kui l, O'Brien M. 2016. Santa Barbara Channel Marine BON: Integrated benthic cover. Environmental Data Initiative. https://doi.org/10.6073/pasta/bf143fa962e1edb822847bc0ee90c2f7. data set accessed 8/19/2019.

Miller R, Rassweiler A, Reed D, Lafferty K, Kui L, O'Brien M. 2017. Santa Barbara Channel Marine BON: Integrated fish. Environmental Data Initiative. https://doi.org/10.6073/pasta/1345f0148e6dfe4df9065e223b4dd783. data set accessed 8/19/2019.

Miller R, Rassweiler A, Reed D, Lafferty K, Kui L, O'Brien M. 2017. Santa Barbara Channel Marine BON: Integrated taxa. Environmental Data Initiative. https://doi.org/10.6073/pasta/d09d4bfd54e6d4e490b4cc34731d808e. data set accessed 8/19/2019.

Appendix A

All files and folders referenced in Appendix A are found in the file:

OCS Study BOEM 2019-063 Appendix A files.zip

A multi-decade time series of kelp forest community structure at the California Channel Islands

David J. Kushner, Andrew Rassweiler, John P. McLaughlin, Kevin D. Lafferty

2013. *Ecology* 94: 2655

INTRODUCTION

The ocean waters surrounding the California Channel Islands include many shallow nearshore reefs. Giant kelp (*Macrocystis pyrifera*) is commonly, but not always, a dominant fixture here (Engle 1994, Cavanaugh et al. 2011). Biotic communities on these reefs vary in space due to differences in depth, substrate, temperature, nutrients, and exposure to waves. In addition to seasonal variation (not measured in this study), rocky reef communities may vary from year to year due to recruitment events, physical disturbance from waves, fluctuations in the climate (including El Ninos and shifts in the Pacific Decadal Oscillation) and community dynamics such as disease outbreaks or changes in grazer behavior (Lafferty 2004).

Humans also affect this ecosystem in space and time. Most of the area is open to commercial and recreational fishing (e.g, Schroeter et al. 2001), and it also supports other recreational activities such as diving and boating. Marine communities may also be affected by chronic and acute pollution from mainland waste disposal and adjacent offshore petroleum development. The five northern California Channel Islands are primarily under the management and jurisdiction of the National Park Service and Nature Conservancy, while the California Department of Fish and Wildlife has jurisdiction over the living marine resources within 3 miles of the island.

To assist in protecting and managing these nearshore ecosystems, the National Park Service created the kelp forest monitoring program (KFM) in 1982. Every summer since then, the KFM has tracked the populations of key species of algae, invertebrates and fishes at locations throughout the park. The resulting data on subtidal community structure and dynamics are nearly unparalleled in their completeness and spatio-temporal extent. These data have contributed substantially to management actions, such as the State of California's decision to close several abalone fisheries in 1997 and the decision to create marine protected areas within the National Park in 2003. The data have also been the basis of a number of general ecological insights (e.g., Reed et al. 2000, Lafferty 2004, Byrnes et al. 2006).

Despite its proven usefulness, several characteristics of this monitoring data set have made it challenging to analyze. The KFM has refined its methods several times (e.g., Davis et al. 1996). These improvements have focused on obtaining the best data possible given a fixed amount of effort, but have come at some

cost to the overall consistency and simplicity of the data set. Researchers wishing to use these data have needed to work closely with the program's extensive documentation, understanding how and when sampling size changed, what species have been counted with which methods in which years, and how choices about which species to count and at what taxonomic resolution have changed over time.

In this data paper, we have summarized observations from five different sampling methods into three core data sets that record the abundance of more than 60 key taxa over a 30-year period. We have corrected for changes in sampling method, changes in taxonomic resolution and other inconsistencies, seeking to make the data easy to use and help researchers avoid potential pitfalls. In addition to these core data sets, we present seven supplementary data sets that record the size structure of key species and provide additional information about fish and kelp densities, patterns of invertebrate recruitment and historical water temperatures. These supplementary data provide valuable information about the processes that underlie the community dynamics described in the core data sets.

The data might be used to describe how these rocky reef communities have changed over time, to evaluate local management practices and human impacts, and to detect the effects of larger-scale anthropogenic effects such as global climate change and ocean acidification. They also give considerable insight into the ecology and population status of the particular species being monitored. Regardless of the use, we strongly recommend that researchers contact the coordinator of the KFM when planning analyses to gain a better understanding of the data set and the system.

METADATA

Metadata table of contents

D. Monitoring methods and data

1. Core community structure data sets

a. Benthic density data

i. Benthic density summary procedures

ii. 1 m^2 quadrat monitoring methods

iii. 5 m^2 quadrat monitoring methods

iv. Band transect monitoring methods

b. Benthic cover data

i. Benthic cover summary procedures

ii. Random point contact monitoring methods

c. Fish density data

i. Fish density summary procedures

ii. Visual fish transect monitoring methods

2. Supplementary data sets

a. Roving diver fish count data

i. Roving diver fish count methods

b. Fish size-frequency data

i. Fish size-frequency methods

c. Invertebrate size-frequency data

i. Invertebrate size-frequency methods

d. Giant kelp size-frequency data

i. Giant kelp size-frequency methods

e. Giant kelp supplementary density data

i. Giant kelp supplementary density methods

f. Artificial recruitment module data

i. Artificial recruitment module methods

g. Subtidal temperature data

i. Subtidal temperature methods

Class III. Data set status and accessibility

A. Latest data update

B. Latest metadata update

C. Data verification

D. Copyright or proprietary restrictions

Class IV. Data set structural descriptors

A. Data files

B. Metadata tables

Class V. Supplemental descriptors

A. Location of completed data forms

B. Data entry verification procedures

C. Publications using the data set

Class I. Data set descriptors

A. **Data set identity**
B. **Data set identification code**
C. **Data set description**
    1. **Data set originator**
    2. **Abstract**
D. **Key words**

I.A. Data set identity: Subtidal community data from the Channel Islands National Park's kelp forest monitoring program.

I.B. Data set identification code: KushnerEtAl2013-KFMP.

I.C. Data set description

I.C.1. Data set originator

Individual:	David J. Kushner
Role:	Program Manager, Corresponding Author
Organization:	Channel Islands National Park
Position:	Marine Biologist
Address:	1901 Spinnaker Dr.
	Ventura, CA 93001
Phone:	(805) 658-5773
Email:	David_Kushner@nps.gov

I.C.2. Abstract:

Studies of temperate rocky reef communities have added much to our understanding of ecology. However, data on these reef communities can be difficult to obtain; wind, waves and poor underwater visibility often prevent research diving, and even under the best conditions only a few hours a day can be spent underwater collecting data. Here, we present data on temperate subtidal reef communities at 33 sites, almost half of which the National Park Service has sampled annually since 1982. We present core data sets describing the population dynamics of 16 fish, 37 invertebrate and 15 algal taxa. We include supplementary data sets documenting the size structure of key species, the relative abundance of all fish, the recruitment of selected invertebrates and the subtidal water temperature at each site through time. Taken together, these data provide one of the most comprehensive descriptions of nearshore reef community dynamics ever assembled.

I.D. Key words: *Kelp Forest, Long-term Monitoring, California Channel Islands, subtidal reef ecology, marine protected areas*

Class II. Research origin descriptors

A. **Project description**
   1. **Objectives**
   2. **System description**
   3. **Guidelines for use of data**

II.A.1. Objectives: When the United States Congress passed legislation creating the Channel Islands National Park (CINP), the National Park Service (NPS) was charged with conducting "an inventory of all terrestrial and marine species, indicating their population dynamics, and probable trends as to future numbers and welfare" (16 USC § 410FF-2, 1980). Shortly after the park was created in 1980, the NPS initiated several ecological inventorying and monitoring programs, including the kelp forest monitoring program (KFM), focusing on the ecological communities found on shallow rocky reefs around the five northernmost Channel Islands. The KFM takes a population dynamics approach to describing the status of communities on these reefs, monitoring the abundance of selected species in the same locations every year. In addition to these records of population status the KFM collects supplementary data from each of these locations to help describe the biological and physical processes that govern community dynamics on these reefs.

II.A.2. System description: This study focuses on shallow (< 20 m deep) rocky reefs in the CINP. The waters of CINP give refuge to an ecologically diverse collection of species. The park is located at the boundary between two major biogeographical provinces: the Oregonian province to the north and the Californian to the south. The western park islands, San Miguel and Santa Rosa, are regularly subject to the cooler waters of the California Current. Waters around the eastern park islands of Anacapa and Santa Barbara come from the south along the mainland coast and are typically warmer. Around Santa Cruz Island, there is a broad transition zone influenced by waters from the north and south (Figure 1).

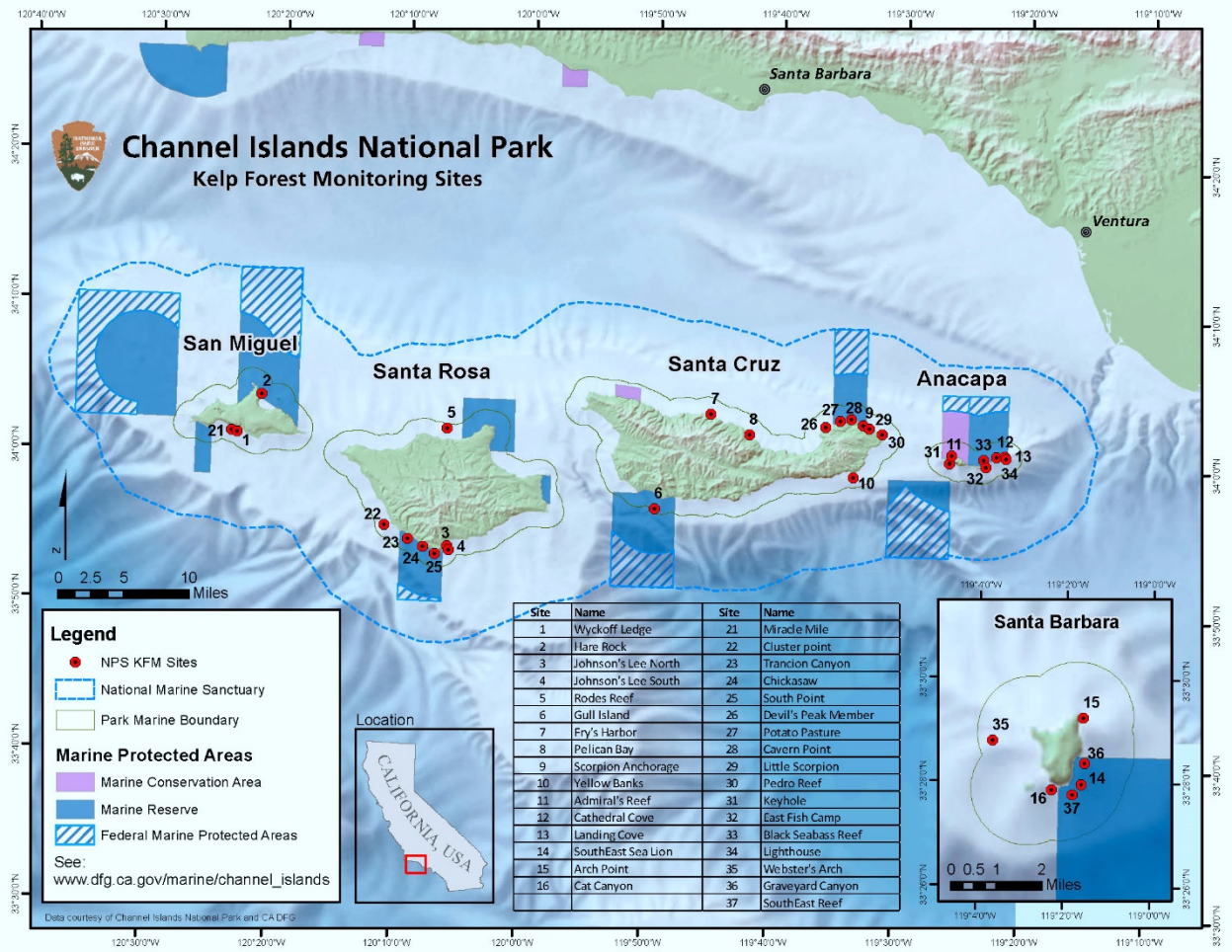
Prevailing winds and the bathymetry of adjacent basins also greatly influence marine communities in the park. Strong north winds and winter swells buffet the north sides of the islands, while the biota of the southern coasts reflects their more sheltered location. Upwelled nutrients from 2,000 meter-deep basins to the south and west of the park contribute to exceptionally productive food webs and create temperature regimes that differ significantly from the relatively shallow northern sides of the islands.

II.A.3. Guidelines for use of data: This data set is part of the ongoing Kelp Forest Monitoring Program. Researchers using these data are encouraged to contact the corresponding author (David Kushner) concerning questions, improvements, and, when appropriate, collaboration. The first author has been managing the National Park Service's kelp forest monitoring program since 1992 and can provide more details on the sampling methodology and the natural history of the system. Additionally, he keeps track of past and ongoing analyses to help reduce duplication of effort (see also section V.C. *Publications using the data set*).

B. **Monitoring design**
   1. **Site selection and layout**
   2. **Human impacts**
   3. **Monitoring method selection**
   4. **Choice of species to be monitored**
   5. **Monitoring design metadata**

II.B.1. Site selection and layout: The Channels Islands National Park's Kelp Forest Monitoring Program began in 1982 and has grown over time to include 33 sites. A single permanent transect at each site is sampled annually, between May and October. A map of the park and monitoring site locations is given in Figure 1.

Figure 1: Map of kelp forest monitoring site locations



Site	Name	Site	Name
1	Wyckoff Ledge	21	Miracle Mile
2	Hare Rock	22	Cluster point
3	Johnson's Lee North	23	Trancion Canyon
4	Johnson's Lee South	24	Chickasaw
5	Rodes Reef	25	South Point
6	Gull Island	26	Devil's Peak Member
7	Fry's Harbor	27	Potato Pasture
8	Pelican Bay	28	Cavern Point
9	Scorpion Anchorage	29	Little Scorpion
10	Yellow Banks	30	Pedro Reef
11	Admiral's Reef	31	Keyhole
12	Cathedral Cove	32	East Fish Camp
13	Landing Cove	33	Black Seabass Reef
14	SouthEast Sea Lion	34	Lighthouse
15	Arch Point	35	Webster's Arch
16	Cat Canyon	36	Graveyard Canyon
		37	SouthEast Reef

The initial 16 monitoring sites were selected from throughout the CINP; 13 sites were established in 1982, one in 1983 and two in 1986. As described below, site selection was not random, and, as a result, average densities for the sites should not be used to extrapolate to average conditions in the system. They are better suited for tracking changes over time, or matching biological communities to site characteristics.

The initial sites were distributed on both the north and south sides of each of the islands, and chosen to cover the east-west transition from the Californian to the Oregonian biogeographic provinces. Within these broad parameters, sites were selected based on the presence of continuous patches of rocky habitat (typically more than 100 m in long-shore extent) and because they were known to be good habitat for

giant kelp forests (i.e., giant kelp was known to be present at all sites in 1982 or shortly prior). Of the initial 16 sites, two were placed within a no-take marine reserve on Anacapa Island (#12 and #13), and a third placed in a nearby seasonal closure on Anacapa Island (#11).

In 2001, the Miracle Mile site (#21) at San Miguel Island was added to the KFM for the purpose of monitoring abalone populations. This site was selected by a commercial fisherman based on its exceptionally high density of the red abalone, *Haliotis rufescens*. In 2005, the Park received funding to expand its monitoring within and adjacent to several newly established marine protected areas (MPAs) at the Channel Islands. The KFM added 16 new sites located inside or adjacent to four of the newly established MPAs: Santa Barbara Island State Marine Reserve, Anacapa Island State Marine Reserve, Scorpion State Marine Reserve, and South Point State Marine Reserve. These four MPAs were chosen for additional monitoring because they were accessible, contained existing KFM monitoring sites and were in areas most likely to be impacted by fishing. New sites were established to complement existing sites so that at least three sites were inside and three adjacent to each of the four MPAs. Permanent sampling locations at these sites were placed in areas of continuous rocky reef. All 33 kelp forest monitoring sites and their locations (WGS 84) are described in Table 1. This information is also available in a format that can be read by geospatial software such as Google Earth. Links to an underwater video transect done at each site in 2010 are provided in Table 1, videos from other years are available from the NPS.

Monitoring at each site is focused on a permanent 100-meter transect, typically laid out parallel to the shoreline along a rough depth isobath. Permanent transects were used to reduce within-site variability in the time series and thus provide measurements of population dynamics with time as the primary independent variable. Each transect is marked by a 12 mm diameter lead-filled woven nylon line permanently affixed to the bedrock with stainless steel eyebolts positioned every 10 meters along the transect. Transects are relocated using GPS and diver search; while individual eyebolts have been lost and needed to be replace from time to time, the permanent line has allowed the transect location to be accurately relocated in all such cases. The transect depths at the original 16 sites vary from site to site, spanning a range from 5-18 meters. Transects at sites established in later years were placed between 7-17 meters in depth.

II.B.2. Human impacts: Many species monitored here are harvested commercially or recreationally. In particular, large and valuable fisheries exist for spiny lobster (*Panulirus interruptus*), California sheephead (*Semicossyphus pulcher*), red sea urchins (*Strongylocentrotus franciscanus*) and sea cucumbers (*Parastichopus parvimensis*). Although most of the sites are open to fishing and have no special restrictions on harvest beyond those common throughout southern California, several sites are currently within marine protected areas where harvest is restricted (Airamé et al. 2003, Davis 2005). All of the protected areas were established in 2003 except the area at the east end of Anacapa Island which predates the monitoring program and includes two of the original monitoring sites. A list of the special restrictions on harvest at each site (if any) and the dates those restrictions were implemented is given in Table 2.

II.B.3. Monitoring method selection: The great diversity of organisms and physical habitats associated with kelp forests require multiple sampling approaches to effectively document species population dynamics. Sampling methods were designed to maximize accuracy and precision while preserving an observer's ability to efficiently sample several target species at once. Additionally, sampling techniques were designed to minimize variation among observers and impacts on the populations of organisms being monitored. Some details of the sampling techniques have been revised several times since the project started. Most notably, several methods have changed their sampling effort one or more times as the result of ongoing power analyses and review of the data. These changes in sampling intensity have been noted in the metadata and in the data sets.

II.B.4. Choice of species to be monitored: The primary objective in selecting taxa for monitoring was to provide a representative cross section of the ecological guilds present in the CINP kelp forests. There are currently 15 algal, 37 invertebrate, and 16 fish taxa specifically targeted in the monitoring, and all fish are monitored with supplementary methods. The species monitored and the data sets in which they appear can be found in Table 3. Though many common organisms are identified to the species level, some are identified only to a coarse taxonomic level (e.g. green algae).

II.B.5. Monitoring design metadata: The tables describing general monitoring methods described above are also available below as comma-separated values files.

Table 1: Monitoring sites and their coordinates

Table 2: History of special fishing restrictions

Table 3: Species monitored by method

C. **Data processing and presentation**
   1. **Core community structure data sets**
   2. **Supplementary data sets**
   3. **Additional documentation and data**

II.C.1. Core community structure data sets: We present data on kelp forest community structure in a format that can be easily analyzed, and so have taken steps to construct data sets that are consistent through time. To achieve this we have combined data from different sampling methods where possible, removed species and sites that were sporadically measured and corrected for changes in sampling effort over time. This has resulted in three summarized "core community structure data sets": benthic density (calculated from counts of organisms in 1 m^2 quadrats, 5 m^2 quadrats and band transects), benthic cover (calculated from random point contact data), and fish density (calculated from visual fish transects). In each of these data sets, the abundance of each species or taxa sampled is expressed as a mean for each site on each sampling date with associated variance where appropriate. Full details on both the methods used to collect these data and the steps taken to process and summarize the raw measurements are given below. Note that the sampling design has been consistent since 1996, so users concerned about changes in sampling effort can limit their analysis to data collected after that year.

II.C.2. Supplementary data sets: In addition to the core data discussed above, the KFM collects several other forms of data that are not simple measures of abundance. We provide these data in seven separate data sets: *roving diver fish counts, fish size-frequency, invertebrate size-frequency, giant kelp size-frequency, giant kelp supplemental density, artificial recruitment module data* and *subtidal temperature data*. These data have not been summarized to a single number for each site and year, so will likely require some processing by users. Although the data are of high quality, many of these data sets are complex and we caution users of these data to carefully read the methods to understand how sampling was done.

II.C.3. Additional documentation and data: Considerable additional documentation is available for the KFM, including the sampling protocols used to guide field teams, annual reports describing changes in the system, and video transects from 1983-2011 for each site. These are available from the National Park Service website (www.nps.gov/chis). Additional data including the unprocessed field data can be obtained by contacting David Kushner, or the current superintendent of the Channel Islands National Park.

D. **Monitoring methods and data**

1. Core community structure data sets:

a. Benthic density data

i. Benthic density summary procedures

ii. 1 m^2 quadrat monitoring methods

iii. 5 m^2 quadrat monitoring methods

iv. Band transect monitoring methods

b. Benthic cover data

i. Benthic cover summary procedures

ii. Random point contact monitoring methods

c. Fish density data

i. Fish density summary procedures

ii. Visual fish transect monitoring methods

II.D.1. Core community structure data sets

II.D.1.a. Benthic density data

Table 4A: Definition of column headers in the benthic density data

Table 4B: Definition of variables in the benthic density data

Benthic density data: **ASCII file in comma separated values format**

II.D.1.a.i. Benthic density summary procedures: Benthic density data are collected by counting the number of individuals in a fixed area and dividing that count by the area sampled. This form of monitoring is used to estimate the abundance of larger benthic organisms such as sea stars and kelps as well as some small benthic fish. In the KFM, three separate methods are used at each site to estimate benthic densities (1 m^2 quadrats, 5 m^2 quadrats and band transects, sampling 1-2 m^2, 5 m^2 and 40-60 m^2 respectively). Because the sampling intensities for these methods have varied through time, data are summarized as the mean density (per m^2) of each taxa at the site, averaged over all replicates within the site for a given year. Also included are the standard errors in density calculated across the sampling replicates, the number of replicates and the area sampled in each replicate.

Different methods are used to count different species, and in some cases the method used to count a particular species changed over time. By combining the densities estimated from each method into a single data set, we assemble continuous time series of the abundance of each species for which density estimates were made. In the few cases where the same species was sampled by multiple methods at the same site in the same year, we used the method that sampled the largest area.

Some species have been added to the monitoring protocols during the 30+ years of monitoring. Thus the absence of these species from the data early in monitoring cannot be taken as evidence of absence. For this reason, instead of a 0 or blank, the code "NA" is entered into the data set as the density for species in years they were not counted. Different size classes of several kelps were counted separately for the later portion of the monitoring. However we have summed across age classes to obtain consistent categories throughout this data set.

II.D.1.a.ii. 1 m^2 quadrat monitoring methods: 1 m^2 quadrats have been sampled since 1982 by counting the number of individuals of each target species occurring within a set of small quadrats at each site. Currently 12 quadrats are sampled, with pairs of quadrats placed every 8.33 m (rounded to the nearest

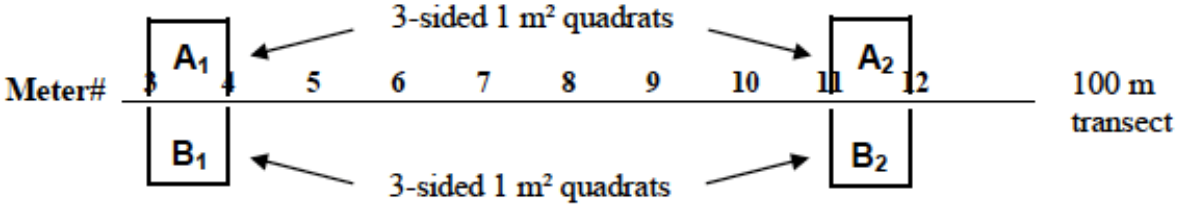
whole meter) along the main transect (Figure 2). Over the course of the program, sample size has varied from 12 to 40 quadrats, and the area of each replicate quadrat has varied from 1 m^2 to 2 m^2 (Table 5).

Table 5: 1 m^2 quadrat sampling intensity by year.

Year	# Quadrats	Size
1982	30	1 m^2
1983-1984	40	1 m^2
1985-1995	20	2 m^2
1996-2011	12	2 m^2

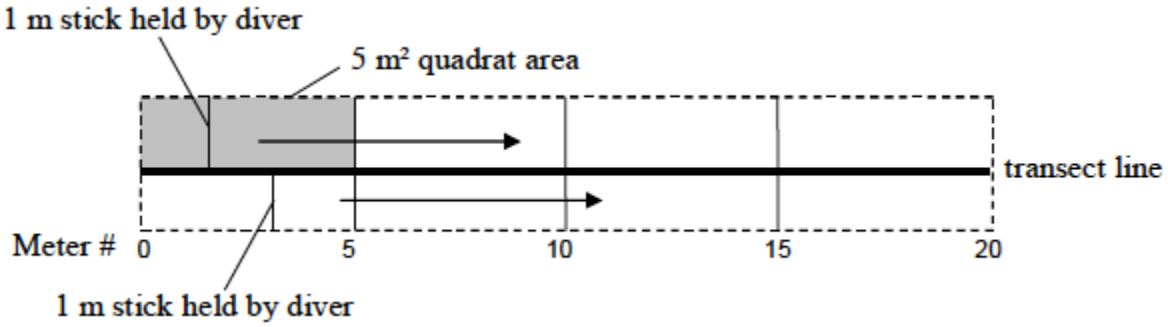
Placement of quadrats is based on a stratified random design, with new points on the transect chosen every year, although the exact method of stratification has changed over time. For this monitoring method, observers place two 1 m^2 frames adjacent to the main transect, so that they form a 2 m^2 area (e.g., A_1 and B_1 in Figure 2) using the transect line as the fourth side of each quadrat. When sampling, divers place the frames, then withdraw to allow island kelp fish, gobies and other easily disturbed species to emerge from crevices before approaching the quadrats slowly. Fishes within the quadrat are counted first; additional fish that enter the quadrat after this first count are ignored. Divers search under ledges and in cracks for all target organisms but do not conduct invasive sampling (e.g., turning over rocks, sifting through sand, etc.).

Figure 2. Placement of 1 m^2 quadrats under current design



II.D.1.a.iii. 5 m^2 quadrat monitoring methods: 5 m^2 quadrats have been sampled since 1996 by counting the number of individuals of each target species occurring within a set of 5 m^2 sampling areas at each site. This sampling method was added to improve sampling of key species that are too rare to be regularly encountered in 1 m^2 quadrats, but too numerous to quickly count in the larger band transects. The 100 m transect is divided into 20 segments, each associated with two 5 m by 1 m sampling areas, placed one on either side of the transect (e.g., Figure 3). Divers search the habitat thoroughly for target species, including cracks and crevices but do not conduct any invasive sampling (e.g., turning over rocks, sifting through sand, etc.).

Figure 3. Placement of 5 m quadrats along the transect line.



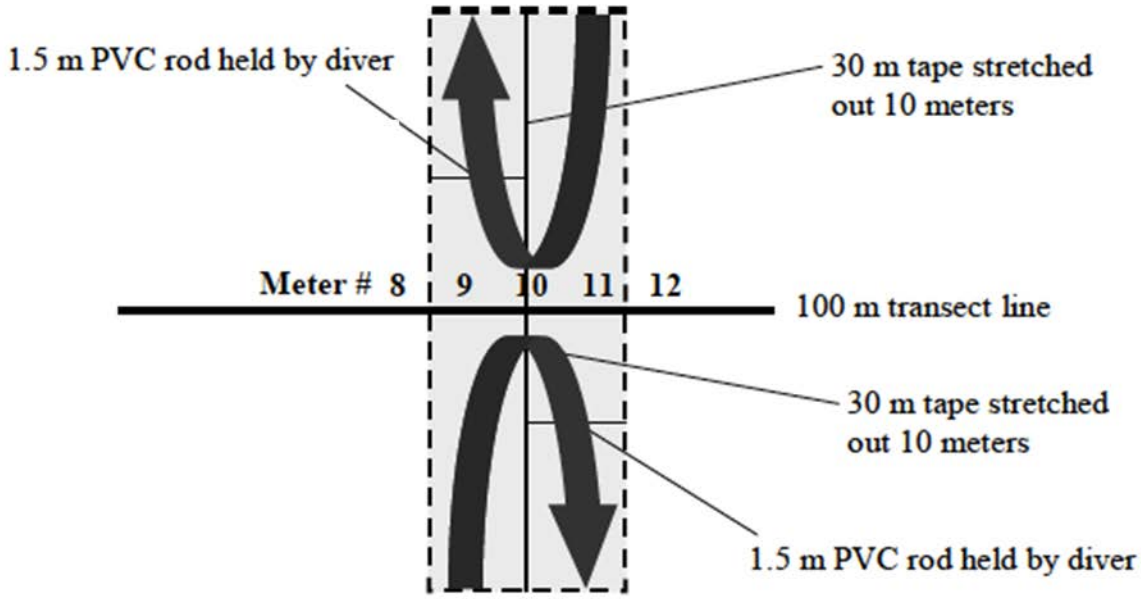
II.D.1.a.iv. Band transect monitoring methods: Band transects have been sampled since 1983 by counting the number of individuals of each target species occurring within a set of band transects at each site. Currently 12 bands are sampled, each covering 60 m^2, but sampling intensity has varied over time (Table 6).

Table 6: Band transects sampling intensity by year.

Year	# Band Transects	Band Transect Size
1983-1984	10	40 m^2
1985-2011	12	60 m^2

Placement of bands is based on a stratified random design, with different points on the main transect chosen every year for sampling. At each sampling point, divers run measuring tapes 10m in each direction, perpendicular to the 100 m transect line, sampling organisms first within 1.5m of one side of the tape and then within that distance of the other tape (Figure 4, before 1985 only 1m on each side was sampled). Divers search the habitat thoroughly, including cracks and crevices, however, they do not conduct any invasive sampling (e.g., turning over rocks, sifting through sand, etc.).

Figure 4. Example of band transect placement under current design.



Shaded portion indicates the area sampled. Arrows indicate direction diver swims.

II.D.1.b. Benthic cover data

Table 7A: Definition of column headers in benthic cover data

Table 7B: Definition of variables in the benthic cover data

Benthic cover data: **ASCII file in comma separated values format**

II.D.1.b.i. Benthic cover summary procedures: Percent cover data for sessile benthic biota (e.g., bryozoans, understory macroalgae) are collected using a random point contact (RPC) method, in which a set of randomly chosen points are superimposed over the bottom and the species contacted by each point are recorded. The number of times each species was counted is divided by the total number of points to give the percent of the bottom occupied by that species. The physical substrate is also categorized during the RPC process and the cover of each substrate is recorded in this data set. The benthic cover data set contains the mean percent cover of each target organism, averaged across all RPC units sampled at that site, the standard error in percent cover calculated across the RPC units and information about the number of RPC units sampled and the number of points per unit. The sampling intensity for this method has changed through time, but by summarizing the data as an average percent cover for the site, we have corrected for this variation in effort.

Some species have been added to this method's target species list during the 30 years of monitoring. Thus the absence of these species from the data early in the monitoring cannot be taken as evidence of absence. The code "NA" is entered into the data set as the percent cover of species in years they were not counted. Several species which were not distinguished individually early in the monitoring were later sampled

separately. However, we have aggregated such species groups in the benthic cover data so that taxonomic resolution is consistent throughout the data set.

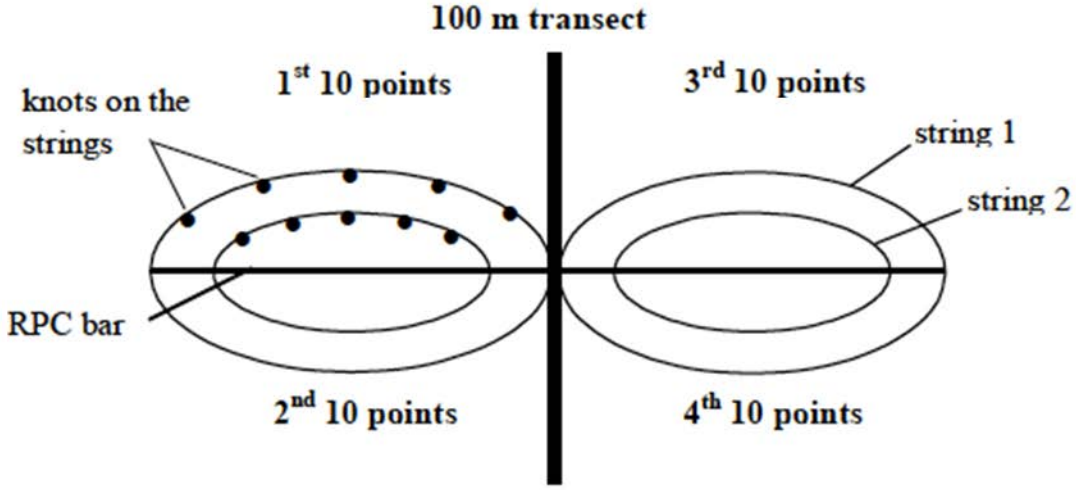
II.D.1.b.ii. Random point contact monitoring methods: Benthic cover data are collected in RPC units of 40 points distributed around a given location on the main transect. Currently 15 units placed approximately every 6.6 m (rounded to the nearest meter) along the transect are sampled at each site, but the intensity of sampling has varied over time (Table 8). Placement of RPC units follows a stratified random design, with new locations on the transect chosen every year, although the exact method of stratification has changed over time. Again, this introduces an element of year to year variation not due to actual changes over time.

Table 8: RPC sampling intensity by year.

Year	Replicates	Points per replicate
1982	25	20
1983	40	10
1984	10	50
1985-1995	25	40
1996-2011	15	40

To perform this sampling, a diver uses an “RPC bar” a 1.5 m rod with 2 strings attached, each string with 5 knots. One string is 1.8 m long and the other measures 1.2 m. The long string attaches to the ends of the bar and the short string attaches 25 cm from each end. At each sampling location, the diver places the RPC bar perpendicular to the transect line (Figure 5), holds the RPC bar in place and stretches the string taut at each knot perpendicular to the bar. This forms a triangle with the bar as one edge and the knot as the opposite vertex. The diver then imagines a line running vertically through the knot from the substrate up to one meter above the substratum and identifies all organisms that intersect this imaginary line. The diver identifies all organisms under or above the knots on both strings on one side of the bar, giving a total of ten points. Next, the diver moves the strings to the opposite side of the RPC bar for another 10 points, then places the bar on the other side of the transect line and repeats the process (Figure 5).

Figure 5: Orientation of RPC bar and strings during sampling.



The diver counts each species only once per point even if multiple individuals of the same species intersect that point. Because of this method of counting multiple layers, total cover of all species can exceed 100% (and usually does), but the cover of any individual species cannot be greater than 100%. The one exception is the category "*Macrocystis*, *Pterygophora*, and *Eisenia* combined" (species code 2008). Each species in this category is counted separately and then these counts are added, so percent covers of as high as 300% are possible for this combined-species category (if all points intersect all three species). Attached or sessile animals providing cover are recorded as "Miscellaneous Invertebrates." Motile invertebrates (except for *Ophiothrix spiculata* and *Pachythyone rubra*, which are counted with "Miscellaneous Invertebrates"), are not counted but are moved to determine what is underneath. The diver also identifies the substrate type encountered at each point.

II.D.1.c. Fish density data

Table 9A: Definition of column headers in fish density data

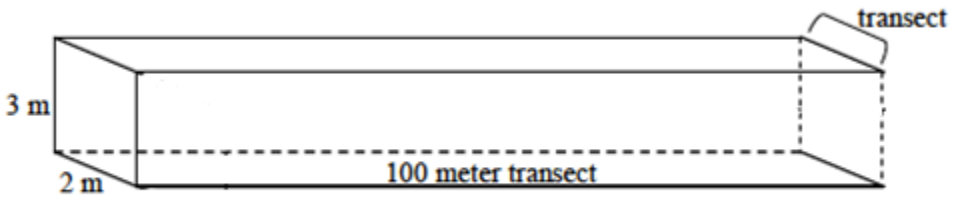
Table 9B: Definition of variables in the fish density data

Fish density data: **ASCII file in comma separated values format**

II.D.1.c.i. Fish density summary procedures: Fish density data are collected by counting the number of individuals in a single fixed plot at each site. This form of monitoring is used to estimate the abundance of 13 species of fish and is based on visual fish transect data. For each fish species counted in a given year, the fish density data set contains the total count of that species; because these data are based on a single transect, no variability is reported. Occasionally the same plot was sampled on multiple dates within a single year; these counts are reported as separate entries in the fish density data set. Sampling for *Semicossyphus pulcher* juveniles began in 1996. Thus their absence from the data early in the monitoring cannot be taken as evidence of absence. The code "NA" is entered into the data set as the density in years they were not counted.

II.D.1.c.ii. Visual fish transect monitoring methods: To count fish, a diver swims along the main 100 m transect line, counting and recording all the individuals of a core set of species within a 3 m tall x 2 m wide x 100 m long area in front of the diver. The diver records any fish seen within this area (Figure 6). Divers typically swim the 100 m transect in 5 minutes. Counts are differentiated by species and by age class—adults or juveniles, with juveniles defined as less than 10 cm in length, except for *Halichoeres semicinctus* and *Semicossyphus pulcher* for which juveniles must be less than 10 cm in length and have a white stripe, and *Hypsypops rubicundus* for which juveniles are not identified by size, but by the presence of blue spots. A few species are also differentiated by sex. Fish transects are conducted once or twice per summer at each site, as time allows, with a minimum of two weeks between sampling events.

Figure 6: Area covered by visual fish transects.



II.D.2. Supplementary data sets

- **a. Roving diver fish count data**
  - **i. Roving diver fish count methods**
- **b. Fish size-frequency data**
  - **i. Fish size-frequency methods**
- **c. Invertebrate size-frequency data**
  - **i. Invertebrate size-frequency methods**
- **d. Giant kelp size-frequency data**
  - **i. Giant kelp size-frequency methods**
- **e. Giant kelp supplementary density data**
  - **i. Giant kelp supplementary density methods**
- **f. Artificial recruitment module data**
  - **i. Artificial recruitment module methods**
- **g. Subtidal temperature data**
  - **i. Subtidal temperature methods**

II.D.2.a. Roving diver fish count data:

Table 10A: Definition of column headers in RDFC data

Table 10B: Definition of variables in the RDFC data

RDFC data: **ASCII file in comma separated values format**

II.D.2.a.i. Roving diver fish count methods:

The roving diver fish count (RDFC) is a method for counting fish that covers a much larger area than the visual transects and attempts to sample all fish species, instead of just the core set of species counted in visual fish transects. It also has the advantage that multiple counts are done by different observers at a site on most sampling dates. As such, the RDFC data provide more complete information on the total fish community than the fish transect data. Because the RDFC seeks to sample fish within a fixed area, it is tempting to convert these data into densities, but because sampling occurs over a fixed time (30 minutes) there are potential pitfalls with treating these data as density estimates.

The RDFC method produces two indices of fish abundance for each species: an abundance category (*single, few, common, many*) and a count of the number of individuals observed (see Table 11 for the relationship between categories and counts). The count data is a better index of abundance, but we retain data on abundance category since fish counts were only added to the RDFC protocol in 2003. Prior to that (from 1996 to 2002) only abundance categories were recorded. Even after 2003 there have been a few cases when observers recorded an abundance category instead of a count for a species—usually because there were too many fish for the observer to reliably count all of the species that were present. Fish are identified to species where possible, and in some cases are separated into adults and juveniles or males and females. Although the taxonomic resolution has remained constant over time, in many cases adults and juveniles were counted together prior to 1999 and separated into adult and juvenile categories from 1999 onward.

During the roving diver fish count, divers gradually swim around the transect line, covering the entire transect area in 30 minutes. The sampling area encompasses ten meters on both sides of the 100 m main transect yielding a total sampling area of 2000 m^2. Throughout the fish count, each diver attempts to search all habitats (i.e. bottom, midwater, under ledges, canopy, etc.) counting all fish observed. Divers attempted to count each fish only once, although some double counting likely occurs, particularly for more mobile species. Each observer is assigned an individual skill level rating based on their experience, but only data from the most experienced observers (i.e., "expert") is included here. An "expert" is defined here as an observer who can confidently identify and count all species of fish that commonly occur at the Channel Islands. For each species, the total count or abundance category are recorded by each observer separately—there may be from 1-8 separate expert counts done at a site on one date and from 1-12 expert counts done at each site in a year.

Table 11: Definition of fish abundance categories

Abundance Category	Equivalent Count
Single	1
Few	2-10
Common	11-100
Many	100+

II.D.2.b. Fish size-frequency data:

Table 12A: Definition of column headers in the fish size-frequency data

Table 12B: Definition of variables in the fish size-frequency data

Fish size-frequency data: **ASCII file in comma separated values format**

II.D.2.b.i. Fish size-frequency methods:

The KFM collects data on fish size-frequency to help capture changes in the size structure of the fish community that would not be apparent from simple counts. This sampling method is performed during or after the roving diver fish count with a minimum sampling time of 30 minutes. As with the RDFC, the observer will sample as much of the 2000 m^2 area (ten meters on either side of the 100 m permanent transect line) as possible. Within this area and time, the sizes of as many individuals as possible are estimated, prioritizing certain species as described below. Observers for this protocol are trained, tested and able to accurately estimate fish sizes underwater to within 20% of the actual total length (TL). Observers estimate the total length of fish to the nearest centimeter searching all habitats (i.e. bottom, midwater, under ledges, water column, canopy, etc.). When schools of fish of the same species are encountered at a site, the mean size of individuals in that group is estimated, as well as the number of fish in the group, and the size of the largest and smallest individuals.

All fish observed are measured except cryptic species (e.g., sculpins and sand dabs) and schooling baitfish (e.g., such as sardines, anchovies and smelt). Some measured species are given priority over others, with species ranked by priority categories from 1 (highest) to 4 (lowest) as indicated in the data set. Fish species are prioritized to ensure that particular target species are sampled even if time does not allow for sampling all fish at the site. As of autumn 2012 it has been possible to sample the fish community completely for the top three priority levels, so the size structure of fish given priority 1, 2 or 3 can be considered complete (i.e., sizes were recorded for all fish of those species encountered).

II.D.2.c. Invertebrate size-frequency data:

Table 13A: Definition of column headers in the invertebrate size-frequency data

Table 13B: Definition of variables in the invertebrate size-frequency data

Invertebrate size-frequency data: **ASCII file in comma separated values format**

II.D.2.c.i. Invertebrate size-frequency methods: The KFM collects data on the size structure of populations of key invertebrate species to help capture changes in those populations that might not be apparent from simple measures of abundance. The number of individuals measured has varied among species, sites and years. A diver records the sizes of individuals of each target species, completely searching swaths running parallel to the main transect, with the spacing and length of these swaths chosen to gather a representative sample from the entire 100 m transect line. When there are relatively low densities of the target species, the observer conducts one long swath; when densities are very high, observers may sample haphazardly placed areas as small as 0.5 m^2. In every case all target organisms

within the chosen area are measured to ensure that an unbiased sample of sizes is obtained. For *Stylaster californica*, *Lophogorgia chilensis*, *Muricea fruticosa* and *Muricea californica* we report the estimated area of the fan in cm^2 based on measurements of height and width (made to the nearest centimeter) and assuming an oval fan shape. All other invertebrate measurements are made in centimeters—measuring the diameter of the test for urchins, the longest dimension of the shell for gastropods and the distance from the mouth to the end of the longest arm for sea stars. Measurements are made to the nearest 0.1 cm *in situ* with minimal disturbance to the organisms, except for sea urchins, which are lifted (if possible) to search for juveniles under the spine canopy of large adults.

II.D.2.d. Giant kelp size-frequency data:

Table 14A: Definition of column headers in the giant kelp size-frequency data

Table 14B: Definition of variables in the giant kelp size-frequency data

Giant kelp size-frequency data: **ASCII file in comma separated values format**

II.D.2.d.i. Giant kelp size-frequency methods: The KFM collects data on the size structure of giant kelp, *Macrocystis pyrifera*, greater than 1 m tall at each site where it is present, to help capture changes in kelp populations that might not be apparent from simple measures of its abundance. Giant kelp sizes are measured within a swath running along the entire length of the permanent 100 m transect. If giant kelp are rare at the site all plants within 10 m of the transect will be sampled, otherwise a narrower swath is sampled, with a width chosen to include at least 100 individuals. Within the sampling area all plants encountered are measured. For each plant, the number of stipes one meter above the bottom is recorded, as well as the greatest diameter of the holdfast, measured to the nearest centimeter. In some cases holdfasts were not measured and the code "NA" indicates this.

II.D.2.e. Giant kelp supplementary density data:

Table 15A: Definition of column headers in the giant kelp supplementary density data

Table 15B: Definition of variables in giant kelp supplementary density data

Giant kelp supplementary density data: **ASCII file in comma separated values format**

II.D.2.e.i. Giant kelp supplementary density methods: In addition to the data on *Macrocystis pyrifera* included in the benthic density data set (section II.D.1.a.), the KFM has collected additional data on the abundance of this species since 1996. The benthic density data set, collected since 1982, reports the abundance of giant kelp adults and juveniles, with adults defined as plants with one or more stipes taller than 1 meter. It provides a very useful time series, but because it is derived from the 1 m^2 quadrats (section II.D.1.a.ii.) the total area sampled is small (less than 40 m^2). Since 1996 the KFM has recorded two additional metrics of kelp abundance, measured in the 5 m^2 quadrats (see section II.D.1.a.iii. for sampling methods): the density of adult giant kelp (plants greater than 1 m tall and possessing haptera above the primary dichotomy) and the density of sub-adult giant kelp (plants greater than 1 m tall without haptera above the primary dichotomy). Note that the adult and sub-adult plants counted here would both be considered "adult" in the core benthic density data set. Since 2007 the KFM has also estimated the density of stipes on plants greater than 1 m tall within the 1 m^2 quadrats (counting stipes without respect to what plant they are found on. Summer stipe density is a strong predictor of total giant kelp biomass and

production (Reed et al. 2009). In the giant kelp supplementary density we report the mean of each measure of density at the site, averaged over all replicates within the site for a given year. Also included are the standard error in density calculated across the sampling replicates, the number of replicates and the area sampled in each replicate.

II.D.2.f. Artificial recruitment module data:

Table 16A: Definition of column headers artificial recruitment module data

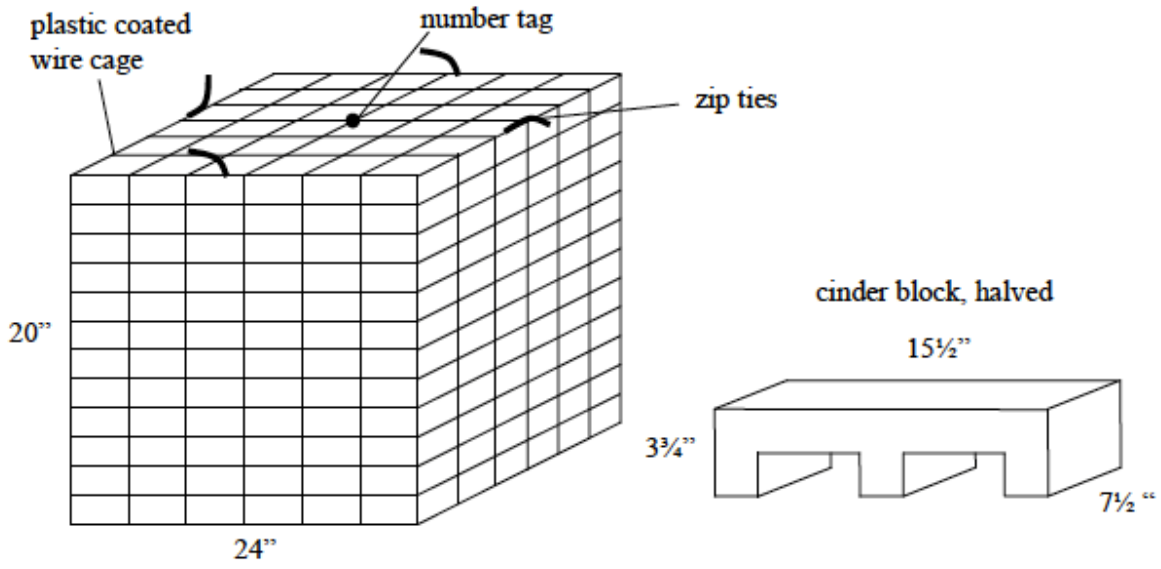
Table 16B: Definition of variables in the artificial recruitment module data

Artificial recruitment module data: **ASCII file in comma separated values format**

II.D.2.f.i. Artificial recruitment module methods: Artificial recruitment modules (ARMs) were placed at 11 sites starting in 1992 and selected organisms accumulating within them have been sampled each summer. The number and location of ARMs varies from site to site (typically 5-15 per site). Each artificial recruitment module consists of a wire cage made of 2" x 4" mesh wire (Figure 7) filled with 20 bricks that serve as recruitment substrates. Bricks are placed in the ARM in five layers with four bricks per layer arranged around the internal perimeter of the ARM, leaving an opening in the center of each row. Bricks are made by cutting a concrete cinder block in half longitudinally to produce 2 bricks, each with a cross section shaped like a lower case "m" (Figure 7). More details on the construction of the ARMs are given in Davis (1995).

To sample each ARM, the bricks are removed from the cage and all target organisms are collected, counted and sized to the nearest mm—measuring the diameter of the test for urchins, the widest dimension of the shell for gastropods and the distance from the mouth to the end of the longest arm for sea stars. *Parastichopus parvimensis* are not measured but are instead classified as larger or smaller than 10cm. Mean size and the associated standard error are reported for all species except *P. parvimensis*, for which only the fraction less than and fraction greater than 10cm are reported. If there are 200 or more *Strongylocentrotus purpuratus* or *S. franciscanus* per ARM, not all of the ARMs are sampled for these two species. In some cases, organisms are brought to the surface to be measured. After sampling all of the bricks and organisms are replaced in the ARM.

Figure 7: Example of an Artificial Recruitment Module (ARM).



II.D.2.g. Subtidal temperature data:

Table 17A: Definition of column headers subtidal temperature data

Table 17B: Definition of variables in the subtidal temperature data

Subtidal temperature data: **ASCII file in comma separated values format**

II.D.2.g.i. Subtidal temperature methods: Since 1993 the KFM has used automated loggers to measure water temperature at the monitoring sites over time. At each site, temperature loggers programmed to record temperature are bolted to the substrate. Loggers record temperature between 5 and 24 times a day, depending on logger memory capacity. Records are then averaged over each 24 hour period. The depth of the loggers at each site is listed in Table 1. In most years, two temperature loggers are deployed at each site. Loggers are serviced and data collected once per year during the summer sampling.

Upon retrieval, the data from the two loggers are compared and if the recorded temperatures differ by more than 0.2 C, the loggers are independently checked to determine which is most accurate and data from that logger is used. On the rare occasion that both loggers fail or are missing, no temperature data is entered into the data set for that site and year, so there are a few gaps in the time series. Several brands of temperature logger have been used at different periods, with some overlap, including HoboTemp, StowAway, Tidbit and UTBI Tidbit V2 loggers (all made by the Onset Computer Corporation).

Class III. Data set status and accessibility

A. **Latest data update**
B. **Latest metadata update**
C. **Data verification**
D. **Copyright or proprietary restrictions**

III.A. Latest data update: The data set may be periodically updated. All updates to the data have been logged in Table 18A. Please check for the latest update before using the data set.

III.B. Latest metadata update: The metadata may also be updated periodically. All updates have been logged in Table 18B. Please check for the latest update before using the data set.

III.C. Data verification: Field sheets were proofed for concerns after every day in the field as well as during data entry.

III.D. Copyright or proprietary restrictions: None

Class IV. Data set structural descriptors

A. **Data files**
B. **Metadata tables**

IV.A. Data Files:

Benthic density data

Benthic cover data

Fish density data

RDFC data

Fish size-frequency data

Invertebrate size-frequency data

Giant kelp size-frequency data

Giant kelp supplementary density data

Artificial recruitment module data

Subtidal temperature data

IV.B. Metadata Files:

Table 1: Monitoring Sites and Their Coordinates

Table 2: History of special fishing restrictions

Table 3: Species Monitored by Method and Year Monitored

Table 4A: Definition of Column Headers in the Benthic Density Data

Table 4B: Definition of Variables in the Benthic Density Data

Table 7A: Definition of column headers in benthic cover data

Table 7B: Definition of variables in the benthic cover data

Table 9A: Definition of column headers in fish density data

Table 9B: Definition of variables in the fish density data

Table 10A: Definition of column headers in RDFC data

Table 10B: Definition of variables in the RDFC data

Table 12A: Definition of column headers in the fish size-frequency data

Table 12B: Definition of variables in the fish size-frequency data

Table 13A: Definition of column headers in the invertebrate size-frequency data

Table 13B: Definition of variables in the invertebrate size-frequency data

Table 14A: Definition of column headers in the giant kelp size-frequency data

Table 14B: Definition of variables in the giant kelp size-frequency data

Table 15A: Definition of column headers in the giant kelp supplementary density data

Table 15B: Definition of variables in the giant kelp supplementary density data

Table 16A: Definition of column headers in the artificial recruitment module data

Table 16B: Definition of variables in the artificial recruitment module data

Table 17A: Definition of column headers in the subtidal temperature data

Table 17B: Definition of variables in the subtidal temperature data

Table 18A: History of data updates

Table 18B: History of metadata updates

Class V. Supplemental descriptors

A. Location of completed data forms
B. Data entry verification procedures
C. Publications using the data set

V.A. Location of completed data forms: Original field data forms are archived at the Channel Islands National Park office. All forms are scanned and stored electronically in PDF form on the park's network server.

V.B. Data entry verification procedures: See III.C.

V.C. Publications using the data set prior to 2013:

Babcock RC, Shears NT, Alcala AC, Barrett NS, Edgar GJ, Lafferty KD, McClanahan TR, Russ GR. 2010. Decadal trends in marine reserves reveal differential rates of change in direct and indirect effects. Proceedings of the National Academy of Sciences 107:18256-61.

Behrens M, Lafferty KD. 2004. Effects of marine reserves and urchin disease on southern Californian rocky reef communities. Marine Ecology Progress Series 279:129-139.

Byrnes J, Stachowicz JJ, Hultgren KM, Randall Hughes A, Olyarnik SV, Thornber CS. 2006. Predator diversity strengthens trophic cascades in kelp forests by modifying herbivore behaviour. Ecology Letters 9:61-71.

Davis G. 1993. Design elements of monitoring programs: the necessary ingredients for success. Environmental Monitoring and Assessment 26:99-105.

Davis GE. 2005. Science and Society: Marine Reserve Design for the California Channel Islands. Conservation Biology 19:1745-1751.

Davis G, Haaker P, Richards DV. 1996a. Status and trends of white abalone at the California Channel Islands. Transactions of the American Fisheries Society 125:42-48.

Davis G, Richards D, Kushner D. 1996b. Kelp forest monitoring design review. Technical Report CHIS-96-01, Channel Islands National Park, California.

Fancy SG, Gross JE, Carter SL. 2009. Monitoring the condition of natural resources in US national parks. Environmental Monitoring and Assessment 151:161-174.

Gelpi CG, Norris KE. 2008. Seasonal temperature dynamics of the upper ocean in the Southern California Bight. Journal of Geophysical Research 113:1-18.

Graham MH. 2004. Effects of local deforestation on the diversity and structure of southern California giant kelp forest food webs. Ecosystems 7:341-357.

Grantham B, Eckert G. 2003. Dispersal potential of marine invertebrates in diverse habitats. Ecological Applications 13:S108-S116.

Halpern BS, Cottenie K. 2007. Little evidence for climate effects on local-scale structure and dynamics of California kelp forest communities. Global Change Biology 13:236-251.

Halpern BS, Cottenie K, Broitman BR. 2006. Strong top-down control in southern California kelp forest ecosystems. Science 312:1230-2.

Karpov K, Haaker P, Albin D, Taniguchi IK, Kushner D. 1998. The red abalone, *Haliotis rufescens*, in California: Importance of depth refuge to abalone management. Journal of Shellfish Research 17:863-870.

Lafferty KD, Kuris AM. 1993. Mass mortality of abalone *Haliotis cracherodii* on the California Channel Islands: tests of epidemiological hypotheses. Marine Ecology Progress Series 96:239-248.

Lafferty KD, Kushner DJ. 2000. Population regulation of the purple sea urchin, *Stronylocentrotus purpuratus*, at the California Channel Islands. Proceedings of the Fifth California Islands Symposium:380-381.

Lafferty K. 2004. Fishing for lobsters indirectly increases epidemics in sea urchins. Ecological Applications 14:1566-1573.

Micheli F, Halpern BS. 2005. Low functional redundancy in coastal marine assemblages. Ecology Letters 8:391-400.

Reed DC, Raimondi PT, Carr MH, Goldwasser L. 2000. The role of dispersal and disturbance in determining spatial heterogeneity in sedentary organisms. Ecology 81:2011-2026.

Rogers-Bennett L, Allen B, Davis GE. 2004. Measuring abalone (*Haliotis* spp.) recruitment in California to examine recruitment overfishing and recovery criteria. Journal of Shellfish Research 23:1201-1207.

Rogers-Bennett L, Haaker P, Karpov KA, Kushner DJ. 2002. Using spatially explicit data to evaluate marine protected areas for abalone in southern California. Conservation 16:1308-1317.

Schroeter SC, Reed DC, Kushner DJ, Estes JA, Ono DS. 2001. The use of marine reserves in evaluating the dive fishery for the warty sea cucumber (*Parastichopus parvimensis*) in California, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 58:1773-1781.

Schroeter S, Page H, Dugan J, Culver CS, Steele B, Guiterrez R, Richards JB, Kushner D. 2009. Scales of variability in larval settlement within the Channel Islands National Marine Sanctuary and along the mainland coast. Proceedings of the 7th California Islands Symposium:151-160.

Zacharias MA, Kushner DJ. 2006. Sea temperature and wave height as predictors of population size structure and density of *Megastraea (Lithopoma) undosa*: Implications for fishery management. Bulletin of Marine Science 79:71-82.

Acknowledgements

The design, implementation and data collection of this ecological monitoring program have been almost entirely supported by the United States National Park Service. The California Department of Fish and Wildlife and the U.S. Department of Commerce's National Oceanographic and Atmospheric Administration have both cooperated greatly with this program. As of 2011, 377 scientists and resource managers have collected data underwater as part of this program. There have also been a large number of other scientists that helped to develop and improve the experimental design and resulting monitoring protocols. In addition to people listed as divers, other support staff including boat captains, dive officers and other topside personnel have all contributed to the success and longevity of the Kelp Forest Monitoring Program since 1982. Of particular importance was the founder of the monitoring program Gary E. Davis who had the vision to develop and support a program that continues to be a valuable component to resources management and support our long-term understanding of kelp forest ecology.

During the preparation of this data paper A.R. and J.M. were supported with funds from the U.S. Department of the Interior, Bureau of Ocean Energy Management, Environmental Studies Program, Washington, DC under Cooperative Agreement Number M11AC00012. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Government. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Government. This product (article, paper, etc.) has been peer reviewed and approved for publication consistent with USGS Fundamental Science Practices (http://pubs.usgs.gov/circ/1367/).

Literature cited

Airamé S, Dugan J, Lafferty K, Leslie H, McArdele D, Warner R. 2003. Applying ecological criteria to marine reserve design: a case study from the California Channel Islands. Ecological Applications 13:170–184.

Byrnes J, Stachowicz JJ, Hultgren KM, Randall Hughes A, Olyarnik SV, Thornber CS. 2006. Predator diversity strengthens trophic cascades in kelp forests by modifying herbivore behaviour. Ecology letters 9:61–71.

Cavanaugh KC, Siegel DA, Reed DC, Dennison PE. 2011. Environmental controls of giant kelp biomass in the Santa Barbara Channel, California. Marine Ecology Progress Series 429:1–17.

Davis GE. 1995. Recruitment of juvenile abalone (*Haliotis* spp.) measured in artificial habitats. Marine and Freshwater Research 46:549–454.

Davis GE. 2005. Science and society: marine reserve design for the California Channel Islands. Conservation Biology 19:1745–1751.

Davis G, Richards D, Kushner D. 1996. Kelp forest monitoring design review. Park Service, Channel Islands National Park. Channel Islands National Park, California. 24 pp.

Engle JM. 1994. Perspectives on the structure and dynamics of nearshore marine assemblages of the California Channel Islands. Proceedings of the Fourth California Islands Symposium:13-26.

Lafferty K. 2004. Fishing for lobsters indirectly increases epidemics in sea urchins. Ecological Applications 14:1566–1573.

Reed D, Raimondi P, Carr M, Goldwasser L. 2000. The role of dispersal and disturbance in determining spatial heterogeneity in sedentary organisms. Ecology 81:2011–2026.

Reed DC, Rassweiler A, Arkema KK. 2009. Density derived estimates of standing crop and net primary production in the giant kelp *Macrocystis pyrifera*. Marine Biology:2077–2083.

Schroeter SC, Reed DC, Kushner DJ, Estes JA, Ono DS. 2001. The use of marine reserves in evaluating the dive fishery for the warty sea cucumber (*Parastichopus parvimensis*) in California, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 58:1773–1781.

Appendix B

All files and folders referenced in Appendix B are found in the file:

OCS Study BOEM 2019-063 Appendix B files.zip

A multi-decade time series of kelp forest community structure at San Nicolas Island, California

Michael C. Kenner, James A. Estes, M. Tim Tinker, James L. Bodkin, Robert K. Cowen, Christopher Harrold, Brian B. Hatfield, Mark Novak, Andrew Rassweiler, Daniel C. Reed

2013. *Ecology* 94: 2654

INTRODUCTION

Reefs at San Nicolas Island are an ideal location for studying the dynamics of shallow subtidal communities. San Nicolas is the most remote of California's Channel Islands, and its relative isolation somewhat protects it from the terrestrial runoff and recreational use faced by other reefs in Southern California. Although the island is small, its gently sloping subtidal shelf provides substantial shallow area for kelp forests and other rocky reef communities. Further, reefs at San Nicolas Island have been the focus of more than three decades of ecological monitoring, and the resulting data represent a valuable tool for understanding kelp forest community dynamics.

The data presented here come from a monitoring program that was initiated in 1980 in anticipation of the potential reintroduction of sea otters (Rathbun et al. 1990, Rathbun and Benz 1991). San Nicolas Island had been identified as a potential site for the establishment of a new breeding population of sea otters (sea otters from central California were translocated to the island starting in 1987). Subtidal monitoring was initiated in autumn of 1980 at six sampling stations positioned around the island to capture communities exposed to a range of wave exposures and other oceanographic conditions (Table 1). One additional station was added in 1986 (chosen to add a location where the bottom had very little physical relief). The intent of this monitoring was to provide baseline information on ecological community structure in order to better evaluate the effects of sea otter introduction. Each station has been sampled twice a year (with a few exceptions). This period of monitoring covers substantial variability in environmental conditions, including two of the strongest El Nino's on record, years with high and low waves, and major shifts in the Pacific decadal oscillation.

In this paper, we present time-series data on the abundance of more than 200 species of fish, algae and invertebrates. These data were collected at each permanent sampling station using a combination of swath counts and point-contact sampling. In addition, we present measurements of the sizes of *Macrocystis pyrifera* (giant kelp), a key foundation species. The long time series and biannual sampling make these data appropriate for exploring ecological dynamics on multiple time scales. The high taxonomic

resolution of the sampling provides a detailed picture of community structure over time, and the fact that fixed quadrats and swaths have been sampled repeatedly means that changes in the community can be examined without the obscuring effects of sampling error. Taken together, these characteristics make these data valuable for investigating the ecology of giant kelp forests, particularly with respect to patterns and drivers of community dynamics across a wide range of physical and biological conditions

---

METADATA CONTENTS

Class I. Data set descriptors

- **E. Data set identity**
- **F. Data set identification code**
- **G. Data set description**
  - **1. Data set originator**
  - **2. Abstract**
- **H. Key words**

Class II. Research origin descriptors

- E. **Project description**
  - **1. Objectives**
  - **2. System description**
  - **3. Guidelines for use of data**
- F. **Monitoring design**
  - **6. Site selection and layout**
  - **7. Human impacts**
  - **8. Monitoring method selection**
  - **9. Choice of species to be monitored**
  - **10. Monitoring design metadata**
- G. **Data processing and presentation**
  - **4. Core community structure data sets**
  - **5. Supplementary data sets**
  - **6. Additional documentation and data**
- H. **Monitoring methods and data**
  - **1. Core community structure data sets:**
    - **d. Benthic density data**
      - **i. Benthic density summary procedures**
      - **ii. Swath count methods**
    - **e. Benthic cover data**
      - **i. Benthic cover summary procedures**
      - **ii. Point contact monitoring methods**
    - **f. Fish density data**
      - **i. Fish density summary procedures**
      - **ii. Fish transect monitoring methods**
  - **2. Supplementary data sets**
    - **a. Giant kelp size-frequency data**
      - **i. Giant kelp size-frequency field methods**

Class III. Data set status and accessibility

- **E. Latest data update**
- **F. Latest metadata update**
- **G. Data verification**
- **H. Copyright or proprietary restrictions**

Class IV. Data set structural descriptors

- **A. Data files**
- **B. Metadata tables**

Class V. Supplemental descriptors

- **D. Location of completed data forms**
- **E. Data entry verification procedures**
- **F. Publications using the data set**

Class I. Data set descriptors

- **A. Data set identity**
- **B. Data set identification code**
- **C. Data set description**
  1. **Data set originator**
  2. **Abstract**
- **D. Key words**

I.A. Data set identity: Rocky reef community structure at San Nicolas Island, California

I.B. Data Set Identification Code: KennerEtAl2013-SNI

I.C. Data set description

I.C.1. Data set originators

Individual: James A. Estes
Role: Initial principal investigator
Organization: University of California, Santa Cruz
Position: Professor of Ecology and Evolutionary Biology
Address: Long Marine Lab, 100 Shaffer Road, Santa Cruz, CA 95060
Phone: 831-459-2820
Email: jestes@ucsc.edu

Individual: Michael C. Kenner
Role: Chief field biologist
Organization: University of California Santa Cruz
Position: Biologist
Address: Long Marine Lab, 100 Shaffer Road, Santa Cruz, CA, 95060
Phone: 831-254-5184
Email: mkenner@ucsc.edu

Individual: M. Tim Tinker
Role: Current principal investigator
Organization: US Geological Survey, Western Ecological Research Center
Position: Research Biologist
Address: Long Marine Lab, 100 Shaffer Road, Santa Cruz, CA, 95060
Phone: 831-459-2357
Email: ttinker@usgs.gov

I.C.2. Abstract:

San Nicolas Island is surrounded by broad areas of shallow subtidal habitat, characterized by dynamic kelp forest communities that undergo dramatic and abrupt shifts in community composition. Although these reefs are fished, the physical isolation of the island means that they receive less impact from human activities than most reefs in Southern California, making San Nicolas an ideal place to evaluate alternative theories about the dynamics of these communities. Here we present monitoring data from seven sampling stations surrounding the island, including data on fish, invertebrate and algal abundance. These data are unusual among subtidal monitoring data sets in that they combine relatively frequent sampling (twice per year) with an exceptionally long time series (since 1980). Other outstanding qualities of the data set are the high taxonomic resolution captured and the monitoring of permanent quadrats and swaths where the history of the community structure at specific locations has been recorded through time. Finally, the data span a period that includes two of the strongest ENSO events on record, a major shift in the Pacific decadal oscillation and the reintroduction of sea otters to the island in 1987 after at least 150 years of absence. These events provide opportunities to evaluate the effects of bottom up forcing, top down control and physical disturbance on shallow rocky reef communities.

I.E. Key words: *long-term monitoring, California Channel Islands, rocky reef ecology, sea otters*

Class II. Research origin descriptors

A. Project description

1. **Objectives**
2. **System description**
3. **Guidelines for use of data**

II.A.1. Objectives: In the early 1980s, as part of the United States Fish and Wildlife Service's research program on southern sea otters, it was decided that a breeding population of sea otters should be established at San Nicolas Island to reduce the species' vulnerability to extinction caused by increasing threats from human activities. San Nicolas was selected because of its isolation and high availability of suitable habitat. In anticipation of the establishment of a new population, the U.S. Fish and Wildlife Service and the University of California, Santa Cruz established six permanent subtidal sampling stations around San Nicolas Island in 1980. One additional station was set up in 1986. It was chosen because the site had virtually no bottom relief and thus kelp, when present, would represent the only three dimensional structure affecting fish density. The study was designed to examine the possible impact of a translocated sea otter population on the local reef communities. The project is currently conducted by the U.S. Geological Survey - Western Ecological Research Center, U.C. Santa Cruz Field Station.

II.A.2. System description: San Nicolas Island (Ventura County, CA, USA) is the most remote of southern California's Channel Islands, both from shore and from the nearest neighboring island, lying approximately 140 km west-south-west of Los Angeles (~110 km offshore, 33°15' N, 119°30' W). San Nicolas is influenced both by the cold southward flowing California Current and the warmer northward edge of the California counter-current; consequently, its kelp forest communities contain biota characteristic of both more northern and more southern regions. The island itself is relatively small (less than 60 km^2 in area) but because of its gently sloping shelf of low-relief soft-sedimentary rock, it has the greatest area of kelp beds of any of the Channel Islands (30% of the total kelp coverage throughout the islands; Engle 1994). San Nicolas is located far from the wave shadows of other islands and the shape of its coastline provides little protection from the swell, so most of its shallow-water habitats are exposed to wave disturbance. Because of its isolation from the populated mainland, shallow subtidal communities at San Nicolas are less exposed to human influences than communities on many reefs in Southern California. However, reefs surrounding the island are the focus of several active commercial fisheries (particularly for sea urchins, lobster and rockfish), modest recreational fishing pressure (particularly for rockfish, ocean whitefish and sheephead) and are potentially impacted by the US Navy's operations on the island.

II.A.3. Guidelines for use of data: This data set is part of an ongoing monitoring program at San Nicolas Island run by the United States Geological Survey Western Ecological Research Center (USGS-WERC).

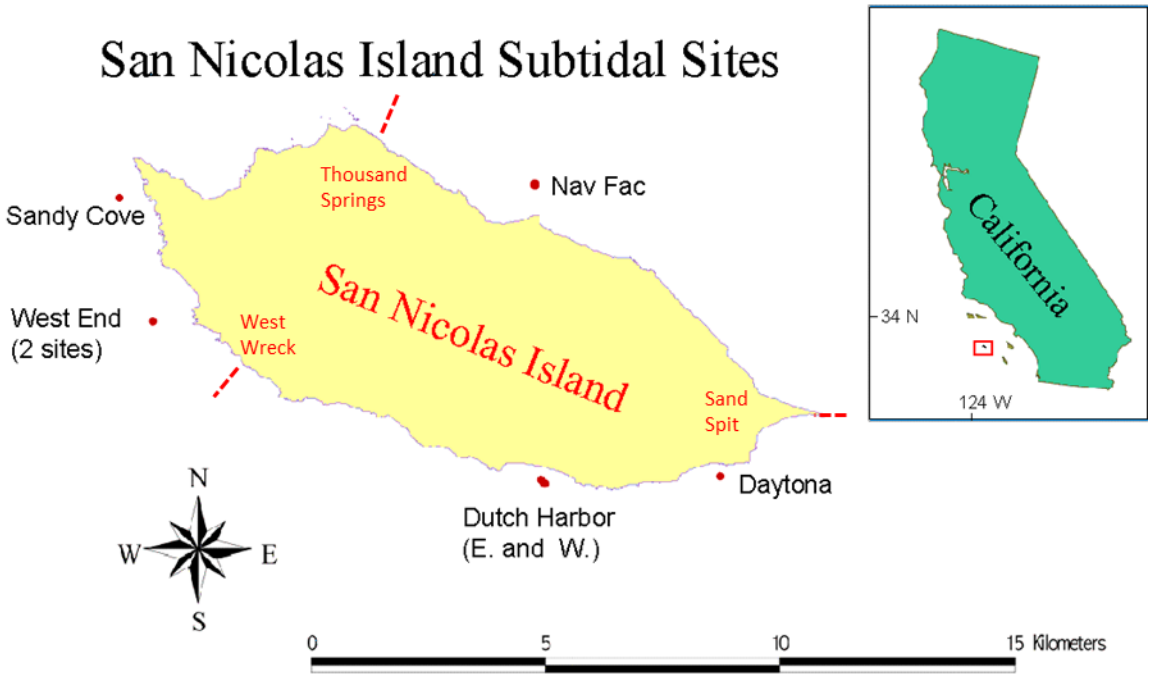
B. Monitoring Design

1. **Site selection and layout**
2. **Human impacts**
3. **Monitoring method selection**
4. **Choice of species to be monitored**
5. **Monitoring design metadata**

II.B.1. Site Selection and Layout: Seven stations have been chosen for monitoring (six in 1980 and one in 1986). These stations are placed around San Nicolas to provide data from each side of the island and to be representative of the available kelp forested habitat. A map of the stations is given in Figure 1. Three

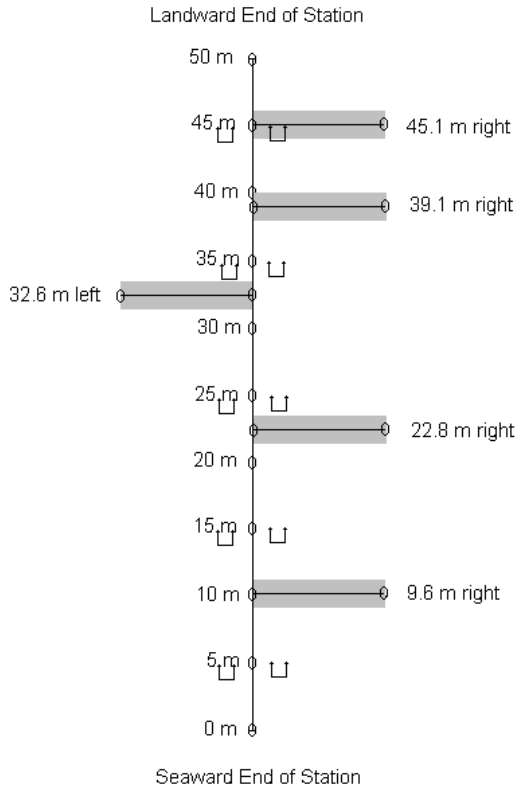
stations are isolated (Sandy Cove, Nav Fac and Daytona), while the other four were placed as two pairs. The pair of stations at West End was chosen to represent an urchin dominated area of reef and a kelp dominated area (although the communities at these stations have changed through time). These two stations are contiguous (the 50 m end of West End Kelp is the zero bolt of West End Urchins). The pair of sampling stations at Dutch Harbor was chosen as experimental (west) and control (east) stations for a sheephead removal experiment run in 1980 through 1982 (Cowen 1983). These two stations are separated by ~140 m. The latitude and longitude of each sampling station are given in Table 1. With a few exceptions, each station has been monitored twice a year (in spring and autumn) since its establishment.

Figure 1: Map of SNI monitoring stations.



At each station, monitoring is focused on a permanent 50-meter transect marked by stainless steel eyebolts placed at 5 m intervals. Each transect is placed approximately perpendicular to the coastline between 10 and 15 meters in depth. Associated with each transect are five permanent 10 m by 2 m swaths running perpendicular to it at fixed intervals. The original orientation of swaths (to the left or right of the transect) was chosen at random. Additionally, ten 1 m^2 quadrats are permanently placed one meter to the right and left of the main transect along its length. These quadrats mark locations where biota is sampled using fixed point contact methods. Finally, five 50 m fish transects are associated with each sampling station. Each fish transect starts at one of the permanent eyebolts and is laid out along a specific compass heading. Figure 2 illustrates the layout of quadrats and swaths along a typical transect.

Figure 2. An example of the layout of quadrats and swaths at a typical sampling station.



II.B.2. Human impacts: A number of significant anthropogenic impacts have occurred at the monitoring stations since their establishment. Perhaps the most important was the translocation of southern sea otters (*Enhydra lutris nereis*) from central California to SNI in 1987-1990. During this period, 140 sea otters were released at SNI. Although there was relatively low retention of these translocated animals, a population of ~15 adult animals remained at the island from 1990 to 1998, and their numbers have been increasing consistently since (Hatfield 2003)

Sea otters are not evenly distributed around the island. Sea otter surveys have been done three to five times per year since 1995, and the sampling date with the highest number of sea otters observed is a good measure of their abundance in space and time (Rathbun et al. 2000). When the coastline is divided into three roughly equivalent segments: the west (West Wreck to Thousand Springs, ~ 11.5 km), the north (Thousand Springs to Sand Spit, ~12 km) and the south (Sand Spit to West Wreck, ~ 13 km), it is apparent that independent sea otters (e.g. older juveniles and adults) are consistently abundant in the west and rare in the north (Table 2). This pattern is also observed in sea otter pups, which are common in the west, occasional in the south and effectively absent from the north (Table 3).

Several other human impacts at individual sampling stations are also worth noting. In September 1980, as part of an experimental manipulation, 220 sheephead (*Semicossyphus pulcher*) were removed from Dutch Harbor West. Cowen (1983) estimated that this manipulation eliminated all sheephead at the site. This manipulation was maintained for 2 years, with additional sheephead removed at bimonthly intervals (10-20 removed per visit). In addition to this scientifically motivated removal, substantial recreational and

commercial fishing occurs at the island. Although fine scale spatial data on fishing pressure are not available, several long term data sets document changes in fishing in southern California (Perry et al 2010).

The U.S. Navy's operations on San Nicolas also had significant effects on nearshore kelp forests. Until 2005 when the new pier opened, the navy brought barges in on the beach at Daytona Beach resulting in considerable sand movement, and relatively heavy sediment loads at that site. Since 2005, this practice has stopped.

In addition to human impacts, other factors have affected the continuity of data collection. In particular, the transect at Daytona could not be located in the fall of 1983, probably because of the sand input described above. Swaths 39R and 45R were never relocated and were eventually replaced with 22L and 39L in fall of 1985.

II.B.3. Monitoring method selection: The great diversity of organisms and physical habitats associated with kelp forests requires multiple sampling approaches to effectively monitor species population dynamics. Sampling methods were designed to maximize accuracy and precision while preserving an observer's ability to efficiently sample many target species at once. Counts within fixed areas are made to describe the abundance of fish as well as large invertebrates and algae. Estimates of percent cover are used to measure the abundance of species for which individuals are numerous, small or otherwise difficult to count.

II.B.4. Choice of species to be monitored: The monitoring captures the abundance of most algae, invertebrates and fish larger than a few centimeters, including all primary benthic spaceholders. The species sampled and data sets in which they appear are found in Table 4.

II.B.5. Monitoring design metadata: The following metadata are descriptive of the monitoring methods and context and are available below as comma separated values files.

Table 1: Monitoring stations and their Coordinates

Table 2: Abundance of independent sea otters

Table 3: Abundance of sea otter pups

Table 4: Species sampled

C. Data processing and presentation

1. **Core community structure data sets**
2. **Supplementary data sets**
3. **Additional documentation and data**

II.C.1. Core community structure data sets: The subtidal monitoring data presented here describe the algal, invertebrate and fish components of the kelp forest community at as high a taxonomic resolution as is practical. We present this information in four "core community structure data sets": benthic density (calculated from counts of organisms within fixed 10 m x 2 m swaths), benthic cover (calculated from point contact data in fixed 1 m X 1 m quadrats), midwater fish density (calculated from 50 m x 5 m fish transects), and benthic fish density (calculated from 50 m x 2 m fish transects). Each of these data sets is presented in both summarized form (with means and standard errors for each station) and in raw form (with counts from each sampling replicate reported separately).

II.C.2. Supplementary data sets: In addition to the core data discussed above, supporting data on *Macrosystis pyrifera* populations are also included. The basal diameters of a subset of giant kelp plants are measured and, in some cases, stipes have also been counted. Additionally, some giant kelp plants have been given unique tags, allowing the size of individual plants to be tracked over time. These data are collected biannually with the core community data.

II.C.3. Additional documentation and data: The subtidal data presented here are part of a larger monitoring program that includes intertidal community data collected twice a year and sea otter population surveys done three to five times a year. Users interested in these data should contact:

M. Tim Tinker
USGS-WERC, Santa Cruz Field Station, Long Marine Lab
100 Shaffer Road, Santa Cruz, CA, 95076
ttinker@usgs.gov

D. Monitoring methods and data

1. Core community structure data sets:

a. Benthic density data

i. Benthic density summary procedures

ii. Swath count methods

b. Benthic cover data

i. Benthic cover summary procedures

ii. Point contact monitoring methods

c. Fish density data

i. Fish density summary procedures

ii. Fish transect monitoring methods

II.D.1.a. Benthic density data

Table 5A: Definition of column headers in the benthic density raw data

Table 5B: Definition of variables in the benthic density raw data

Benthic density raw data: **ASCII file in comma separated values format**

Table 6A: Definition of column headers in the benthic density summary data

Table 6B: Definition of variables in the benthic density summary data

Benthic density summary data: **ASCII file in comma separated values format**

II.D.1.a.i. Benthic density summary procedures: Density data for solitary benthic algae and macro-invertebrates are collected by counting the number of individuals in a fixed swath and dividing that count by the area sampled. For each sample date and station, density estimates are reported for each swath in the "raw" data set and as a mean and standard error across all swaths in the "summary" data set.

II.D.1.a.ii. Swath count methods: 19 target species (Table 4) are sampled by counting the number of individuals of each that occur within five 10 m by 2 m swaths at each sampling station. Swaths are laid out perpendicular to the main transect at specified intervals (see Figure 2 for an example of swaths

placement). On each swath, divers count seven species or categories of brown algae (primarily kelps) and 12 species of macro-invertebrates. Only organisms that can be seen without extensive or destructive searching are counted.

II.D.1.b. Benthic cover data

Table 7A: Definition of column headers in the benthic cover raw data

Table 7B: Definition of variables in the benthic cover raw data

Benthic cover raw data: **ASCII file in comma separated values format**

Table 8A: Definition of column headers in the benthic cover summary data

Table 8B: Definition of variables in the benthic cover summary data

Benthic cover summary data: **ASCII file in comma separated values format**

II.D.1.b.i. Benthic cover summary procedures: Percent cover data for the benthic community of algae and sessile invertebrates are collected using a fixed point contact method in which a series of points is superimposed over the bottom, and the species intersecting each point is recorded. The number of times each species is contacted is divided by the total number of points sampled and multiplied by 100 to give the percent of the bottom occupied by that species. Percent cover of each species is calculated separately for each of the 10 1 m^2 quadrats. For each sample date and station, percent cover estimates for each taxon are reported for each quadrat in the "raw" data set and as a mean and standard error calculated across all quadrats in the "summary" data set. Note that the raw (quadrat level) estimates of cover are based on 20 points each, so give cover estimates in increments of 5% (one point => 5% cover). The summary (station level) data gives cover estimates in increments of .05% (one point => .05% cover). Thus, while the raw data are very useful for looking at small scale associations between common species, the increased precision given by the summary data make them more appropriate for describing population dynamics, particularly of rare species.

II.D.1.b.ii. Point contact monitoring methods: Benthic percent cover data are collected in 1 m^2 permanent quadrats, with a quadrat placed 1 m to the left or right of the main 50 m transect at 10 fixed locations. Within each quadrat 20 points are distributed in a fixed pattern. The diver imagines a line running vertically through each point up to one meter above the substratum and identifies all organisms that intersect this imaginary line. The diver counts each species only once per point even if multiple individuals of the same species intersect that point. Because of this method of counting multiple layers, total cover of all species can exceed 100% (and usually does), but the cover of any individual species cannot be > 100%. The list of taxa recorded is open and ranges from actual individual species to species groups like "orange encrusting sponge". A few organisms which could not be identified have been recorded as "Unidentified spp." Substrate type is also recorded if exposed.

II.D.1.c. Fish density data

Table 9A: Definition of column headers in the midwater fish density raw data

Table 9B: Definition of variables in the midwater fish density raw data

Midwater fish density raw data: **ASCII file in comma separated values format**

Table 10A: Definition of column headers in the midwater fish density summary data

Table 10B: Definition of variables in the midwater fish density summary data

Midwater fish density summary data: **ASCII file in comma separated values format**

Table 11A: Definition of column headers in the benthic fish density raw data

Table 11B: Definition of variables in the benthic fish density raw data

Benthic fish density raw data: **ASCII file in comma separated values format**

Table 12A: Definition of column headers in the benthic fish density summary data

Table 12B: Definition of variables in the benthic fish density summary data

Benthic fish density summary data: **ASCII file in comma separated values format**

II.D.1.c.i. Fish density summary procedures: Fish density data have been collected since autumn 1981 by counting the number of individuals within fixed transects and dividing that count by the area of bottom sampled. At each sampling date, fish are counted in five fish transects per sampling station. A diver passes over each transect twice, counting midwater fish in a 5 m wide transect on the first pass and benthic fish in a 2 m wide transect on the second. These data are presented separately here, because the same individuals could be present in both counts depending on their behavior. Counts are presented as densities per m^2 of horizontal area sampled. For each sample date and sampling station, density estimates are reported for each fish transect separately in the "raw" data sets and as a mean and standard error across all transects in the "summary" data sets. On one occasion (at Daytona Beach in October 1988) a huge school of California anchovy (*Engraulis mordax*) was observed on two transects, but they were too numerous to get an estimate of their abundance. They have been omitted from the data.

II.D.1.c.ii. Fish transect monitoring methods: Five 50 m visual fish transects are sampled at each station on each sampling date. Each of these transects is defined by a beginning point on the main transect and a compass heading, so that the same areas are counted each sampling period (see Table 11 for fish transect placement). On each transect, two counts are made. First the observer swims the entire transect about 2-3 m off the bottom, sampling a 5 m wide transect and counting all fish found between the surface and 2 m above the bottom. Sheephead and all schooling fish (e.g. blacksmith) are counted even if they are within 2 m of the bottom. Fish observed in this count are recorded as midwater fish; the total horizontal area sampled is 250 m^2. After reaching the end of the transect, the observer works backwards more slowly, counting all fish found in a 2 m wide swath within 2 m of the bottom. Midwater fish (such as sheephead) are excluded from this count, even if they are found within 2 m of the bottom. This is slow and careful work, because many of these fish are mixed in with the benthic algae, so despite the smaller area sampled, this second count typically takes longer than the first. Fish observed in the second count are recorded as benthic fish. The total horizontal area sampled is 100 m^2. In each count, an open species list is used, so all fish encountered are sampled. Fish are identified to the highest taxonomic resolution possible in the field, typically to species or genus. In both counts, juveniles are scored separately when observed, and male and female sheephead (*Semicossyphus pulcher*) are recorded separately.

Table 13: Starting point on the main transect and orientation of fish transects

Station	StationName	1	2	3	4	5
1	Nav Fac	0 m E	10 m W	20 m E	30 m W	40 m W
2	West End, Urchin	0 m N	10 m N	20 m N	30 m N	ON
3	West End, Kelp	0 m S	10 m N	20 m S	30 m N	ON
4	West Dutch Harbor	0 m W	10 m W	30 m W	45 m W	ON
5	East Dutch Harbor	0 m E	10 m W	20 m E	30 m W	40 m E
6	Daytona Beach	0 m E	0 m W	10 m E	20 m W	30 m E
7	Sandy Cove	0 m W	10 m W	20 m E	30 m E	ON

Note: "ON" means the fish transect was performed along the main transect, N, E, S and W refer to compass headings of 0, 90, 180 and 270 degrees respectively.

II.D.2. Supplementary data sets

a. Giant kelp size-frequency data

i. Giant kelp size-frequency field methods

II.D.2.a. Giant kelp size-frequency data:

Table 14A: Definition of column headers in the giant kelp size-frequency data

Table 14B: Definition of variables in the giant kelp size-frequency data

Giant kelp size-frequency data: **ASCII file in comma separated values format**

IV.D.2.a.i. Giant kelp size-frequency methods: Beginning in spring 1981, the basal diameters (holdfast width) of all *Macrocystis pyrifera* plants greater than 1 m tall encountered on swaths have been measured. In some cases the number of stipes on each plant has been counted as well. Stipe counts are typically done on the first two swaths at each sampling station with kelp present, but counts have not been made in all such cases. To obtain information on patterns of recruitment, mortality and lifespan of giant kelp, some individuals were tagged with unique numbers, written on plastic tape and cable-tied loosely around the primary dichotomous branch above the holdfast and followed over time. The number of tagged plants varied through time, sometimes including all plants > 1 m tall, and at other times including only the first 10 plants encountered on each swath. Tagging was discontinued in 2004 to reduce overall sampling effort. Missing holdfast sizes, frond counts or tag numbers are noted with "NA" in the data set.

Class III. Data set status and accessibility

I. **Latest data update**
J. **Latest metadata update**
K. **Data verification**
L. **Copyright or proprietary restrictions**

III.A. Latest data update: The data set may be periodically updated. All updates to the data have been logged in Table 15A. Please check for the latest update before using the data set.

III.B. Latest metadata update: The metadata may also be updated periodically. All updates have been logged in Table 15B. Please check for the latest update before using the data set.

III.C. Data verification: Field data sheets are proofed for accuracy after every day in the field as well as during data entry.

III.D. Copyright or proprietary restrictions: None

Class IV. Data set structural descriptors

C. **Data files**
D. **Metadata tables**

IV.A. Data Files:

Benthic density raw data

Benthic density summary data

Benthic cover raw data

Benthic cover summary data

Midwater fish density raw data

Midwater fish density summary data

Benthic fish density raw data

Benthic fish density summary data

Giant kelp size-frequency data

IV.B. Metadata Files:

Class V. Supplemental descriptors

A. Location of completed data forms
B. Data entry verification procedures
C. Publications using the data set

V.A. Location of completed data forms: USGS-WERC, Santa Cruz Field Station, Long Marine Lab 100 Shaffer Road, Santa Cruz, CA, 95076.

V.B. Data entry verification procedures: See III.C.

V.C. Publications using the data set prior to 2013:

Cowen R. 1983. The effects of sheephead (*Semicossyphus pulcher*) predation on red sea urchin (*Strongylocentrotus franciscanus*) populations: an experimental analysis. Oecologia 58:249-255.

Cowen RK, Bodkin JL. 1993. Annual and spatial variation of the kelp forest fish assemblage at San Nicolas Island, California. Proceedings of the Third California Islands Symposium:463–474.

Cowen RK. 1986. Site-specific differences in the feeding ecology of the California sheephead *Semicossyphus pulcher* (Labridae). Environmental Biology of Fishes 16:193-203.

Harrold C, Reed DC. 1985. Food availability, sea urchin grazing, and kelp forest community structure. Ecology 66:1160–1169.

Hatfield BB. 2005. The translocation of sea otters to San Nicolas Island: An update. Proceedings of the Sixth California Islands Symposium:473-475.

Newsome SD, Bentall GB, Tinker MT, Oftedal OT, Ralls K, Estes JA, Fogel ML. 2010. Variation in delta 13C and delta 15N diet-vibrissae trophic discrimination factors in a wild population of California sea otters. Ecological Applications 20:1744-1752.

Perry WM, Gustafson KB, Sanders GS, Takekawa JY. 2010. Pacific Coast Fisheries GIS Resource Database. U.S. Geological Survey, Western Ecological Research Center, Dixon and Vallejo, CA and Bureau of Ocean Energy Management, Regulation and Enforcement, Camarillo, CA. BOEMRE study profile #2010-023. [CD-ROM]

Rathbun GB, Hatfield BB, Murphey TG. 2000. Status of translocated sea otters at San Nicolas Island, California. The Southwestern Naturalist 45:322.

Schroeter SC, Reed DC, Kushner DJ, Estes JA, Ono DS. 2001. The use of marine reserves in evaluating the dive fishery for the warty sea cucumber (*Parastichopus parvimensis*) in California, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 58:1773-1781.

Tinker MT, Bentall G, Estes JA. 2008. Food limitation leads to behavioral diversification and dietary specialization in sea otters. Proceedings of the National Academy of Sciences of the United States of America 105:560-565.

Tinker MT, Guimarães PR, Novak M, Marquitti FMD, Bodkin JL, Staedler M, Bentall G, Estes JA. 2012. Structure and mechanism of diet specialization: testing models of individual variation in resource use with sea otters. Ecology Letters 15(5) 475-483.

Zacharias MA, Kushner DJ. 2006. Sea temperature and wave height as predictors of population size structure and density of *Megastraea (Lithopoma) undosa*: implications for fishery management. 79:71-82.

Acknowledgements

Funding for fieldwork and data collection was provided by the US Geological Survey, US Fish and Wildlife Service and University of California Santa Cruz. Logistical support for fieldwork and access to the waters around San Nicolas Island were provided by the US Navy and facilitated by Grace Smith, John Ugoretz and Ron Dow of the Naval Air Warfare Center Weapons Division Pt. Mugu Sea Range Sustainability Office. We are grateful to the dozens of volunteers and agency divers that contributed to field work over the years, in particular Linda Brown, Ken Collins, Keith Miles and Amy Story. During the preparation of this data paper A.R. was supported with funds from the U.S. Department of the Interior, Bureau of Ocean Energy Management, Environmental Studies Program, Washington, DC under Cooperative Agreement Number M11AC00012. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Government. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Government. This product (article, paper, etc.) has been peer reviewed and approved for publication consistent with USGS Fundamental Science Practices (http://pubs.usgs.gov/circ/1367/).

Literature cited

Cowen R. 1983. The effects of sheephead (*Semicossyphus pulcher*) predation on red sea urchin (*Strongylocentrotus franciscanus*) populations: an experimental analysis. Oecologia 58:249-255.

Engle JM. 1994. Perspectives on the structure and dynamics of nearshore marine assemblages of the California Channel Islands. Proceedings of the Fourth California Islands Symposium:13-26.

Hatfield BB. 2005. The translocation of sea otters to San Nicolas Island: An update. Proceedings of the Sixth California Islands Symposium:473-475.

Rathbun GB, Jameson RJ, VanBlaricom GR, Brownell RL Jr. 1990. Reintroduction of sea otters to San Nicolas Island, California: Preliminary results for the first year. Pages 99-114 in P. J. Bryant and J. Remington, editors. Memoirs of the Natural History Foundation of Orange County: Endangered wildlife and habitats in southern California.

Rathbun GB, Benz CT. 1991. Third year of sea otter translocation completed in California. Endangered Species Technical Bulletin XVI(3):1, 6-8.

Tinker MT, Bentall G, Estes JA. 2008. Food limitation leads to behavioral diversification and dietary specialization in sea otters. Proceedings of the National Academy of Sciences of the United States of America 105:560-565.

USFWS (United States Fish and Wildlife Service). 1982. Southern Sea Otter Recovery Plan. USFWS, Portland, Oregon.

Appendix C

All files and folders referenced in Appendix C are found in the file:

OCS Study BOEM 2019-063 Appendix C files.zip

DOI partnership integrated data package - Metadata

DATA SET OVERVIEW

As part of the DOI partnership project (Cooperative Agreement Number M11AC00012) we have produced a group of coordinated data sets which include biological data from four different monitoring programs: the National Park Service's kelp forest monitoring program (KFM), the San Nicolas Island monitoring program initiated by US Geological Survey (SNI), the privately funded Partnership for the Interdisciplinary Studies of Coastal Oceans (PISCO) reef monitoring program, and the NSF funded Santa Barbara Coastal Long-Term Ecological Research program (LTER). These data have been supplemented with environmental data from a range of sources including output from remote sensing analyses and regional models. Important community metrics such as measures of biodiversity have also been calculated from the biological data and presented in additional data sets.

A key challenge to producing these data sets has been to determine what biological variables have been measured in comparable ways by each program. In particular, the taxonomic resolution and the range of species counted varies widely from program to program. Our focus here was to produce data sets that are ready for analysis; thus, it is important that each species (or taxonomic group) included be counted using comparable methods by all data sources that are being synthesized. As a consequence, the taxonomic list for each integrated data set can only include species that are shared across all source data. Each data set included lacks some taxa (or taxonomic resolution) contained in the other data sources; therefore, the greater the number of sources included, the smaller the resulting usable taxa. We therefore include five versions of the biological data here, a core data set including all four sources (tagged as KFM_LTER_PISCO_SNI), and four additional data sets - representing all three-way combinations of these sources. Although we do not include all pairwise combinations in this report, the computer code we attach will produce any pairwise combination except for the LTER-SNI combination (this pair is unlikely to be used in isolation because of the geographic layout of the sites, with LTER in the north, SNI in the south and PISCO and KFM in between).

Inside the attached file "OCS Study BOEM 2019-063 Appendix C files.zip" is a folder named "DATA_SETS" containing 5 subfolders, each containing one version of the biological data. Subfolders are named:

KFM_LTER_PISCO_SNI
KFM _PISCO_SNI
KFM_LTER_ SNI
KFM_LTER_PISCO
LTER_PISCO_SNI

Within each version of the data set, we present three primary biological data sets, a “fish” data set representing counts of fish and the area sampled, a “quadswath” data set representing counts of algae and invertebrates and the area sampled, and finally a “benthic” data set containing the percent cover of different taxonomic groups. We also present diversity data sets for each of these sampling types, which include values for richness, Shannon, and Simpson’s diversity indices calculated from the included taxa. Throughout, data are presented from 2000 onward, as both PISCO and LTER data do not start until the early 2000s.

Within each folder are files for the specified combination of data sources with the prefixes:

benthic_Post2000_XXXXXX.csv – percent cover data (by replicate)
benthic_diversity_XXXXXX.csv – biodiversity metrics for percent cover data (by sampling event)
fish_Post2000_XXXXXX.csv – fish count data (by replicate)
fish_diversity_XXXXXX.csv – biodiversity metrics for fish count data (by sampling event)
quadswath_Post2000_XXXXXX.csv – count data for algae and invertebrates (by replicate)
quadswath_diversity_XXXXXX.csv – biodiversity metrics for algae/invertebrate count data (by sampling event)

In each case the suffix “XXXXXX” indicates the data sources included and the date the data was processed. For example “KLPS_20190604” indicates that all 4 data sources are included (the 4 initials) and that the data was processed June 4, 2019.

To supplement these biological data, we present environmental data sets that have been processed to match the spatial and temporal patterns of sampling (three environmental data sets, one matching “fish”, one for “quadswath”, and one for “benthic”). These data sets include information about sea surface temperature, wave climate, biomass of giant kelp, and sea surface chlorophyll in the area immediately surrounding a sampling site. Environmental data are averaged over two temporal scales; the six months prior to a sampling date, and the 12 months prior to a sampling date.

Inside the attached file “OCS Study BOEM 2019-063 Appendix C files.zip” there is a folder named “environmental” which contains the following files:

benthic_environmental_by_subsite_XXXXXXXX.csv
fish_environmental_by_subsite_XXXXXXXX.csv
quadswath_environmental_by_subsite_XXXXXXXX.csv percent cover sampling events

NOTES ON CITATION AND ATTRIBUTION

The curated data products we produce here rely on extensive integration and processing done by researchers at the Santa Barbara Channel Marine Biodiversity Observing Project, particularly Dr. Li Kui. Any use of this data should acknowledge their contribution, for example:

Miller R, Rassweiler AR, Caselle JE, Kushner DJ, Reed D, Lafferty KD, et al. Santa Barbara Channel Marine BON: Nearshore kelp forest integrated fish, 1981-ongoing. Environmental Data Initiative; 2018.

Ultimately the data presented here are based on laborious collection and processing done by teams of researchers at the long term research programs which are the source of these data. Any use of these data must cite the underlying data sets directly. Appropriate citations are listed here for convenience:

Kenner, Michael C., et al. "A multi-decade time series of kelp forest community structure at San Nicolas Island, California (USA)." Ecology 94.11 (2013): 2654-2654.

Kushner, David J., et al. "A multi-decade time series of kelp forest community structure at the California Channel Islands." Ecology 94.11 (2013): 2655-2655.

PISCO, Mark Carr, and Jenn Caselle. 2009a. Channel Islands OPC: Subtidal: Community Surveys: Swath Surveys. PISCO MN. doi:10.6085/AA/CI_subtidal_swath.1.2.

PISCO, Mark Carr, and Jenn Caselle. 2009b. Channel Islands OPC: Subtidal: Community Surveys: Fish Survey. PISCO MN. doi:10.6085/AA/CI_subtidal_fish.1.2.

Reed, D. C. 2016a. SBC LTER: Reef: Kelp Forest Community Dynamics: Abundance and size of Giant Kelp (*Macrocystis pyrifera*), ongoing since 2000. Santa Barbara Coastal LTER. doi:10.6073/pasta/fa27c7877a083cf3ea5a9bde55dbcf06

Reed, D. C. 2016b. SBC LTER: Reef: Kelp Forest Community Dynamics: Fish abundance. Santa Barbara Coastal LTER. doi:10.6073/pasta/c3977bff17b302f12e2bb832642808a4

Reed, D. C. 2016c. SBC LTER: Reef: Kelp Forest Community Dynamics: Invertebrate and algal density. Santa Barbara Coastal LTER. doi:10.6073/pasta/62803c95783c4e771695d1c6cc3d23ac

Reed, D. C. 2016d. SBC LTER: Reef: Kelp Forest Community Dynamics: Cover of sessile organisms, Uniform Point Contact. Santa Barbara Coastal LTER. doi:10.6073/pasta/ca1a92c63e57993a09111f1e33dc9210

Lafferty, K.D., Morton, D.N., Gotschalk, C.C., Henderikx, F., Rassweiler, A., and Washburn, L., 2018, Hourly wave height and period hindcasts at 32 sites throughout the Channel Islands National Park and San Nicolas Island from 2000-2017: U.S. Geological Survey data release, https://doi.org/10.5066/P91KH2Q7.

In addition, the programs that collected the data should appear in the acknowledgments of any publication derived from this integrated data set. For example, PISCO requests the following text:

“This study utilized data collected by the Partnership for Interdisciplinary Studies of Coastal Oceans: a long-term ecological consortium funded by the David and Lucile Packard Foundation”.

BIOLOGICAL DATA SET COLUMN DESCRIPTORS

FISH

1. **auth_taxon_id** : the authoritative taxon_id for that taxa group
2. **filled count** : count of individuals for that taxa group at that transect on the sampling date
3. **phylum**
4. **class**
5. **order**
6. **family**

7. **genus**
8. **species**
9. **site_id** : identification code for each site
10. **subsite_id** : identification code for each subsite (e.g. a site-side sampling event for PISCO)
11. **data_source** : one of four sources (KFM, LTER, PISCO, SNI)
12. **sample _method** : method of sampling the fish (fish for SNI, PISCO, LTER, rdfc = roving diver fish count for KFM)
13. **sample_subtype** : subtype of sampling (BOT = bottom, CAN = canopy, MID = middle, FUL = full, CNMD = canopy and middle)
14. **date** : sampling date
15. **mean_area** : area of the transect (m^2)
16. **transect_id** : identification of the transect
17. **replicate_id** : identification of the replicate (for KFM this is the diver id number)
18. **year**
19. **taxon_name**

QUAD SWATH

1. **auth_taxon_id** : the authoritative taxon_id for that taxa group
2. **site_id** : identification code for each site
3. **subsite_id** : identification code for each subsite (e.g. a site-side sampling event for PISCO)
4. **data_source** : one of four sources (KFM, LTER, PISCO, SNI)
5. **sample _method** : method of sampling the area (*band, quad, swath, 5mquad, 1mquad*)
6. **date** : sampling date
7. **transect_id** : identification of the transect
8. **replicate_id** : identification of the replicate
9. **phylum**
10. **class**
11. **order**
12. **family**
13. **genus**
14. **species**
15. **sum_of_count** : count of individuals for that taxa group at that transect on the sampling date
16. **mean_area** : area of the transect (m^2); averaged only if aggregated taxa over a transect (but will be same area regardless)
17. **taxon_name**
18. **year**

BENTHIC

1. ***auth_taxon_id*** *: the authoritative taxon_id for that taxa group*
2. **filled count** : count of individuals for that taxa group at that transect on the sampling date
3. **kingdom**
4. **phylum**
5. **class**
6. **order**
7. **family**
8. **genus**
9. **species**

10. **site_id** : identification code for each site
11. **subsite_id** : identification code for each subsite (e.g. a site-side sampling event for PISCO)
12. **data_source** : one of four sources (KFM, LTER, PISCO, SNI)
13. **sample _method** : method of sampling the area (*rpc* = random point contact; *upc* = uniform point contact)
14. **date** : sampling date
15. **points** : number of sampling points along the area; averaged only if aggregated taxa over a transect (but will be same area regardless)
16. **transect_id** : identification of the transect
17. **replicate_id** : identification of the replicate
18. **year**
19. **taxon_name**

BIODIVERSITY

1. **site_id** : identification code for each site
2. **subsite_id** : identification code for each subsite (e.g. a site-side sampling event for PISCO)
3. **data_source** : one of four sources (KFM, LTER, PISCO, SNI)
4. **date** : sampling date
5. **richness** : species richness by sampling date per transect (count of species on the transect)
6. **shannon** : Shannon's diversity index (H) by sampling date per transect
7. **simpson** : Simpson's diversity index by sampling date per transect
8. **year**

ENVIRONMENTAL DATA SET COLUMN DESCRIPTORS

Note that the environmental data sets corresponding to each biological sampling method contain environmental data for all sampling events. The environmental data sets can be merged with the corresponding biological data sets by selecting rows from the environmental data that match in site_id, subsite_id, data_source, date.

ENVIRONMENTAL DATA SETS

1. **site_id** : identification code for each site
2. **subsite_id** : identification code for each subsite (e.g. a site-side sampling event for PISCO)
3. **data_source** : one of four sources (KFM, LTER, PISCO, SNI)
4. **date** : sampling date
5. **year** : year
6. **month** : month (in integers)
7. **day** : day
8. **sitevar_1** : Code used to describe the sampling site within the source data set (for KFM this is the site number; for LTER, SNI and PISCO it is the site name)
9. **sitevar_2** : for PISCO, this variable is the sampling site side (CEN = center + the cardinal directions (NSEW) from the site); for KFM and SNI this is the site name
10. **latitude** : latitude
11. **longitude** : longitude

12. **island** : variable that tells which island the sampling site is located (San Miguel, Santa Rosa, San Nicolas, Santa Cruz, Anacapa, Santa Barbara, San Clemente, Mainland)
13. **reserve**: reserve status for the site; 0 if the site is unprotected, 1 if the site is protected
14. **SST.6mo** : mean of the SST's from the previous 6 months (data by day) prior to the sampling date (C°); by subsite_id
15. **SST.12mo** : mean of the SST's from the previous 12 months (data by day) prior to the sampling date (C°); by subsite_id
16. **mean.wave.6mo** : mean of the mean wave height (m) from the previous 6 months (data by day) prior to the sampling date; by subsite_id
17. **mean.wave.12mo** : mean of the mean wave height (m) from the previous 12 months (data by day) prior to the sampling date; by subsite_id
18. **max.wave.6mo** : mean of the max wave height (m) from the previous 6 months (data by day) prior to the sampling date; by subsite_id
19. **max.wave.12mo** : mean of the max wave height (m) from the previous 12 months (data by day) prior to the sampling date; by subsite_id
20. **biomass.6mo** : mean of the wet kelp canopy biomass (kg/m^2) from the previous 6 months (data by month) prior to the sampling date; by subsite_id
21. **biomass.12mo** : mean of the wet kelp canopy biomass (kg/m^2) from the previous 12 months (data by month) prior to the sampling date; by subsite_id
22. **chloro.point.6mo** : mean of chlorophyll (mg/m^3) from the previous 6 months (data by month) prior to the sampling date; by subsite_id; point is at the lat/lon for the sampling subsite
23. **chloro.point.12mo** : mean of chlorophyll (mg/m^3) from the previous 12 months (data by month) prior to the sampling date; by subsite_id; point is at the lat/lon for the sampling subsite
24. **chloro.3km.buffer.6mo** : mean of chlorophyll (mg/m^3) from the previous 6 months (data by month) prior to the sampling date; by subsite_id; 3 km buffer around the lat/lon for the sampling subsite
25. **chloro.3km.buffer.12mo** : mean of chlorophyll (mg/m^3) from the previous 12 months (data by month) prior to the sampling date; by subsite_id; 3 km buffer around the lat/lon for the sampling subsite
26. **chloro.5km.buffer.6mo** : mean of chlorophyll (mg/m^3) from the previous 6 months (data by month) prior to the sampling date; by subsite_id; 5 km buffer around the lat/lon for the sampling subsite
27. **chloro.5km.buffer.12mo** : mean of chlorophyll (mg/m^3) from the previous 12 months (data by month) prior to the sampling date; by subsite_id; 5 km buffer around the lat/lon for the sampling subsite
28. **diffuse.attenuation.6mo.site** : mean of the attenuation coefficient for turbidity (m^{-1}) from the previous 6 months (data by month) prior to the sampling date; by site_id; at the lat/lon for the sampling site
29. **diffuse.attenuation.12mo.site** : mean of the attenuation coefficient for turbidity (m^{-1}) from the previous 12 months (data by month) prior to the sampling date; by site_id; at the lat/lon for the sampling site

30. **diffuse.attenuation.1km.6mo.site** : mean of the attenuation coefficient for turbidity (m^{-1}) from the previous 6 months (data by month) prior to the sampling date; by site_id; 1 km buffer around the lat/lon for the sampling site
31. **diffuse.attenuation.1km.12mo.site** : mean of the attenuation coefficient for turbidity (m^{-1}) from the previous 12 months (data by month) prior to the sampling date; by site_id; 1 km buffer around the lat/lon for the sampling site
32. **mean.Hs.corrected.6mo.site** : mean of corrected wave height (m; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
33. **meanlog.Hs.corrected.6mo.site** : mean of the fitted log normal distribution of the corrected wave height (m; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
34. **sdlog.Hs.corrected.6mo.site** : standard deviation of the fitted log normal distribution of the corrected wave height (m; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
35. **mean.Hs.corrected.12mo.site** : mean of corrected wave height (m; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
36. **meanlog.Hs.corrected.12mo.site** : mean of the fitted log normal distribution of the corrected wave height (m; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
37. **sdlog.Hs.corrected.12mo.site** : standard deviation of the fitted log normal distribution of the corrected wave height (m; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
38. **mean.OV.6mo.site** : mean of corrected bottom orbital velocity (m/s ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
39. **meanlog.OV.6mo.site** : mean of the fitted log normal distribution of the corrected bottom orbital velocity (m/s ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
40. **sdlog.OV.6mo.site** : standard deviation of the fitted log normal distribution of the corrected bottom orbital velocity (m/s ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
41. **mean.OV.12mo.site** : mean of corrected bottom orbital velocity (m/s ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
42. **meanlog.OV.12mo.site** : mean of the fitted log normal distribution of the corrected bottom orbital velocity (m/s ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
43. **sdlog.OV.12mo.site** : standard deviation of the fitted log normal distribution of the corrected bottom orbital velocity (m/s ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
44. **mean.WaveEnergy.6mo.site** : mean of corrected wave energy (kW/m ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
45. **meanlog.WaveEnergy.6mo.site** : mean of the fitted log normal distribution of the corrected wave energy (kW/m ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id

46. **sdlog.WaveEnergy.6mo.site** : standard deviation of the fitted log normal distribution of the corrected wave energy (kW/m ; CDIP models) from the previous 6 months (data by hour) prior to the sampling date; by site_id
47. **mean.WaveEnergy.12mo.site** : mean of corrected wave energy (kW/m ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
48. **meanlog.WaveEnergy.12mo.site** : mean of the fitted log normal distribution of the corrected wave energy (kW/m ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id
49. **sdlog.WaveEnergy.12mo.site** : standard deviation of the fitted log normal distribution of the corrected wave energy (kW/m ; CDIP models) from the previous 12 months (data by hour) prior to the sampling date; by site_id

COMPUTER CODES

In addition to the five analysis-ready data sets we supply inside the attached file "OCS Study BOEM 2019-063 Appendix C files.zip", that zip file also includes a folder named "Aggregation_Biodiversity" containing the computer codes used to produce these integrated data sets. Of particular interest are the count_fish, count_quadswath, and pct_cover R scripts. At the top of each of these scripts are user-friendly parameters that allow someone running the script to select which of the four source data sets are included in the output (using a zero to exclude the data and a 1 to include it). The fish and invertebrate processing scripts also include a parameter that allows the user to set the balance between keeping more observations and obtaining better taxonomic resolution (taxaRatioCutoff, described in the notes on biological data below).

These codes curate the data, and draw on a large integrated data table produced by the Santa Barbara Channel Marine Biodiversity Observing Program. To run the codes for aggregating and calculating the diversity metrics, open the corresponding code and set the working directory to the "Aggregation_Biodiversity" folder. Once in that working directory, all the data will be called in the code.

We included version of that integrated data table updated as of 3/31/18. It is also available online at the following DOIs.

doi:10.6073/pasta/1345f0148e6dfe4df9065e223b4dd783
doi:10.6073/pasta/bf143fa962e1edb822847bc0ee90c2f7
doi:10.6073/pasta/51d2db26e90d4b8687db81fb40bc58c4
doi:10.6073/pasta/d09d4bfd54e6d4e490b4cc34731d808e

Code – in Aggregation_Biodiversty folder

count_fish.R – script that merges and aggregates the fish count data based on user decisions of which data sets and cutoff to use
count_quadswath.R – script that merges and aggregates the algae/invertebrate count data based on user decisions of which data sets and cutoff to use
pct_cover.R – script that merges and aggregates the percent cover data based on user decisions of which data sets to use

biodiversity_data.R – script that calculates the species richness, Simpson's and Shannon's diversity index for every sampling event for fish, quadswath and percent cover data

Data – in Aggregation_Biodiversty folder

cover_integrated_XXXXXX.csv – integrated percent cover data set, most recently updated on 01/21/2018, but newer versions can be downloaded online (DOI)
fish_integrated_XXXXXX.csv – integrated fish count data set, most recently updated on 01/21/2018, but newer versions can be downloaded online (DOI)
quadswath_integrated_XXXXXX.csv – integrated algae/invertebrate count data set, most recently updated on 01/21/2018, but newer versions can be downloaded online (DOI)
TaxaTable_XXXXXXXX.csv – table of all the taxa codes and identifiers for each program, most recently updated 02/26/2018 (has had codes removed for things not being used i.e. substrate, superlayers, etc.)
taxaTableForSubstrate_XXXXXXXX.csv – table to be utilized for substrate data, most recently updated 01/21/2018

SUPPORTING COMPUTER CODES

Inside the attached file, "OCS Study BOEM 2019-063 Appendix C files.zip" is a folder named "Codes for integration data set", which contains all codes used to produce the integrated data from the underlying source data sets. Running these codes would require requesting source data sets from the underlying data owners; the codes are provided here for reference.

Inside the attached file, "OCS Study BOEM 2019-063 Appendix C files.zip" is a folder named "Environmental/Site Info" which contains all the code in the 1_Code subfolder and data used to produce the environmental data from the underlying data sets included in the 2_Data subfolder. Running the code (environmental_data.R – script that calculates all the environmental data for each sampling event in benthic, fish and algae/invertebrate) requires setting the working directory to the "Environmental/Site Info" folder.

NOTES ON BIOLOGICAL DATA SETS

For all data, we discarded any data pre-2000 as LTER and PISCO data only exist from 2000 on.

FISH

User choices

- ***Data Sources***: 0 = do not use this data source in the aggregation, 1 = use this data source in the aggregation
    - use_SNI = 1
    - use_KFM = 1
    - use_LTER = 1
    - use_PISCO = 1

- ***Taxa ratio cutoff***: This is the maximum fraction of observations within one taxonomic level that you would be willing to discard in order to reach the next finer level of taxonomic resolution.

Numbers closer to zero will aggregate data at higher taxonomic levels, but keep more of the observed fish. Larger numbers (closer to 1) will throw away more observations, but retain finer taxonomic resolution.

- taxaRatioCutoff = 0.2
- ** For all the fish file outputs, a taxa ratio cutoff of 0.2 was used by default, meaning we were willing to get rid of 20% of the data in order to move to a finer level of taxonomic resolution.

Manual decisions

- The fish data sets were zero filled so that all taxa are included at all sampling dates and transects. Because fish counts were done with an open species list, the absence of an observation is treated as a zero.

- In order to safely zero-fill the data we needed to be sure that a fish would have been counted if observed. For this reason we chose to exclude the following taxa, which were not treated consistently by the different sampling programs:
  - Order
    - Clupeiformes
    - Pleuronectiformes
    - Tetraodontiformes
    - Batrachoidiformes
    - Gobiesociformes
    - Ophidiiformes
    - Carcharhiniformes
  - Family
    - Carangidae
    - Chaenopsidae
    - Chaetodontidae
    - Pholidae
    - Scombridae
    - Atherinidae
    - Atherinopsidae
    - Clinidae
    - Gobiidae
    - Pholidae
    - Stichaeidae
    - Agonidae
    - Cottidae
      - Except for *Scorpaenichthys marmoratus* (cabezon)
    - Hemitripteridae
    - Hexagrammidae
      - Except for *Ophiodon elongatus* (lingcod)

        - Liparidae
        - Scorpaenidae
        - Labrisomidae
- A few taxa in these groups have been retained in the following special case:
    - Fish species in LTER crypticfish surveys are retained if either PISCO or KFM counts that species in quadrats (and thus comparable data are available) but NOT if SNI data are included (SNI does not survey these cryptic fish in a similar way)
- We removed KFM visualfish surveys (pre-1996 issue)
- We treated olive-yellowtail as olive rockfish & yellowtail rockfish as olive rockfish

QUAD SWATH

User choices

- ***Data Sources***: 0 = do not use this data source in the aggregation, 1 = use this data source in the aggregation
    - use_SNI = 1
    - use_KFM = 1
    - use_LTER = 1
    - use_PISCO = 1
- ***Taxa ratio cutoff***: This is the maximum fraction of observations within a taxonomic level that can be discarded in order to reach the next finer level of taxonomic resolution. Small numbers will aggregate more species at higher taxonomic levels, but keep more of the observed taxa. Larger numbers will discard more observations, but result in finer taxonomic resolution.
    - taxaRatioCutoff = 1
    - ** For all the quad swath file outputs, a taxa ratio cutoff of 1 was used, meaning the code went to the lowest possible taxonomic resolution as possible.

Manual decisions

- We treated *Megastraea* in the LTER swath as *M. undosa*
- Exclude quads from SNI (added in 2016)
- Filter out things known as generally incompatible (juveniles, open classes, incompatible methods):
    - Family
        - Muricidae
        - Calliostomatidae
    - Genus
        - *Cucumaria*
        - *Tegula*
        - *Pomaulax*
        - *Undaria*
    - Taxon name
        - *Sargassum horneri*
        - *Sargassum muticum*

- Remove *K. kelletti* from quads
- Remove open classes
- Remove juveniles/recruits
- Remove *Egregia* from quads
- Remove "no organism in this sample"
- Remove combined *Laminaria* and *Pterygophora* groups before 1996 (lumps juveniles and adults)
- Keep adult only *Macrocystis* adults
- Remove *Tethya, Crassadoma, Urticina* incompatibles
- Remove all PISCO *Styela montereyensis* because incompatible dates/methods
- Remove KFM band/1mquad for consistency *Pisaster giganteus*
- Remove KFM *Pycnopodia helianthoides* in 1mquad (probable mistake + inconsistent method)
- Remove duplicated LTER quad (two proj_taxon_id for same thing, NA in count)
- Remove STRPURAD_COVER percent cover count (*Strongylocentrotus purpuratus*)
- Remove all KFM *Lytechinus pictus* because inconsistent monitoring (changed sampling method)
- Remove PISCO quad *Haliotis corrugata* (possible mistake count)
- Remove KFM quad *Haliotis rufescens* for method compatiblity
- Remove KFM 1mquad & 5mquad (both juv & adult) *Sargassum horneri* for method compatibility
- Remove broad *Laminaria* spp. PISCO
- Remove broad *Laminaria* spp. KFM
- Remove KFM 1mquad *Macrocystis pyrifera* (pre-1996)
- Remove KFM 1mquad adult & juv *Eisenia arborea* (pre-1996),
- Remove PISCO quad for *Strongylocentrotus purpuratus* because PISCO quads ended 2011,
- Remove PISCO quad for *Stephanocystis osmundacea* because PISCO quads ended 2011,
- Remove PISCO quad for *Lytechinus pictus* because PISCO quads ended 2011
- Remove PISCO quad for *Neobernaya spadicea* because PISCO quads ended 2011
- Remove PISCO quad *for Mesocentrotus franciscanus* because PISCO quads ended 2011
- Remove KFM band for *Haliotis sorenseni* because not observed by other groups

- Filter out taxa if one group would not have recorded it if it was observed (not open species list)

BENTHIC (PERCENT COVER)

User choices

- ***Data Sources***: 0 = do not use this data source in the aggregation, 1 = use this data source in the aggregation
  - use_SNI = 1
  - use_KFM = 1
  - use_LTER = 1
  - use_PISCO = 1

Manual decisions

- Change genus *Gigartina* to *Chondracanthus* (switch in taxonomy) for SNI
- Add order Actiniaria for where it is missing in data and change auth_taxon_id = 1360
- Add order Alcyonacea to PISCO (missing in data) and change auth_taxon_id = 1365
- Add order Canalipalpata to *Phragmatopoma californica* (missing in data, causes aggregation issues)
- Remove SNI mollusca
- Percent cover was zero filled so that all taxa are included at all sampling dates and transects because it was open species list

For filtering, it was decided what taxonomic groups to keep based on which data sources were being included. Note: If a lower taxonomic level is included in a higher taxonomic level, it was pulled from the higher level to put a separate entity.

1. **If both PISCO and KFM are being included, taxa are aggregated at the following levels:**

Name	Included	Excluded
KINGDOM		
Animalia		*Corynactis californica, Phragmatopoma californica, Thylacodes squamigerus, Pachythyone rubra, Stephanocystis osmundacea*; anything else below, mollusks
PHYLUM		
Chlorophyta	*Bryopsis corticulans, Codium fragile, Codium setchellii, Derbesia marina, Cladophora graminea, Cladophora, Ulva, Chlorophyceae*	
Porifera		
Bryozoa		
Rhodophtya		Corallinaceae
CLASS		
Ascidiacea	*Clavelina huntsmani, Pycnoclavella stanleyi, Didemnum carnulentum, Diplosoma listerianum, Trididemnum opacum. Euherdmania claviformis, Cystodytes lobatus, Eudistoma molle, Eudistoma psammion, Aplidium, Polyclinum planum, Synoicum, Ritterella aequalisiphonis, Chelyosoma productum, Boltenia villosa, Metandrocarpa dura, Styela montereyensis*	
ORDER		
Scleractinia	*Paracyathus stearnsii, Balanophyllia (Balanophyllia) elegans, Astrangia haimei*	
FAMILY		
Corallinaceae	*Amphiroa beauvoisii, Bossiella orbigniana, Lithothrix aspergillum, Jania rosea, Corallina officinalis, Calliarthron cheilosporioides, Calliarthron, Lithothrix, Bossiella*	
GENUS		
Desmarestia	*Desmarestia ligulata*	
SPECIES		
Stephanocystis osmundacea		
Corynactis californica		
Phragmatopoma californica		
Thylacodes squamigerus		
Pachythyone rubra		

2. If KFM was not included, taxa are aggregated at the following levels:

Name	Included	Excluded
KINGDOM		
Animalia		*Corynactis californica, Stephanocystis osmundacea,* Mollusks; anything else below
PHYLUM		
Chlorophyta	*Bryopsis corticulans, Codium fragile, Codium setchellii, Derbesia marina, Cladophora graminea, Cladophora, Ulva, Chlorophyceae*	
Porifera		
Bryozoa		
Rhodophtya		Corallinaceae
CLASS		
Ascidiacea	*Clavelina huntsmani, Pycnoclavella stanleyi, Didemnum carnulentum, Diplosoma listerianum, Tridemnum opacum. Euherdmania claviformis, Cystodytes lobatus, Eudistoma molle, Eudistoma psammion, Aplidium, Polyclinum planum, Synoicum, Ritterella aequalisiphonis, Chelyosoma productum, Boltenia villosa, Metandrocarpa dura, Styela montereyensis*	
Polychaeta	*Diopatra ornata, Eudistylia polymorpha, Sabellidae, Spirobranchus spinosus, Serpulidae, Timarete luxuriosa, Pista elongata, Chaetopterus*	*Salmacina tribranchiata, Dodecaceria fewkesi, Phragmatopoma californica*
Hydrozoa	*Schuchertinia milleri, Stylantheca papillosa, Stylaster californicus, Obelia, Plumularia, Abietinaria, Sertularia, Symplectoscyphus turgidus*	*Aglaophenia*
Bivalvia	*Pectinidae, Crassadoma gigantea, Mytilus californianus, Pholadidae, Parapholas californica, Chaceia ovoidea*	
ORDER		
Scleractinia	*Paracyathus stearnsii, Balanophyllia (Balanophyllia) elegans, Astrangia haimei*	
Alcyonacea	*Discophyton rudyi. Leptogorgia chilensis, Muricea californica, Octocorallia*	
Dictyotales	*Dictyopteris undulata, Dictyota binghamiae, Dictyota flabellata, Dictyota coriacea, Taonia lennebackerae, Zonaria farlowii, Dictyotaceae*	
Dendrochirotida	*Cucumaria piperata, Cucumaria salma, Cucumaria, Lissothuria nutriens, Eupentacta quinquesemita*	*Pachythyone rubra*
FAMILY		
Corallinaceae	*Amphiroa beauvoisii, Bossiella orbigniana, Lithothrix aspergillum, Jania rosea, Corallina officinalis, Calliarthron cheilosporioides, Calliarthron, Lithothrix, Bossiella*	
Balanidae	*Balanus*	
Vermetidae	*Thylacodes squamigerus, Petaloconchus montereyensis*	
GENUS		
Desmarestia	*Desmarestia ligulata*	
Sargassum	*Sargassum horneri, Sargassum muticum*	
Aglaophenia	*Aglaophenia struthionides*	
SPECIES		
Stephanocystis osmundacea		
Corynactis californica		
Phragmatopoma californica		
Salmacina tribranchiata		
Dodecaceria fewkesi		
Pachythyone rubra		

3. If KFM is included and PISCO is not, taxa are aggregated at the following levels:

Name	Included	Excluded
KINGDOM		
Animalia		*Stephanocystis osmundacea, Astrangia haimei, Corynactis californica, Phragmatopoma californica, Balanophyllia (Balanophyllia) elegans, Diopatra ornata, Thylacodes squamigerus, Pachythyone rubra;* anything else below, mollusks
PHYLUM		
Chlorophyta	*Bryopsis corticulans, Codium fragile, Codium setchellii, Derbesia marina, Cladophora graminea, Cladophora, Ulva, Chlorophyceae*	
Porifera		
Bryozoa		
Rhodophtya		genus *Chondracanthus,* genus *Gelidium,* family Corallinacae
CLASS		
Ascidiacea	*Clavelina huntsmani, Pycnoclavella stanleyi, Didemnum carnulentum, Diplosoma listerianum, Trididemnum opacum. Euherdmania claviformis, Cystodytes lobatus, Eudistoma molle, Eudistoma psammion, Aplidium, Polyclinum planum, Synoicum, Ritterella aequalisiphonis, Chelyosoma productum, Boltenia villosa, Metandrocarpa dura, Styela montereyensis*	
FAMILY		
Corallinaceae	*Amphiroa beauvoisii, Bossiella orbigniana, Lithothrix aspergillum, Jania rosea, Corallina officinalis, Calliarthron cheilosporioides, Calliarthron, Lithothrix, Bossiella*	
GENUS		
Desmarestia	*Desmarestia ligulata*	
Gelidium	*Gelidium robustum*	
Chondracanthus	*Chondracanthus corymbiferus, Chondracanthus exasperatus*	
SPECIES		
Stephanocystis osmundacea		
Astrangia haimei		
Corynactis californica		
Phragmatopoma californica		
Balanophyllia (Balanophyllia) elegans		
Diopatra ornata		
Thylacodes squamigerus		
Pachythyone rubra		

NOTES ON ENVIROMENTAL DATA SETS

SST Data

Information regarding SST data can be found and downloaded at:
https://podaac.jpl.nasa.gov/dataset/MUR-JPL-L4-GLOB-v4.1

SST data were largely compiled in 2016 using the Marine Geospatial Ecology Toolbox in ArcGIS (http://mgel.env.duke.edu/mget/):

SST Fish

- subsites: b-p-025, b-p-026, b-p-027, b-p-048, b-p-084, b-p-104, b-p-107, b-p-108 - no data in SST file
- subsites: b-p-050 - missing data for 2002, 2003, 2004
- 2017 data not complete

SST Quad Swath

- subsites: b-p-047, b-p-107, b-p-108, b-p-138, b-p-139 - no data in SST file
- subsites: b-p-050 - missing data for 2002, 2003, 2004
- 2017 data not complete

SST Benthic

- subsites: b-p-047, b-p-138, b-p-139 - no data in SST file
- subsites: b-p-050 - missing data for 2002, 2003, 2004
- 2017 data not complete

Wave Height (mean & max)

- Data from 2003 – 2015 only
- These data were compiled by Tom Bell, and represent the daily significant mean/ median and maximum wave height (m) at 200-m resolution.
- Data were downloaded at: https://cdip.ucsd.edu/m/documents/data_access.html#data-use-and-acknowledgements

Kelp Canopy Biomass

- Data from 2004 – 2014 only
- Lots of months with NA recorded in raw biomass data set
  This is a monthly kelp patch biomass data set compiled by Tom Bell. Both Landsat 5 and 7 were used in this data set. Methods described in Bell et al. *in press*
- The SBC LTER data set was downloaded at : http://sbc.lternet.edu/cgi-bin/showDataset.cgi?docid=knb-lter-sbc.74

Chlorophyll (point, 3km buffer, 5km buffer)

- Data from 2004 – 2016 only
- Many months have entries with NA in raw chlorophyll data set

- Data were downloaded here: http://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMWchlamday.graph?chlorophyll[(2010-12-16T12:00:00Z)][(0.0)][(31.0):(35.0)][(238.6):(243.0)]&.draw=surface&.vars=longitude|latitude|chlorophyll&.colorBar=Spectrum|C|Log|||
- Data are monthly composites from 31°- 35° N and 238.6° - 243° E

Turbidity (point, 1km)

- Data from 2003 – 2016 only
- Many months have entries with NA in raw turbidity data set (not true for 1km data set)
- The monthly diffuse attenuation coefficient data were downloaded here : https://coastwatch.pfeg.noaa.gov/erddap/griddap/erdMWk490mday.html
- NASA distributes Diffuse Attenuation Coefficient at 490 nm Wavelength data from NASA's Aqua satellite. Measurements are gathered by the Moderate Resolution Imaging Spectroradiometer (MODIS) carried aboard the spacecraft.

CDIP Wave Data (Hs, Orbital Velocity, Wave Energy)

- Sites without hindcasts: a-k-17, a-k-18, a-k-19, a-k-20, a-k-23, a-k-26, a-k-27, a-k-29, a-k-33, a-p-03, a-p-05, a-p-11, a-p-12, a-p-13, a-p-16, a-p-23, a-p-31, a-p-36, a-p-49, a-p-51, a-p-54, a-s-01
- Paper describing Hs and bottom orbital velocity functions attached in Metadata folder (Wiberg and Sherwood 2008)
- Paper describing wave energy functions attached in Metadata folder (Williams et al. 2013)
- The original hourly hindcasts were generated by Chris Gotschalk and Tom Bell using the CDIP MOP model (see above). Kevin Lafferty generated the correction coefficients. Methods used to create and apply these coefficients are described in Lafferty et al. (2018).
- For most sites, the coefficients for Hs were simply multiplied by the CDIP model predictions for those sites. There were a few sites where the coefficients were applied in slightly different ways. Those sites are described in the report.
- The hindcasting procedure is explained in Bell et al. 2015

Literature cited

Bell TW, Cavanaugh KC, Reed DC, Siegel DA. 2015. Geographical variability in the controls of giant kelp biomass dynamics. Journal of Biogeography 42(10):2010–2021.

Lafferty KD, Morton DN, Gotschalk CC, Henderikx F, Rassweiler A, Washburn L. 2018. Hourly wave height and period hindcasts at 32 sites throughout the Channel Islands National Park and San Nicolas Island from 2000-2017: U.S. Geological Survey data release, https://doi.org/10.5066/P91KH2Q7.

Wiberg PL, Sherwood CR. 2008. Calculating wave-generated bottom orbital velocities from surface-wave parameters. Computers and Geosciences 34(10):1243–1262.

Williams GJ, Smith JE, Conklin EJ, Gove JM, Sala E, Sandin SA. 2013. Benthic communities at two remote Pacific coral reefs: Effects of reef habitat, depth, and wave energy gradients on spatial patterns. PeerJ 2013(1):1–26.

Appendix D

Detecting human impacts in a highly variable ecosystem using long-term monitoring data

ABSTRACT

Evaluating the effects of unexpected human impacts is a major challenge for applied ecology. When ecological monitoring data are available, a Before-After-Control-Impact (BACI) analysis is often applied, which can control for natural spatial and temporal variation to better isolate an impact. The virtues of BACI-type designs are well established, but the degree to which the approach may be compromised by patchy population distributions and dynamics has not been systematically explored in nearshore marine systems. Here we quantify the potential for BACI analyses of long term monitoring data to detect anthropogenic impacts with the goal of better understanding its reliability when applied to particular species (or sets of species). We start with a spatially and temporally extensive monitoring data set from reefs in southern California and simulate impacts of different spatial-scales and severities. We find the BACI approach had substantial potential to detect local reductions in the population sizes of 28 species of fish, invertebrate and macroalgae, as long as temporal autocorrelation is accounted for in the statistical model (neglecting this temporal structure lead to very high false positive rates). However, the power to detect impacts varies widely from species to species and false positives exceed target levels in some cases because of spatial synchrony in dynamics (if impact sites are scattered across the region, rather than clustered, false positives rates are lower). Impacts are most likely to be detected in species with relatively stable, widely distributed populations, but the same factors increase false positive rates. Finally, we found that combing even a small number of impacted species in our analyses greatly improved the method's success rates, but that the choice of species to be combined was crucial. These results provide guidance to long-term monitoring projects, both suggesting characteristics of species that would serve as good indicators of an impact and setting expectations for what sizes and severities of impacts are likely to be detectable.

INTRODUCTION

Rigorously quantifying the ways in which human activities affect natural systems is a central goal of applied ecology. One of the most challenging contexts for such quantification occurs when a human impact is unreplicated, as they often are. Such impacts can result from accidents such as oil spills (Foster et al. 1971) or can be side-effects of planned activates such as waste discharge (Roberts et al. 2010), fishing and hunting (Gray et al. 2006, Jupiter et al. 2012) or construction and removal of major infrastructure (Pen et al. 2006, Kibler et al. 2011, Rytwinski et al. 2017). Detecting the ecological consequences of human impacts is important for a range of political, legal, and scientific reasons.

In many cases, impacts occur at locations where baseline data are unavailable, but they are much easier to evaluate at locations with established monitoring programs. Indeed many monitoring programs are at least in part intended to provide baseline data for evaluating anthropogenic impacts (Field et al. 2007, Fancy et al. 2009). When monitoring data are available from the region where an impact occurs, a Before-After-Control-Impact (BACI) design is often used to test for effects on the ecological system. BACI-type statistical designs have evolved considerably since their introduction (Stewart-Oaten et al. 1986,

Underwood 1992, Schmitt and Osenberg 1996), but the general approach is to use contrasts between control and impact sites prior to an event to control for spatial variation, and contrasts between pre-impact and post-impact time periods at the control sites to account for natural temporal variation. For example, BACI designs are commonly used to estimate the effects of oils spills such as the 2010 Deepwater Horizon spill (Dietl and Durham 2016, Lauritsen et al. 2017).

There are numerous examples of the advantages of BACI-type designs (Osenberg and Schmitt 1996), and these analyses are common and have been increasing over the past two and a half decades since the term was introduced. The BACI approach has known weaknesses as well however (Stewart-Oaten et al. 1992, Underwood 1992, Stewart-Oaten and Bence 2001, Parker and Wiens 2005, Paul 2011). At the most basic level, to evaluate an impact on a species using BACI requires data on the abundance of that species both inside and outside the hypothesized impact area, and such locations might be difficult to predict in advance. A more subtle problem is the challenge of distinguishing between natural fluctuations in populations and human impacts. When an impact is unreplicated, a BACI analysis may detect a statistically significant change at the impacted sites in the post-impact time period due to spatial variability in temporal dynamics rather than any human impact. While the power of BACI designs are routinely evaluated using traditional power analysis (based on expected variance and effect sizes; (Carey and Keough 2002)), the prevalence of confounding spatial dynamics has not been systematically explored in natural systems.

The potential for confounds of BACI designs is particularly acute in marine systems because the availability of ecological data is often limited and the systems are subject to extreme natural variability on all spatial-scales (Reed et al. 2016). On reefs in Southern California, for example, ecosystem state varies dramatically from location to location and over time, and the combination of strong environmental gradients, stochastic year to year recruitment dynamics and decadal scale environmental forcing can cause different areas within a region to exhibit contrasting dynamics. Nevertheless BACI approaches have commonly been applied in California's marine systems (Schroeter et al. 2001a, Parnell et al. 2005, Martin et al. 2012) and substantial investment in ecological monitoring is being made within this region by federal, state and private institutions. Understanding the value of these data for assessing impacts is key to decisions about continued investment in such monitoring (Legg and Nagy 2006).

Here we apply BACI analyses to monitoring data from subtidal reefs in Southern California, evaluating the likelihood that these data would be able to detect a spatially-localized anthropogenic impact such as an oil spill as a function of the spatial scale and severity of the impact. We evaluate a synthesized data set generated by combining monitoring data collected by the U.S. National Park Service and the U.S. Geological Survey. These data are large in spatial, temporal and taxonomic scope and they encompass the extreme variability characteristic of shallow reef ecosystems. As such, the data provide an exceptional opportunity for evaluating the strengths and weaknesses of a BACI approach within such a context.

METHODS

To examine how nearshore species vary in their value for estimating and detecting impacts, we conducted numerical analyses using observational surveys for a broad spectrum of nearshore marine taxa. Specifically, we use these spatially replicated data to ask how impact area and impact severity interact with species characteristics and model assumptions to determine our ability to detect impacts (or lack thereof) and estimate impact severity. We estimate an "informedness" score that measures capacity to both accurately detect real impacts and avoid false positives when impacts are absent. In addition to species by species analysis, we grouped species using various strategies to evaluate if analyzing clusters of species improves or diminishes informedness.

Study region and data

This analysis focuses on nearshore waters in Southern California, which has a population of more than 18 million people, much of it concentrated on the coastline, and has near-shore ecosystems that are subject to wide range of human impacts including fishing, pollution and coastal development. In spite of these threats, near-shore reefs support productive and diverse ecological communities, and are the focus of intensive management and conservation attention. Several programs measure abundances of fish, invertebrates, and algae at sites throughout the region. Two of the longest running monitoring programs are administered by the Channel Islands National Park (CINP) and the U.S. Geological Survey (USGS). The CINP has been monitoring 16 sites at San Miguel, Santa Rosa, Santa Cruz, Anacapa, and Santa Barbara Island annually for more than 30 years. The USGS has annually monitored seven sites at San Nicolas Island over a similar time. Although there are some differences between the program's methods, both count mobile invertebrates, fish, and kelps within defined areas, and census sessile invertebrates and smaller algae using point contact methods (full details on sampling methods are given in (Kenner et al. 2013, Kushner et al. 2013).

We obtained a broad study area around California's Northern Channel Islands by combining 23 sites from the CINP and USGS monitoring programs. Combining programs meant focusing on 28 species that have been sampled in compatible ways for the whole length of both programs. These include 5 algal species, 10 fish species, 7 mobile invertebrate species, and 6 sessile invertebrate species. CINP samples once a year (in summer), while the USGS monitoring program samples twice a year, so only the autumn samples from the USGS data were used in this analysis to maintain compatibility in temporal resolution between the two data sets. For each site and sampling year, we aggregated the data for each species to produce a total count (individuals or points observed) along with the total area or number of points sampled.

Simulating human impacts

Our general approach for evaluating the utility of these data in a BACI-type analysis was to start with the actual data, simulate a hypothetical human impact by reducing species abundance over a period of time at "impacted" sites, and then analyze this altered data set to determine whether the reduction in abundance was statistically detectable and how this likelihood varied among species for a range of impact sizes and severities.

More specifically, for each simulation, we explored the effect of impact size by repeatedly simulating circular impacts of various sizes, choosing a random location for the center of each impact between a latitude of 33.2 and 34.1 degrees north and a longitude between 119.0 and 120.4 degrees west, a domain that included all sites in the data set. Based on the chosen center location and the radius of impact, we determined which monitoring sites were affected. We then defined the timing of an impact by selecting a year at random between 1996 and 2006. The 10 years before the selected impact year were used as pre-impact data and the five years after were used as our post-impact period.

For each simulated impact, we constructed a new data set in which the actual observed abundances of each species were used for all sites in the pre-impact period and the control sites during the post-impact period, but data from the impacted sites in the "post" period were replaced with altered abundances. The degree to which impacted abundances were altered was based on effect *severity*, which was the average fraction by which the abundance of each species within the affected area would be reduced (so a *severity* of 0 meant no effect was simulated, while a *severity* of -100% meant the species were entirely removed from the affected sites). Within a given simulation, we assumed all monitoring sites within the chosen radius of the center of the event were similarly affected over the course of the post-impact period (so effects were spatially uniform and persisted over time). Because the data are counts, simulated severity

was applied probabilistically by rounding to the nearest integer with probability equal to the remaining decimal (e.g. 10.25 is rounded to 10 with probability 0.25 and 11 with probability 0.75). This probabilistic rounding was done to maintain the integer nature of the data while maintaining overall mean impact severity.

To evaluate how individual or collective analysis affected power and false positives, we applied the impacts and analyses to a) each individual species separately, b) all species of the same functional group (algae, fish, mobile invertebrates, sessile invertebrates), c) a series of species subsets generated by sequentially adding species ranked in informedness from highest to lowest and lowest to highest (based on single species analysis) while accounting for species identity. For (c) we combined only species with counts (not point contact) in order to maintain a consistent response variable and likelihood. For each impact severity and each species/functional group, we ran 1000 simulations and summarized statistics across runs.

Statistical models

We used generalized linear mixed effects models for all analysis with a before/after (BA) effect, a control/impact (CI) effect, and an interaction between BA and CI terms representing the BACI effect. For count data, we used a Poisson likelihood and for UPC data we used a binomial likelihood. We attempted to the keep the models simple to minimize convergence issues and model fitting errors in the simulation algorithm. We avoided both the negative binomial and beta-binomial, which also account for additional dispersion, because these implementations converged less reliably over many randomization runs. Instead, to account for both over dispersion and zero-inflation, we included a random lognormal error dispersion term in the model. We also excluded sites from our analysis if the focal species was present at fewer than 15% of sampling events, to reduce zero-inflation. To account for among-site variation, we added random effects for site in the single species models, and site within species for the multi-species models. For the single species models we either tried two approaches, a) accounting for serial correlations using a first order autoregressive model on the error terms or b) ignoring within site serial correlation. Except where explicitly stated, we report only results including an AR(1). We estimated statistical models using glmmTMB (Magnusson et al. 2016) in R (R Development Core Team and R Core Team 2017).

Performance metrics

Performance metrics for each run include the following. 1) one-sided significance tests (negative and positive) of the estimated BACI effect. 2) whether or not the model converged (represents when a model failed despite having sufficient data for analysis). 3) data sufficiency, in this case requiring at least two sites in each of the BA and CI categories with at least 15% of the years with the species present in the original data. This threshold was chosen for efficiency as data below this threshold consistently provide unreliable model estimates. 4) relative error from the true effect measured as (estimated impact – true impact)/true impact (only calculated when we simulated a non-zero impact). 5) Finally we used incidence of false positives and power to calculate the Youden's J informedness statistic. Functionally, this metric measures the probability of a correct inference, assuming an even prior probability of no impact versus a true impact.

Additional statistics

We calculated several statistics for each simulation run that we expected might be associated with power and false positive rates. These included: (1) the coefficient of variation; (2) average serial correlation; (3) number of sites with sufficient data for analysis; (4) spatial synchrony among sites; and (5) mean number of years where the species of interest is present at the given sites. The coefficient of variation was

calculated as the within-site standard deviation divided by the within-site mean, averaged across all sites for the 10 years before the impact. Serial correlation was calculated as the within-site empirical partial autocorrelation function, averaged across all sites. Spatial synchrony among sites was estimated as the variance of the mean time series divided by the sum of the covariance matrix of the group of individual time series (Loreau and de Mazancourt 2008). We compared whether these metrics could explain power and probability of false detection using generalized linear models with a beta-binomial likelihood in a multiple regression framework. We centered and standardized all variables, including a log-transformation for those variables constrained by zero (coefficient of variation, number of sites with sufficient data, synchrony, and number of yeas with sufficient data).

RESULTS

Impact severity and area

The likelihood of detecting an impact with a BACI design increased with severity, and had a hump-shaped relationship with impact size - illustrated for black surfperch (*Embiotoca jacksoni*) and giant kelp (*Macrocystis pyrifera*)(Figure 2a, b). Intermediate sized impacts were more likely to be detected than either large or small impacts due mostly to issues of data sufficiency; if an impact was very small there were unlikely to be any monitoring sites within its radius, while if an impact was large there might be few or no control sites available (Figure A1). Whereas the general shape of this pattern was consistent across most species analyzed, the peak's height varied from species to species (Figure A1).

In contrast to detection probability, error in estimated impact severity decreased monotonically with impact area (Figure 2c & d). As the number of impact sites increased, relative error in the estimated impact declined even as the number of control sites diminished to zero (Figure 2c & d). In such cases without control sites, the change over time could still be estimated based on comparison with densities in the pre-impact period, but the degree to which this represented change due to the impact vs natural temporal trends could not be evaluated.

Among species variation in power and false positives

We examined among-species variation in power and the rate of false positives, holding the spatial scale of an impact constant. When one-quarter of sites were affected by the impact (5 or 6 sites), a severe impact was almost sure to be detected for some species, but impossible to detect for others (Figure 3a-d). Likewise, false impact detection was more prevalent than we expected. In scenarios with no simulated impact (zero severity, Figure 3e-f) only 11 out of 28 species had <5% probability of false detection and only 23 out of 28 had a <10% probability of false detection.

These high rates of false detections were likely due to violations of the assumptions of our simple statistical models. Standard BACI models assume no serial autocorrelation, no spatial autocorrelation, and homogeneity of variance among sites and through time. Of these confounding factors, our results in Figure 3 only accounted for within site serial autocorrelation (using an AR(1) error model). Accounting for this variation is key as it reduced false detection rates relative to a model that did not account for such autocorrelation, with the reduction dependent on the degree of serial correlation in the time series (Figure 4). While including an AR(1) model can reduce the incidence of false detections, they linger at elevated probabilities when serial correlation is high.

While our models accounted for temporal autocorrelation, it is considerably more complicated to adjust for patterns of spatial autocorrelation (in which clusters of sites behave more similarly in time). Our monitoring sites are sparsely and irregularly distributed, with large gaps between sites, and spatial structure in these populations is driven by complex environmental, biological and anthropogenic factors

that are often not well represented by geographic distance (Watson et al. 2011, Lamy et al. 2018). Results shown in Figure 3-7 do not attempt to account for spatial autocorrelation, but we can show its importance by removing spatial structure of impacts. When impacts were randomized in space, so that distant sites were as likely to be co-affected as nearby ones, false detections were further reduced (Figure A2). In this case, 15 rather than 11 out of 28 species have less than 5% probability of false positives. Thus, violations of model assumptions (many of which we do not explore or include here for simplicity – e.g. heteroscedastic error variance, nonstationary trends, higher order spatiotemporal correlation structures, observation error, immigration and emigration, etc.) can dramatically impact false detection probabilities.

Power to detect even strong impacts was variable in part because some species exhibited enormous year-to-year variation in abundance. Results of our post-hoc regressions on both power and false detections show that power diminishes as a function of mean variability in the time series, data sufficiency, serial correlation (even after applying an AR(1) error model) and potentially spatial synchrony among sites (Table 1, Figure 5). Unfortunately, the same properties that provided high power also provided high probabilities of false detections, as power shows a weak positive correlation with probability of false detection (Figure 6). However, some species yielded low false positives but only modest power (e.g. *Macrocystis*), whereas others exhibited similarly low false positives but high power (e.g. black surfperch).

Value of combined species inference

Fortunately, in cases when groups of species are expected to be similarly impacted, power can be increased by estimating BACI models in a nested framework. When species within a taxonomic group were combined, it resulted in dramatically increased power in the nested BACI analysis relative to the power obtained when analyzed separately (Figure 3 indicted by “all algae”, “all sessile inverts”, “all mobile inverts”, “all fish”). For three of the four groups, false detections in the combined analyses resembled the mean independent false detection rate (a regression to the mean); the exception was algae, where only three species were pooled because of differences in measurement types (point contact versus counts). For algae, combined analysis actually increased probability of false detections while increasing power (Figure 3).

While combining inference can benefit power and false positives, the individual species informedness has a large impact on clustered informedness, and thus the decision to undertake a multispecies analysis should be taken with care. When adding species in order of individual informedness, grouped informedness rapidly saturates with the 5-6 most informed species, and even eventually begins to erode when adding species with lower informedness (Figure 7 - blue points). In contrast, when adding species from lowest to highest informedness to the grouped analysis, saturation is exceedingly slow (Figure 7 – red points). Thus, although grouping species for analysis is promising, the approaches efficiency of depends largely on composition of species used in the analysis.

DISCUSSION

Our analyses of the effectiveness of BACI in marine systems provides guidance on the kinds of impacts that are likely to be successfully detected. As might be expected, we found that more severe impacts are more detectable, but that the relationship between impact size and detectability was more complex, and depends on the spatial extent of monitoring locations and their density. Impacts that were small relative to the average spacing between monitoring sites were unlikely to be detected, as there is unlikely to be data from the impacted area. By contrast, impacts that approach the spatial scale of the whole monitoring program often leave few or no unimpacted control sites that could be used in a BACI analysis.

Our results emphasize that impacts on some species may be difficult to detect even with an exceptionally extensive monitoring program. The macroalgae in this data set provide a striking example of this; a severe impact would be detected less than half the time for any of the five species examined. The BACI approach worked better on average for the other taxonomic groups, but with exceptions in each case.

Most of the failures of this approach occurred either because the species was highly variable or because it was not found at many of the sites being monitored. For example, the five species with the lowest power *(Phragmatopoma, Eisenia, Desmarestia, Pterygophora,* and olive rockfish) all had very high temporal coefficients of variation, with standard deviations well larger than the mean in each case) and limited distributions, being regularly observed at less than 77% of the sites. These observations can guide the selection of species to be included in marine monitoring programs and help set expectations for the power to detect impacts. For example, the coefficient of variation of species abundance and the number of sites at which a species is present can be calculated before an impact occurs, potentially flagging species which would serve as poor indicators of an impact. Spatial and temporal synchrony can also be calculated, although they serve as less reliable indicators of power. Although we found limited power to detect impacts on some species some species, BACI has been used successfully with these data sets, even for single species (Schroeter et al. 2001b, Parnell et al. 2005).

The simulations performed here are at root a power analysis, but one performed using actual data rather than one based solely on expected variation, effect size, and sample size. This data-oriented approach allows us to grapple with the variability and spatio-temporal structure of real ecological communities. Still, the simulation approach contains several assumptions that may well be violated in many scenarios of human impact. First, we assume that the spatial scale of the impact is known, so the set of monitoring sites can be unerringly classified into control and impact groups. Second, our simulated impacts persist at a constant level for a known length of time (species were reduced by a fixed fraction for five years). Both of these characteristics were well matched to the relatively simple statistical approach we implemented; in cases where there is considerable uncertainty about the spatial scale, duration, or temporal course of an impact, more sophisticated methods would be necessary.

Our results emphasize the importance of accounting for serial correlation in analyses of impacts in marine systems. We found that false positive rates more than doubled if the first order autoregressive term was not included in the model. This inflation in false positive rates was particularly apparent for localized spatial impacts (the main focus of this analysis) but also occurred when impacts are scattered across the region. False positives were common even when first order autocorrelation was accounted for. For example, we detected a significant positive or negative impact more than 20% of the time for most invertebrate species examined, even when no impact had been simulated into the data. False positives likely derived in part from our using a universal statistical model for a variety of taxa. For example, in each case we assumed a first order autoregressive model with a common parameter among sites. This is a robust and commonly used approach but as does not always eliminate serial autocorrelation because of higher order correlations, variance among sites in the degree of autocorrelation, or heteroskedastic error variances.

In many cases, however, false positives do not represent statistical failures but instead are the detection of population trends that differ across the study region. An underlying assumption of the BACI design is that in the absence of an impact, species would be changing in a similar way at both control and impact sites (Stewart-Oaten et al. 1986, Underwood 1992), but species dynamics are likely spatially correlated within the region and thus may yield a significant result if a species' trend within a set of putative impact sites differs from its trend elsewhere in the region. False positives due to spatial autocorrelation may be reduced by incorporating underlying environmental drivers as covariates (Parker and Wiens 2005, Kalies et al. 2010) or with analyses that explicitly account for spatial correlation structure. We were not able to

account for spatial autocorrelation in the underlying ecosystem dynamics because of the sparse and an irregular distribution of sampling, but the effects of spatial autocorrelation are apparent in the contrast between our analyses of spatially aggregated impacts and scattered impacts.

The power of BACI analyses is generally improved if impacts affect multiple species simultaneously. Although many BACI designs focus on detecting species-by-species impacts, some analyze changes in the whole community (Clarke et al. 2006, Pitcher et al. 2009). We found that a severe impact that locally reduced the abundance of all fish, sessile or mobile invertebrates would be virtually certain to be detected if all species in the class were included in the analysis. For these three groups false power rates are still elevated but not above the average of the constituent individual species. Analyzing macroalgae as a group was less promising; the grouped analysis yielded improved power but much higher false positive rates as well. Aggregating species is not a silver bullet however. Our analysis of the level of informedness (a statistic that accounts for both power and false positives) shows that combining species which are individually poor indicators of an impact does not improve informedness markedly until more than a quarter of the species are included.

Monitoring is and will continue to be a fundamental element of basic and applied ecological studies. In applied assessments of human impacts statistical significance is especially important because of litigation damage compensation and mitigation. Using a comprehensive long-term data set from Southern California reefs we show that the ability to detect such impacts not only varies with the magnitude of the impact but also depends on the spatial scale of the impact and the species affected. We found that a simple relatively standard BACI-type analysis would be likely to detect a severe impact of medium-size in the region, particularly if multiple species were involved. However, false positive rates were elevated over target levels, and would be unacceptably high if serial correlation was not accounted for. The same general findings will likely apply to the statistical assessment of environmental impacts from marine monitoring data more broadly and we urge that similar analyses be undertaken at the beginning of any monitoring program to increase the likelihood of data being suitable for the program's desired goals.

ACKNOWLEDGMENTS

Funding for fieldwork and data collection was provided by the United States National Park Service, US Geological Survey, US Fish and Wildlife Service and University of California Santa Cruz. Analyses were supported with funds from the U.S. Department of the Interior, Bureau of Ocean Energy Management, Environmental Studies Program, Washington, DC under Cooperative Agreement Number M11AC00012. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of the U.S. Government. Mention of trade names or commercial products does not constitute their endorsement by the U.S. Government.

REFERENCES

Carey JM, Keough MJ. 2002. The variability of estimates of variance, and its effect on power analysis in monitoring design. Environmental Monitoring and Assessment 74:225–241.

Clarke KR, Somerfield PJ, Airoldi L, Warwick RM. 2006. Exploring interactions by second-stage community analyses. Journal of Experimental Marine Biology and Ecology 338:179–192.

Dietl GP, Durham SR. 2016. Geohistorical records indicate no impact of the Deepwater Horizon oil spill on oyster body size. Royal Society Open Science 3.

Fancy SG, Gross JE, Carter SL. 2009. Monitoring the condition of natural resources in US national parks. Environmental Monitoring and Assessment 151:161–74.

Field SA, O'Connor PJ, Tyre AJ, Possingham HP. 2007. Making monitoring meaningful. Austral Ecology 32:485–491.

Foster M, Neushul M, Zingmark R. 1971. The Santa Barbara oil spill Part 2: Initial effects on intertidal and kelp bed organisms. Environmental Pollution (1970) 2:115–134.

Gray JS, Dayton P, Thrush S, Kaiser MJ. 2006. On effects of trawling, benthos and sampling design. Marine Pollution Bulletin 52:840–843.

Jupiter SD, Weeks R, Jenkins AP, Egli DP, Cakacaka A. 2012. Effects of a single intensive harvest event on fish populations inside a customary marine closure. Coral Reefs 31:321–334.

Kalies EL, Chambers CL, Covington WW. 2010. Wildlife responses to thinning and burning treatments in southwestern conifer forests: A meta-analysis. Forest Ecology and Management 259:333–342.

Kenner MC, Estes JA, Tinker MT, Bodkin JL, Cowen RK, Harrold C, Hatfield BB, Novak M, Rassweiler A, Reed DC. 2013. A multi-decade time series of kelp forest community structure at San Nicolas Island, California (USA). Ecology 94:2654–2654.

Kibler KM, Tullos DD, Kondolf GM. 2011. Learning from dam removal monitoring: Challenges to selecting experimental design and establishing significance of outcomes. River Research and Applications 27:967–975.

Kushner D J, Rassweiler A, Mclaughlin JP, Lafferty KD. 2013. A multi-decade time series of kelp forest community structure at the California Channel Islands. Ecology 94:2655.

Lamy T, Reed DC, Rassweiler A, Siegel DA, Kui L, Bell TW, Simons RD, Miller RJ. 2018. Scale-specific drivers of kelp forest communities. Oecologia 186:217–233.

Lauritsen AM, Dixon PM, Cacela D, Brost B, Hardy R, MacPherson SL, Meylan A, Wallace BP, Witherington B. 2017. Impact of the Deepwater Horizon oil spill on loggerhead turtle *Caretta caretta* nest densities in northwest Florida. Endangered Species Research 33:83–93.

Legg CJ, Nagy L. 2006. Why most conservation monitoring is, but need not be, a waste of time. Journal of Environmental Management 78:194–9.

Loreau M, de Mazancourt C. 2008. Species synchrony and its drivers: Neutral and nonneutral community dynamics in fluctuating environments. The American Naturalist 172:E48–E66.

Magnusson A, Skaug HA, Nielsen C, Berg K, Kristensen M, Maechler K, Bentham V, Bolker B, Brooks M. 2016. glmmTMB: Generalized Linear Mixed Models using Template Model Builder. R package version 0.0 2.

Martin C, Allen B, Lowe C. 2012. Environmental impact assessment: Detecting changes in fish community structure in response to disturbance with an asymmetric multivariate BACI sampling design. Bulletin Southern California Academy of Sciences 111(2):119-132.

Osenberg CW, Schmitt RJ. 1996. Detecting Ecological Impacts Caused by Human Activities. Pages 3–16 *in* R. J. Schmitt and C. W. Osenberg, editors. Detecting Ecological Impacts: Concepts and Applications in Coastal Habitats. Academic Press.

Parker KR, Wiens JA. 2005. Assessing recovery following environmental accidents: environmental variation, ecological assumptions, and strategies. Ecological Applications 15:2037–2051.

Parnell P, Lennert-Cody C, Geelen L, Stanley L, Dayton P. 2005. Effectiveness of a small marine reserve in southern California. Marine Ecology Progress Series 296:39–52.

Paul WL. 2011. A causal modelling approach to spatial and temporal confounding in environmental impact studies. Environmetrics 22:626–638.

Pen O, Carstensen J, Henriksen OD, Teilmann J. 2006. Impacts of offshore wind farm construction on harbour porpoises: acoustic monitoring of echolocation activity using porpoise detectors (T-PODs). Marine Ecology Progress Series 321:295–308.

Pitcher CR, Burridge CY, Wassenberg TJ, Hill BJ, Poiner IR. 2009. A large scale BACI experiment to test the effects of prawn trawling on seabed biota in a closed area of the Great Barrier Reef Marine Park, Australia. Fisheries Research 99:168–183.

R Development Core Team, R., and R Core Team. 2017. R: A language and environment for statistical computing. R: A language and environment for statistical computing.

Reed DC, Rassweiler AR, Miller RJ, Page HM, Holbrook SJ. 2016. The value of a broad temporal and spatial perspective in understanding dynamics of kelp forest ecosystems. Marine and Freshwater Research 67:14–24.

Roberts DA, Johnston EL, Knott NA. 2010. Impacts of desalination plant discharges on the marine environment: A critical review of published studies. Water Research 44:5117–28.

Rytwinski T, Algera DA, Taylor JJ, Smokorowski KE, Bennett JR, Harrison PM, Cooke SJ. 2017. What are the consequences of fish entrainment and impingement associated with hydroelectric dams on fish productivity? A systematic review protocol. Environmental Evidence 6:1–9.

Schmitt RJ, Osenberg CW (eds). 1996. Detecting Ecological Impacts: Concepts and Applications in Coastal Habitats. Elsevier.

Schroeter SC, Reed DC, Kushner DJ, Estes JA, Ono DS. 2001. The use of marine reserves in evaluating the dive fishery for the warty sea cucumber (*Parastichopus parvimensis*) in California, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 58:1773–1781.

Stewart-Oaten AA, Bence JR, Osenberg CW. 1992. Assessing effects of unreplicated perturbations: No simple solutions. Ecology 73:1396.

Stewart-Oaten A, Bence JR. 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs 71:305–339.

Stewart-Oaten A, Murdoch WW, Parker KR. 1986. Environmental impact assessment: "Pseudoreplication" in time? Ecology 67:929-940.

Underwood AJ. 1992. Beyond BACI: the detection of environmental impacts on populations in the real, but variable, world. Journal of Experimental Marine Biology and Ecology 161:145–178.

Watson JR, Hays CG, Raimondi PT, Mitarai S, Dong C, McWilliams JC, Blanchette CA, Caselle JE, Siegel DA. 2011. Currents connecting communities: nearshore community similarity and ocean circulation. Ecology 92:1193–1200.

	Power		False Detections	
Source	**Estimate**	**95% CI**	**Estimate**	**95% CI**
Coefficient of Variation	-0.71	[-1.01 : -0.42]	-0.24	[-0.49 : 0.01]
Number of Sufficient Sites	0.99	[0.68 : 1.31]	0.37	[0.06 : 0.69]
Spatial Synchrony	0.15	[-0.13 : 0.43]	-0.18	[-0.45 : 0.08]
Serial Correlation	-0.51	[-0.85 : -0.17]	0.28	[-0.02 : 0.59]
Number of Years	0.12	[-0.15 : 0.39]	-0.08	[-0.40 : 0.24]

Table 1: Standardized coefficients and 95% confidence intervals from the beta-binomial GLM. Models for both power and false detections included all shown variables. For interpretation, the coefficients represent the rate of change of the response to a single standard deviation change in the predictor. Thus, all coefficients within a model are directly comparable.

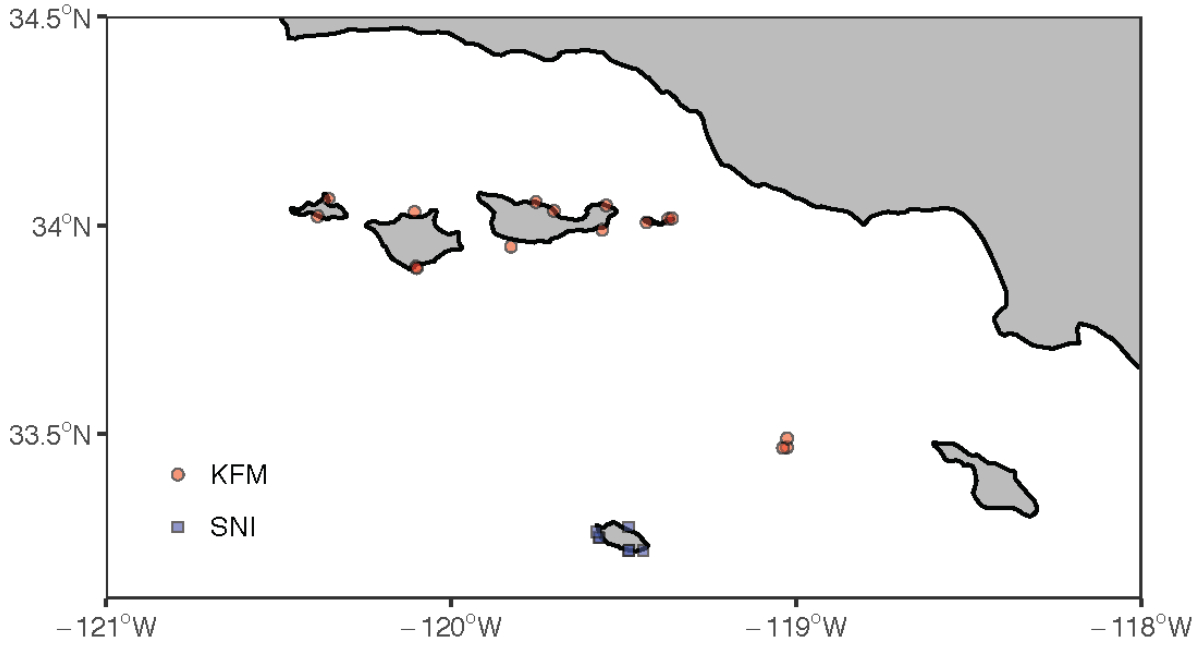


Figure 1: Map of study sites in the Southern California Bight. Colors and shapes indicate the programs, designated as either San Nicolas Island (SNI) or the Channel Islands Kelp Forest Monitoring Program (KFM).

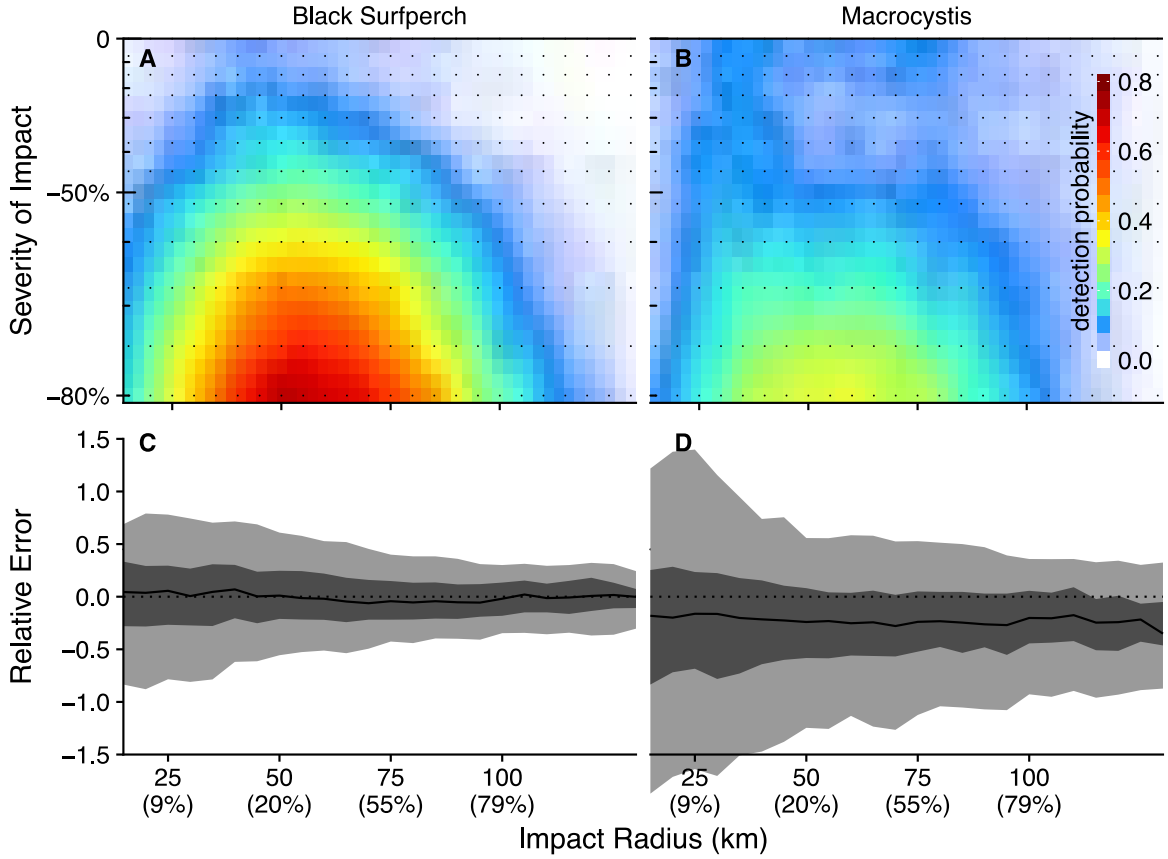


Figure 2: Detection probability for BACI estimates (a & b) and relative error and mean bias of BACI estimates (c & d) for two species as a function of the impact radius and proportion of sites impacted (in parentheses). Top panels (a & b) illustrate power as a function of impact severity and impact radius. Lower panels (c & d) present Relative Error at an impact severity of -80%, where bands are the 90% quantiles (light grey) interquartile range (dark grey) and median bias (black line). Relative error is the log-estimate minus the log-true value divided by the log-true value. See figures A1 and A2 for results for different species.

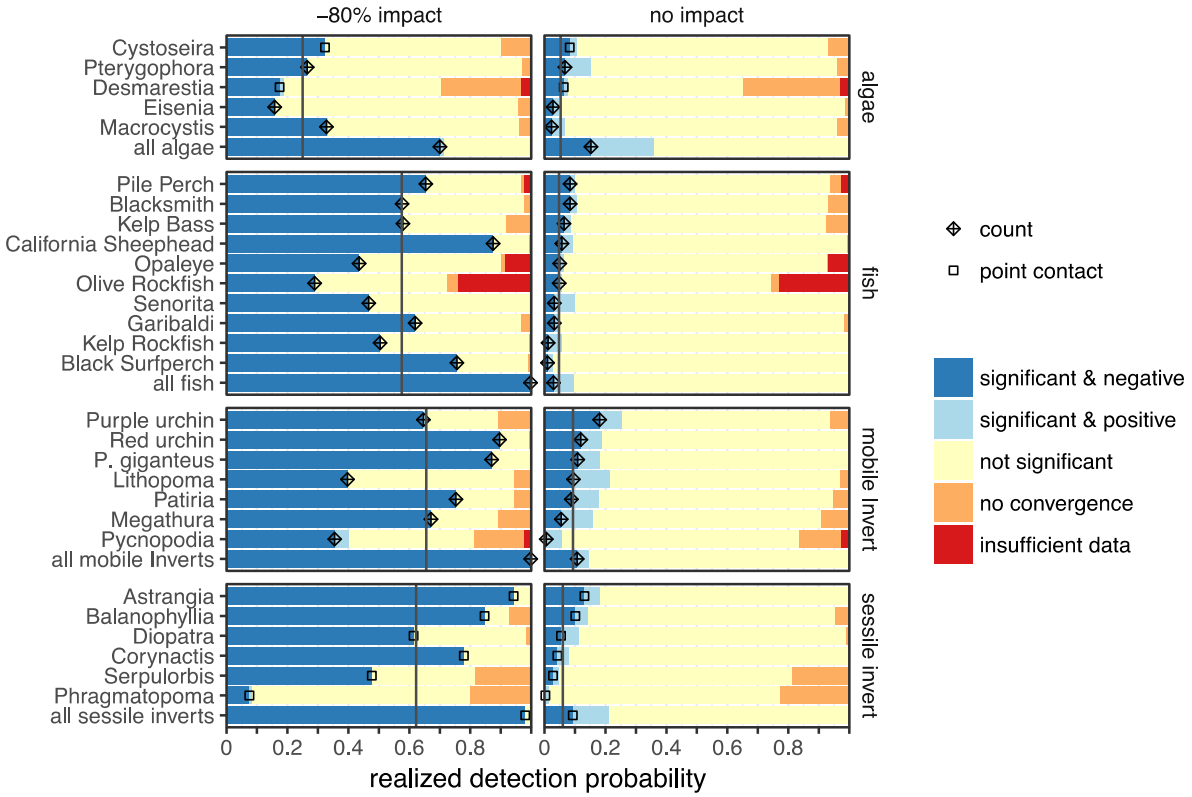


Figure 3: Detection probabilities for scenarios with a large negative impact (-80% or an 80% reduction) versus no impact for individual species, and using all species within the group in a nested framework (the lowest bar in each panel). Colors represent proportions of different outcomes from the simulations: Significant & negative and significant & positive indicate a significant effect at alpha = 0.025 along with the direction of the estimated BACI effect. No convergence indicates models that failed to converge despite having sufficient data for analysis, where data sufficiency is at least two sites in each of the BA and CI categories with at least 15% of the years with the species present. Symbols represent different data types (counts versus point contact).

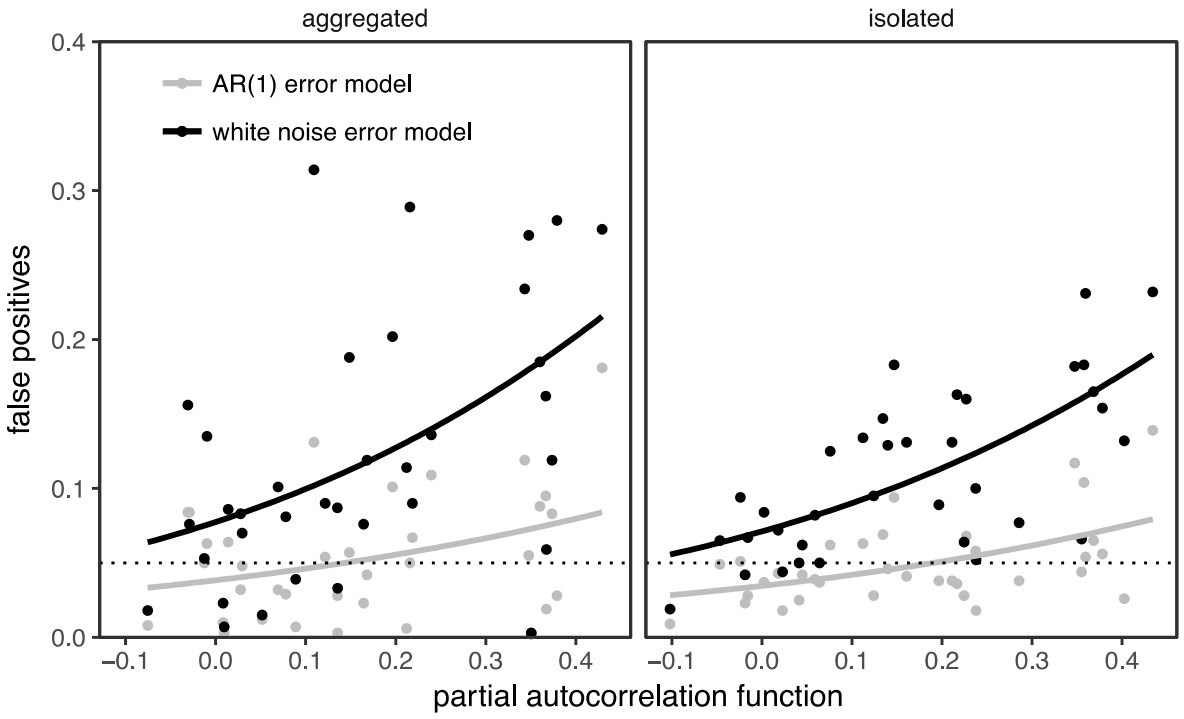


Figure 4: Effect the mean partial autocorrelation function on rate of false positives, with (grey) and without (black) accounting for first order residual correlations in the models. Each point represents the mean for an individual species across 1000 simulations. Panels represent scenarios where the impacts are spatially aggregated or isolated (i.e. no spatial pattern in the impact).

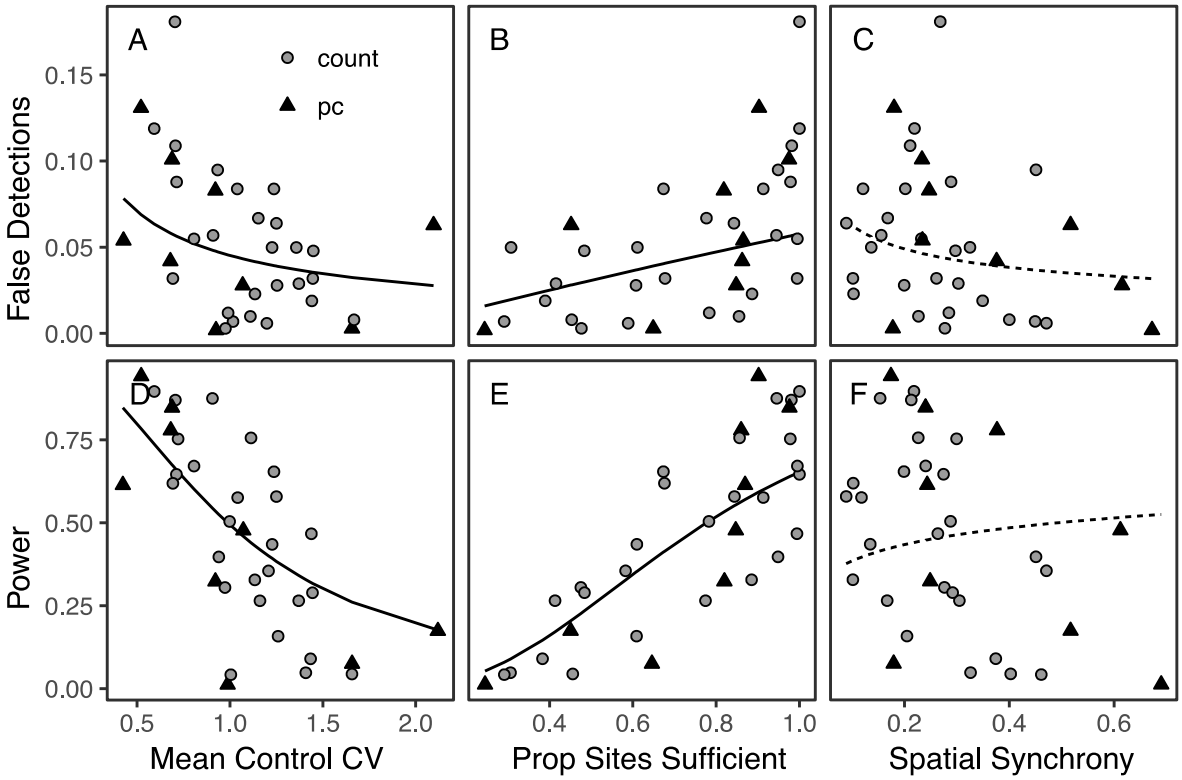


Figure 5: Correlations between statistical properties of the time series and power (at a -80% impact; top row) and false positives (significant BACI effect despite no impact; bottom row). Solid lines indicate a significant correlation after accounting for all other variables in the GLM shown in Table 1.

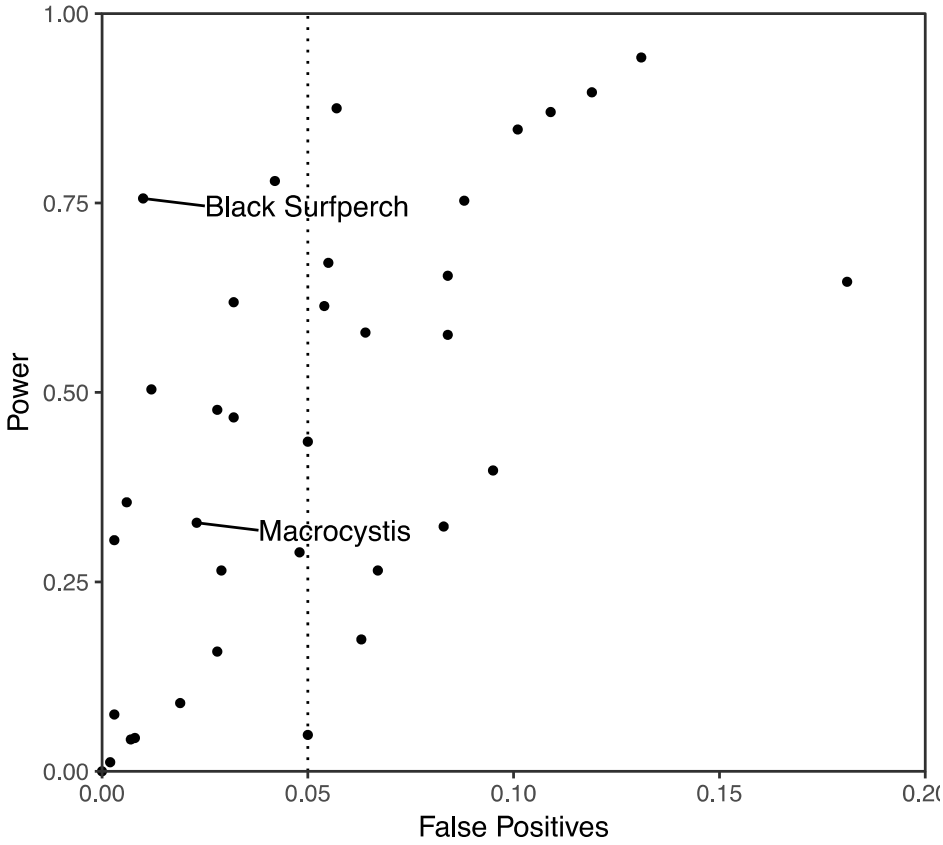


Figure 6: Power (detection probability under a -80% impact) versus false positives (detection probability under no impact), where points represent individual species.

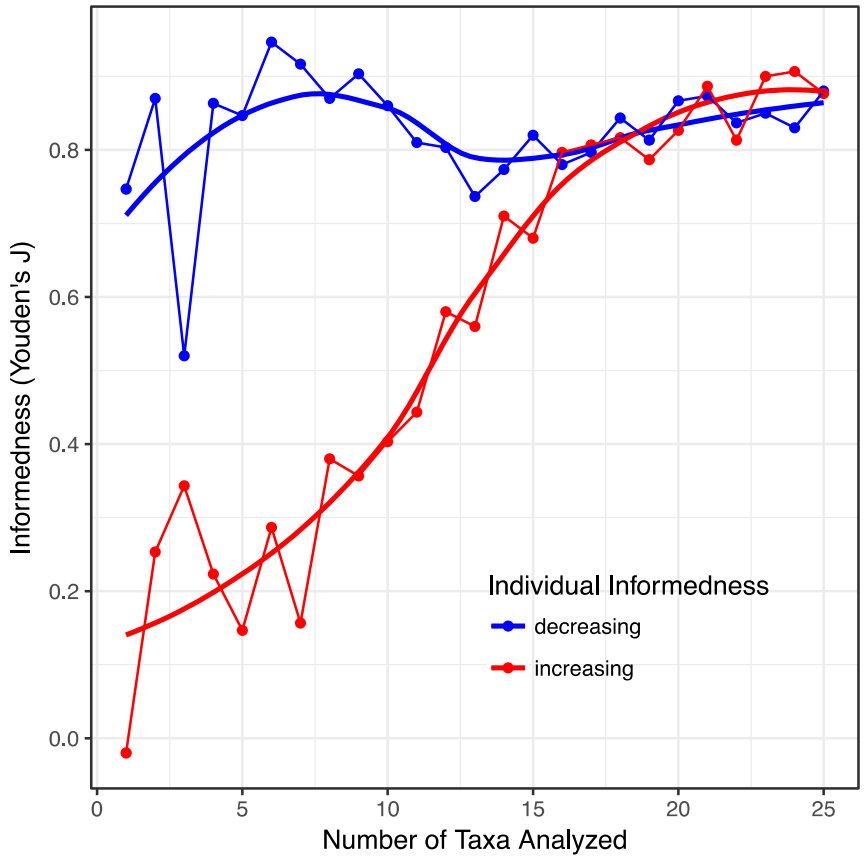


Figure 7: Informedness as a function of the number of taxa in the grouped analysis and order in which species are added by their individual informedness. Informedness is the probability of correct inference under equal prior odds of impact or no impact) as a function.

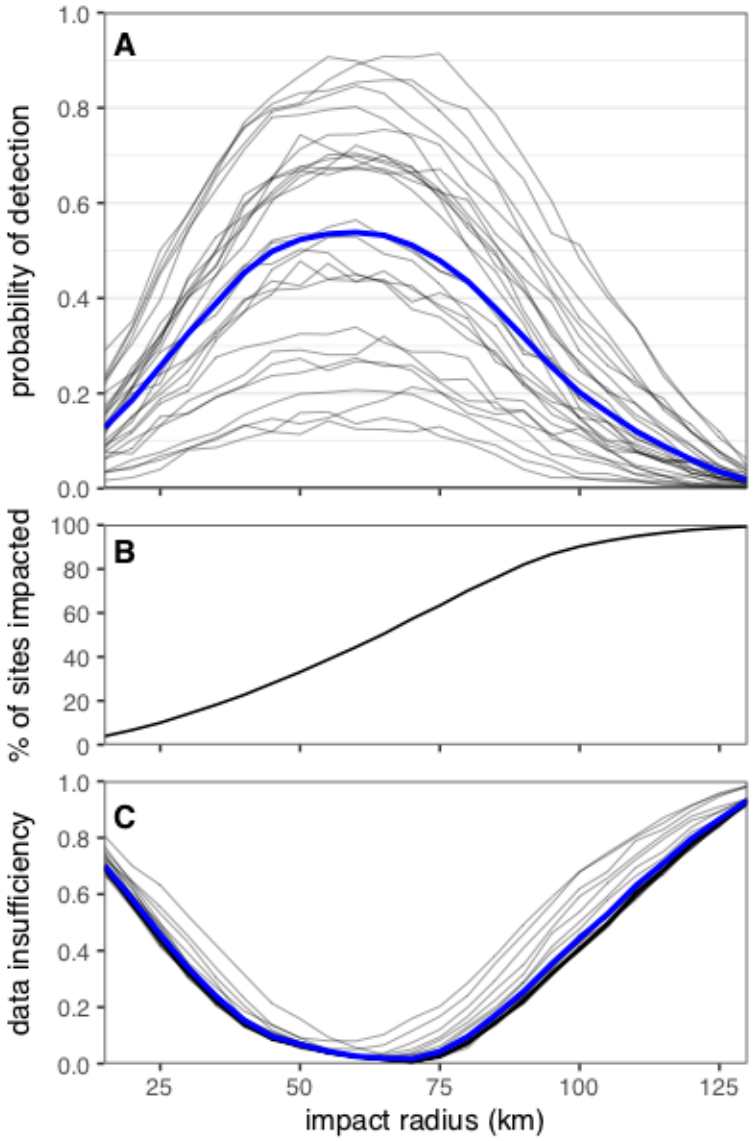


Figure A1: Effect of impact area on A) detection probability for BACI estimates for each species (thin grey lines) and the mean probability (blue line), B) % of sites impacted and C) probability of data insufficiency for analysis for each species (thin lines) and the mean probability (blue line). Results are shown for an -80% impact. Note that overlapping thin grey lines appear as a heavier black line in panels B and C.

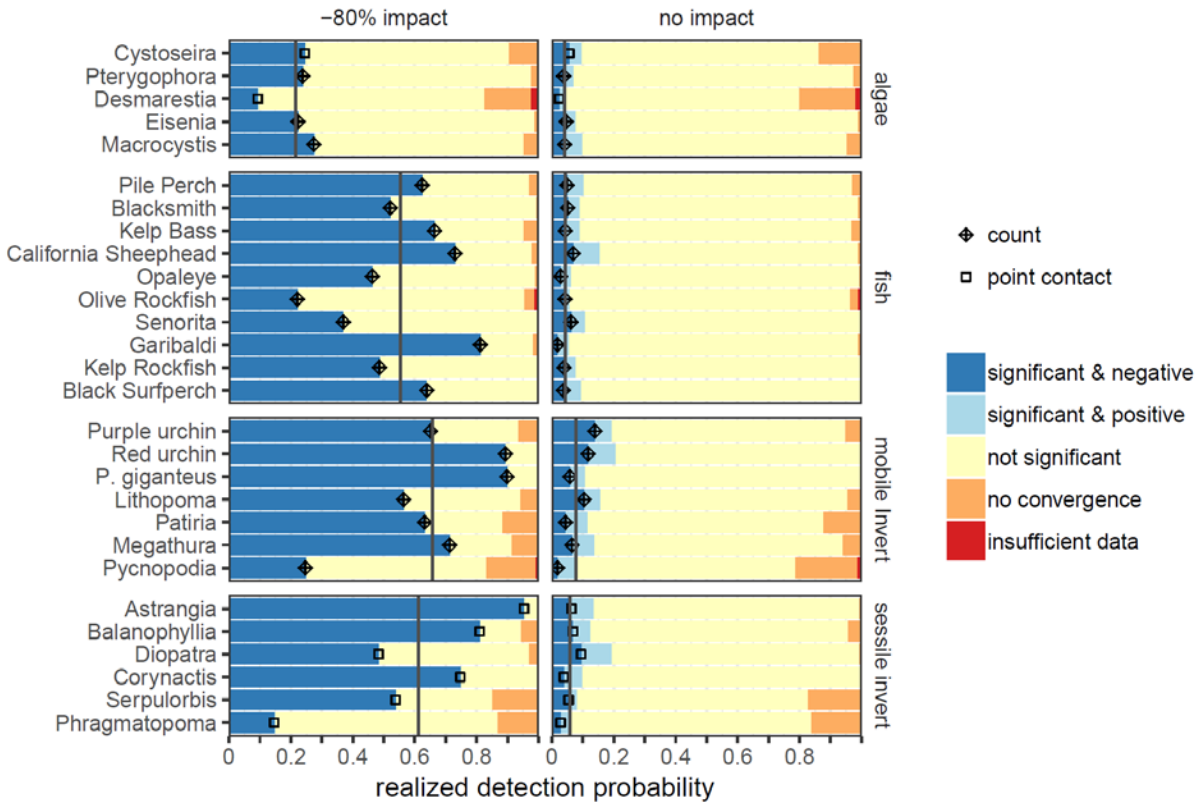


Figure A2: Detection probabilities for scenarios in which impacts were randomly distributed across the region, instead of spatially clustered. Left panels represent a large negative impact (-80% or an 80% reduction) whereas right panels represent scenarios where no impact was imposed. Individual species are plotted as well as results using all species within the group in a nested framework (the lowest bar in each panel). Colors represent proportions of different outcomes from the simulations: Significant & negative and significant & positive indicate a significant effect at alpha = 0.025 along with the direction of the estimated BACI effect. No convergence indicates models that failed to converge despite having sufficient data for analysis, where data sufficiency is at least two sites in each of the BA and CI categories with at least 15% of the years with the species present. Symbols represent different data types (counts versus point contact).

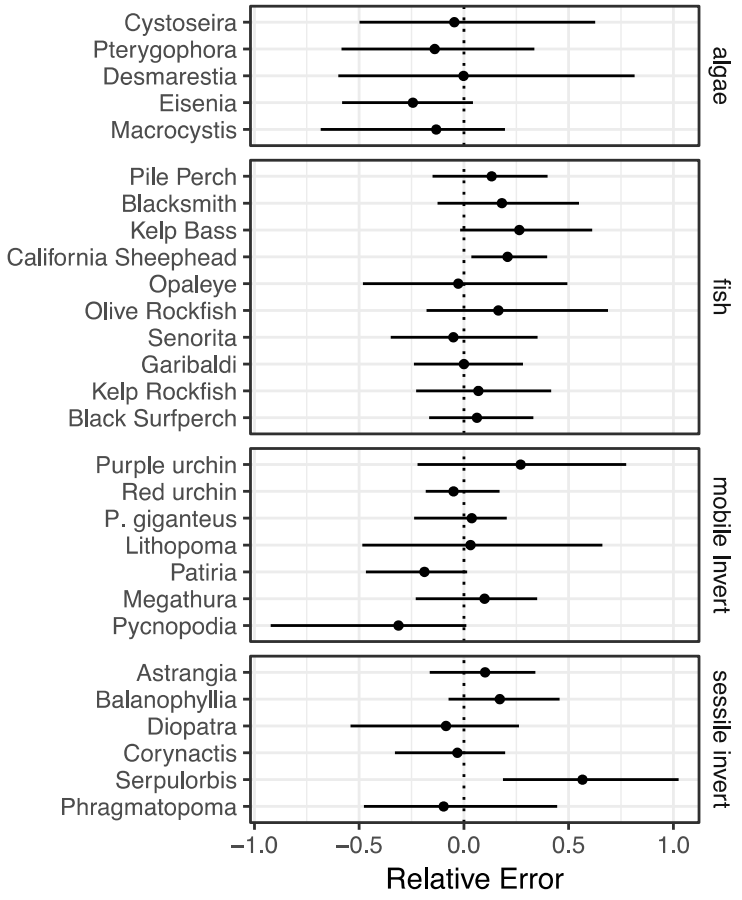


Figure A3: Relative error of the log estimated impact severity under a -80% impact for individual species.

Appendix E

Scale-specific drivers of kelp forest communities

Thomas Lamy, Daniel C. Reed, Andrew Rassweiler, David A. Siegel, Li Kui, Tom W. Bell, Rachel D. Simons, and Robert J. Miller

Abstract

Identifying spatial scales of variation in natural communities and the processes driving them is critical for obtaining a predictive understanding of biodiversity. In this study, we focused on diverse communities inhabiting productive kelp forests on shallow subtidal rocky reefs in southern California, USA. We combined long-term community surveys from 86 sites with detailed environmental data to determine what structures assemblages of fishes, invertebrates and algae at multiple spatial scales. We identified the spatial scales of variation in species composition using a hierarchical analysis based on eigenfunctions, and assessed how sea surface temperature (SST), water column chlorophyll, giant kelp biomass, wave exposure and potential propagule delivery strength contributed to community variation at each scale. Spatial effects occurring at multiple scales explained 60% of the variation in fish assemblages and 52% of the variation in the assemblages of invertebrates and algae. Most variation occurred over broad spatial scales (> 200km) consistent with spatial heterogeneity in SST and potential propagule delivery strength, while the latter also explained community variation at medium scales (65-200km). Small scale (1-65km) community variation was substantial but not linked to any of the measured drivers. Conclusions were consistent for both reef fishes and benthic invertebrates and algae, despite sharp differences in their adult mobility. Our results demonstrate the scale dependence of environmental drivers on kelp forest communities, showing that most species were strongly sorted along oceanographic conditions over various spatial scales. Such spatial effects must be integrated into models assessing the response of marine ecosystems to climate change.

Introduction

Ecological communities are complex systems. Identifying the processes responsible for their spatial distribution is an enduring, yet important goal of community ecology (Leibold et al. 2004, Legendre and Legendre 2012). One of the main challenges in achieving this goal is that many ecological processes act across multiple spatial scales, resulting in complex spatial patterns of community structure. Although the importance of spatial scale in community ecology is widely recognized (Menge and Olson 1990, Levin 1992, Schneider 2001, Dungan et al. 2002), recent changes in global biodiversity (Dornelas et al. 2014, McGill et al. 2015, Newbold et al. 2015, García Molinos 2016) highlight the need for a better understanding of how anthropogenic and natural processes interact at multiple scales to influence community structure and biodiversity conservation (Tzanopoulos et al. 2013; Socolar et al. 2016). Our ability to precisely quantify the effects of a given ecological process or environmental driver on biodiversity critically depends upon an understanding of the spatial scales over which they operate. Failure to account for this can lead to unproductive debates over the importance of a given process or driver on biodiversity loss (e.g., invasive species: Sax et al. 2002, Vilà et al. 2011).

Beta diversity, the compositional variation among communities (sensu Anderson et al. 2011, Legendre and Legendre 2012), offers a way to account for varying spatial scales ranging from small areas to larger regions, and has provided novel insight into the processes driving community structure (Chase 2010; Chase and Myers 2011; Chase et al. 2011; Vellend et al. 2014). However, spatial variation among communities (commonly referred to as metacommunity structure) can also occur across multiple scales

when the underlying processes generating this variation are themselves spatially structured. Complex and multiscale compositional variation among communities will ultimately arise when communities are controlled by multiple environmental factors that differ in their spatial structure (Legendre and Legendre 2012). A key assumption here is that dispersal limitation does not prevent species from tracking their environmental optimum, but instead allows them to be sorted along environmental gradients (Leibold et al. 2004). Not only does dispersal play a fundamental role in community ecology, as put forward by neutral theory (Hubbell 2001) and metacommunity theory (Leibold et al. 2004), but it can also strongly influence spatial variation among communities (Hubbell 2001) due to "true" spatial autocorrelation, the spatial variation due to dispersal limitation and/or local processes (Legendre 1993).

Spatial scales of measurement are frequently arbitrary and reflect either our perception of a system or practical sampling limitations. Multiscale modelling using eigenfunction analysis can be used as an objective tool to detect and define hierarchical scales of spatial variation among communities (Borcard and Legendre 2002; Dray et al. 2006, 2012). This approach comprises a family of methods in which eigenvectors computed from matrices of the spatial arrangement among localities are used as predictors in multivariate analyses. These spatial proxies can further serve to decompose the total spatial variation among communities into the most relevant scales of variation and aid in characterizing the underlying drivers. Spatial eigenfunction analysis has been successfully applied to a variety of systems and has led to an improved understanding of metacommunity structure (Borcard et al. 2004; Laliberté et al. 2009; Declerck et al. 2011). More importantly, the integration of spatial scales into metacommunity theory has helped resolve the ways that niche-based processes (e.g., species sorting), stochastic processes, and spatial processes structure natural communities at different spatial scales (Laliberté et al. 2009; Meynard et al. 2013; Ryberg and Fitzgerald 2015). For instance, evidence suggests that broad-scale patterns in metacommunity structure tend to arise primarily due to niche-based processes, while biotic interactions and stochastic processes can generate smaller-scale variation that can be difficult to identify (Jones et al. 2006; Laliberté et al. 2009; Declerck et al. 2011).

The interplay between environmental heterogeneity and dispersal limitation on metacommunity structure has received much theoretical attention (Leibold et al. 2004; Leibold and McPeek 2006), but disentangling the relative importance of these processes across spatial scales in nature remains challenging. Uncertainty about the role of these processes is especially high in complex marine metacommunities consisting of species that vary widely in their mobility, dispersal abilities and environmental optimum. Many marine species have complex life cycles, with dispersive propagule stages that can remain in the plankton for several weeks. The resulting potential for long distance dispersal is often assumed to reduce the likelihood of dispersal limitation. In addition, environmental heterogeneity can be high in marine ecosystems, and propagule survival and recruitment is often stochastic and unpredictable. Finally, marine species markedly differ in their dispersal abilities as adults. For instance, sessile algae and invertebrates remain stationary following recruitment, while mobile fishes have home ranges of widely varying size. Because of this multi-level complexity, the extent to which different environmental processes drive spatial variation in marine communities remains challenging to quantify (but see: Fraschetti et al. 2005; Terlizzi et al. 2007; MacNeil et al. 2009, Yeager et al. 2011, Menge et al. 2015).

Here we combined detailed community data collected from shallow subtidal reefs across a broad region in the northern portion of the Southern California Bight with environmental data obtained from satellite remote sensing and oceanographic models to understand the multiscale patterns and drivers of kelp forest communities. Forests of the giant kelp *Macrocystis pyrifera* on rocky reefs in the study region are extremely productive ecosystems (Reed et al. 2008) whose dynamics are influenced by a range of environmental factors operating at different spatial scales (Cavanaugh et al. 2011, 2013; Bell et al. 2015). Thus, the communities inhabiting this iconic ecosystem experience a spatially complex environment, at scales ranging from the local reef (~1000s m^2) to the regional Southern California Bight (~1000s km^2).

We seek to characterize the multiscale spatial patterns in kelp forest communities and understand the environmental conditions that underlie their variation across scales. The species included in this study span a wide variety of life history and dispersal abilities ranging from sessile to highly mobile adult forms, and non-dispersive to highly dispersive propagules (Reed et al. 2000; Shanks et al. 2003). We investigated whether differences in adult dispersal abilities affect these scales of variability by comparing assemblages of extremely mobile fishes with benthic assemblages of weakly mobile and sessile invertebrates and algae. Although propagule dispersal in these two groups varies widely, it tends to be higher in fishes than in invertebrates and algae (Kinlan and Gaines 2003; Shanks et al. 2003), which is consistent with patterns of adult movement. Thus, we hypothesized that the strong environmental heterogeneity occurring in the study region should underpin compositional variation in the assemblages of these two functional groups at multiple spatial scales. If both types of assemblages display compositional variation along the same spatial scales, it would suggest in kelp forest communities, the structuring effects of environmental sorting prevail over effects of dispersal limitation related to adult movement or propagule dispersal. We predicted that the metacommunity structure of both types of assemblages should vary across multiple scales due to the influence of regional oceanographic conditions at the Southern California Bight scale, and the effect of local kelp dynamics, wave disturbances or biotic interactions at smaller spatial scales. We tested this prediction using eigenfunction analyses aimed at: (1) characterizing the most important scales of variation in assemblages of kelp forest fish and benthic invertebrates and algae and (2) assessing the environmental driver(s) underlying each of these spatial scales.

Materials and methods

Ecological data

Our study area spans the mainland and offshore islands that bound the Santa Barbara Channel, the Santa Barbara Island and the San Nicolas Island to the south (Fig. 1a). It provides a uniquely advantageous setting for investigating the multiscale structure of nearshore marine communities in a relatively small region (*c.a.* 200 km long) that is characterized by strong environmental heterogeneity. Here, the cool southward flowing California Current and the warm northward flowing Southern California Counter Current meet to create highly heterogeneous oceanographic conditions (Harms and Winant 1998; Henderikx Freitas et al. 2017), while the coastal topography and offshore islands generate much variability in wave exposure. Giant kelp forests in the region are spatially structured due to the patchy distribution of rocky habitat and environmental processes (e.g. oceanographic conditions, wave exposure, grazing and recruitment) operating across a range of scales (Cavanaugh et al. 2011, 2013, Castorani et al. 2015).

Kelp forest communities were surveyed annually at 86 sites (Fig. 1a) between summer to early autumn from 2005 to 2014 by four monitoring programs: (1) the National Park Service Kelp Forest Monitoring program (KFM, 33 sites, Kushner et al. 2013), (2) the Partnership for Interdisciplinary Studies of Coastal Oceans (PISCO, 37 sites), (3) the Santa Barbara Coastal Long Term Ecological Research Project (SBC LTER, 11 sites, Reed 2016), and (4) the San Nicolas Island monitoring program (SNI, 5 sites, Kenner et al. 2013). We focused on 45 demersal fish species recorded at all 86 sites and 29 invertebrate and algal species that were recorded at 74 of the 86 sites (Electronic supplementary material Table S1 and Fig. S1). We computed the annual mean density of each fish species at the site level based on visual counts of individuals in spatially defined transects established by each monitoring program. We similarly computed the annual mean density of 19 benthic species of invertebrates and algae at the site level based on replicate counts of individuals in defined areas of varying size ranging from 1-60 m^2 depending on the species and monitoring program. We examined the abundance of 10 additional species of invertebrates and algae using percent cover data collected by random or uniform point contact sampling. The specific sampling methods used by each monitoring program to estimate the density or percent cover of each

species is summarized in Table S1. Species that were not recorded by all four programs, were rare, or not characteristic of rocky reef habitats were not included in the analysis.

Environmental variables

We used five variables to capture the spatial heterogeneity in oceanographic conditions and kelp forest structure across sites: (1) Sea surface temperature (SST); (2) chlorophyll-*a* concentration in surface waters (a surrogate for phytoplankton biomass); (3) wave exposure; (4) potential propagule delivery strength and (5) giant kelp canopy biomass, which is known to influence assemblages of reef fishes, invertebrates and algae. The variables were each averaged over a ten-year period (2005-2014) for each study site.

SST and chlorophyll *a* concentration. Coastal SST (°C) and chlorophyll *a* concentration (mg m^{-3}) were estimated using composite satellite measurements. SST was taken from the Multi-scale Ultra-high Resolution Sea Surface Temperature data set (MUR SST, http://mur.jpl.nasa.gov/), which provides daily SST values at 1 km spatial resolution. Chlorophyll *a* concentration was based on Aqua MODIS, which provides monthly chlorophyll *a* concentration at a 1.25 km spatial resolution (Henderikx Freitas et al. 2017). We extracted long-term SST and chlorophyll-*a* concentration values from the pixel overlaying each site to compute its average over the 10-year period.

Wave exposure. Maximum significant wave height at each site was estimated using data from the Coastal Data Information Program (CDIP; http://cdip.ucsd.edu). CDIP uses a wave propagation model to generate spatially-explicit significant wave height (H_s; mean height of the highest 1/3 of the waves) at 200 m resolution. Hourly estimates of H_s at a depth of 10 m were used to generate the daily maximum significant wave height (the top 1/3 of the hourly estimates). We extracted daily maximum significant wave height for each site as the value of the closest 200 m pixel to obtain the average over the 10-year period following Bell et al. (2015). We chose maximum daily values over mean daily values because extreme waves are more likely to influence kelp forests and their associated communities in the study region (Reed et al. 2008; Byrnes et al. 2011). In addition, these two variables were highly correlated among sites ($r = 0.99$).

Potential propagule delivery strength. A three-dimensional Regional Ocean Modeling System (ROMS) solution for the Southern California Bight (SCB) region (Dong et al. 2009) was used to estimate connectivity among sites via the delivery of propagules (i.e. spores and larvae) by ocean currents following methods discussed in Mitarai et al. (2009), Simons et al. (2013) and Siegel et al. (in review). More than 500 million Lagrangian particles were released each month from 135 nearshore ROMS connectivity cells of approximately 80 km^2, every 5 meters in depth from 5 to 30 meters below the surface (Fig. S2). Spatial probability density function (PDF) distributions for each source connectivity cell and release month were computed for particle advection times ranging from 1 to 70 days. The PDFs were used to form monthly connectivity matrices for each advection time among every ROMS connectivity cell in the SCB network and were used to calculate mean Lagrangian particle transit times among each pair of cells for each month (Siegel et al. in review). To quantify the relative "attractiveness" of a ROMS connectivity cell for propagule delivery, we computed the potential propagule delivery strength by summing the monthly connectivity evaluated at the mean transit time over all potential sources in the domain (i.e. excluding the focal ROMS cells; Siegel et al. in review). Finally, we averaged monthly potential propagule delivery strength between January 1996 and December 2007 to estimate the long-term average potential propagule delivery strength for each ROMS connectivity cell. Both long-term annual, as well as the vernal, averages were determined since propagule production for many species tends to be highest in the spring. These two indices were not as highly correlated ($r = 0.65$) as the other long-term seasonal averages (all three seasons vs. yearly average: $r > 0.94$), which suggested they convey

different aspects of the seasonal variability in coastal oceanographic connectivity. Because we were interested in characterizing spatial patterns of diverse communities, we were not able to impose general propagule characteristics (e.g., planktonic larval duration, size of the propagule source) on particles, as can be done when modelling the distribution of a single species (Watson et al. 2011b). Consequently, our approach only captured the physical aspect of coastal oceanographic connectivity (Grober-Dunsmore et al. 2009).

Giant kelp biomass. Giant kelp forms a dense floating canopy at the sea surface that influences the species composition and abundance of reef fishes, invertebrates and algae (reviewed in Schiel and Foster 2015). Data on kelp canopy biomass at 30 m spatial resolution were derived from Landsat 5 Thematic Mapper and Landsat 7 Enhanced Thematic Mapper + satellite images and calibrated using *in situ* data collected by divers (Bell et al. 2017) following methods introduced by Cavanaugh et al. (2011). We assessed canopy biomass of giant kelp in a 2 km buffer area around each site every 1 to 2 months to compute long-term average kelp canopy biomass from 2005 to 2014. The choice of a 2-km radius was guided both by the spatial scales over which these remote sensing techniques are reliable and the fact that choosing larger (e.g., 3 km) or smaller (e.g., 1 km) buffer areas resulted in highly correlated indices ($r >$ 0.90). Because kelp forests are extremely dynamic (Cavanaugh et al. 2011, 2013; Bell et al. 2015; Castorani et al. 2015), we also computed the temporal coefficient of variation of kelp canopy biomass within each buffer area to capture information on the temporal dynamics of kelp forests.

Space–time interaction

We hypothesized that kelp forest communities are primarily structured across space rather than time due to the strong spatial gradients in the oceanographic conditions of the Santa Barbara Channel (SBC). We tested for a space-time interaction in community composition using a multivariate analysis of variance by canonical Redundancy Analysis (RDA: Legendre and Anderson 1999; Legendre et al. 2010). Helmert contrasts were used to test for the main factors (space and time), while the space–time interaction was modelled based on the product of the first s/2 and t/2 spatial and temporal Principal Coordinates analysis of Neighbourhood Matrix (PCNM) variables, respectively, following Legendre et al. (2010), where s stands for the number of sites and t the number of years. We used type III sums-of-squares due to unequal replication in our sampling design (9% and 20% of site by year combinations were missing for fish surveys and invertebrate and algal surveys respectively, resulting in space factors, time factors and space-time interaction factors that were not fully orthogonal: see Fig. S1). Despite significant space-time interactions in our analysis (see Results), the spatial component of the total variation in both types of communities was predominant. In addition, our results provided similar conclusions when modelled separately for each year (see Table S2). Thus, because we were primarily interested in understanding the spatial structure of communities, we focus on results of analysing ecological data at the site level using long-term averages over the ten-year period.

Spatial variables

Based on the sampling design of the 86 sites, we developed three groups of spatial variables to analyse the multi-scale spatial variation in fish assemblages and benthic assemblages of invertebrates and algae at broad scale (> 200 km), medium scale (65-200 km) and small scale (1-65 km) respectively. We performed a hierarchical analysis of the multiscale spatial variation of communities, whereby the spatial variation was modelled sequentially from the largest to the smallest spatial scale by means of partial RDA.

The broad scale of variation in the two assemblages was modelled using the geographic coordinates of each site as spatial variables to capture spatial trends at broad scales. We then used spatial eigenfunctions derived from a connection matrix of neighbouring sites, where the connection among sites was defined as the overwater distance, to analyse finer-scale spatial variation in both assemblages. This method

corresponds to distance-based Moran's eigenvector maps (dbMEM; Dray et al. 2006) previously referred to as PCNM (Principal Coordinates analysis of Neighbor Matrices: Borcard and Legendre 2002, Borcard et al. 2004). dbMEM analysis provides a means of dissecting all the possible scales of variation encompassed by a sampling design. Each eigenfunction generated represents an orthogonal spatial variable that displays a given spatial correlation (Moran's *I;* Borcard and Legendre 2002, Dray et al. 2006). We first modelled the medium scale of variation in both assemblages based on eleven dbMEM variables obtained from the connection matrix among all sites, representing significant positive spatial structure at scales ranging from 65 km (*i.e.* the truncation distance: Borcard and Legendre 2002) to the entire Channel (~200 km). We then modelled variation in both assemblages at the small scale (1-65 km), using a staggered matrix of dbMEM variables for spatially clustered sites following Declerck et al. (2011). In this analysis, we built dbMEM variables for each spatial cluster (*i.e.* the six islands and the coastline; Fig. 1a) independently and assembled them into a staggered matrix, arranging dbMEM variables in blocks corresponding to each spatial cluster. Within these blocks, sites from the other spatial clusters were assigned the value 0 (Declerck et al. 2011). A dummy variable coding for the six spatial clusters was used to control for spatial differences among the six spatial clusters.

Lastly, we computed a multivariate spline correlogram to complement our hierarchical analysis for identifying the scales of variation in the two assemblages. The correlogram allowed us to investigate how spatial autocorrelation, measured as the Mantel statistic, varied as a function of overwater distance among sites. The confidence interval of the Mantel statistic was constructed based on bootstrap resampling with 999 replications (Bjørnstad and Falck 2001).

Spatial scales of variation in community composition and their underlying drivers

We used RDA to independently model the spatial structure of fish assemblages and benthic assemblages of invertebrates and algae as a function of the three groups of spatial predictors (broad, medium and small scales). For each group, we first pre-selected spatial variables by forward selection (Blanchet et al. 2008), performed RDAs, and tested the significance of the RDA axes through marginal tests (Legendre et al. 2011). Each significant RDA axis represents a distinct and significant pattern of spatial variation in the corresponding assemblage. The eigenvalue associated with each significant RDA axis was used to assess the variance explained by that scale as expressed by adjusted R^2.

To determine the extent to which environmental variables explain different spatial scales of assemblage composition, we modelled the variation along significant RDA axes as a function of the five categories of environmental variables (SST, chlorophyll-*a* concentration, potential propagule delivery strength, giant kelp biomass and wave exposure) using variation partitioning (Borcard et al. 1992; Peres-Neto et al. 2006). This allowed us to quantity how much variation at each scale was explained by: (1) environmental heterogeneity, *i.e.* the overall contribution of the five categories of variables, and (2) each of the five variables. In the latter case, variation partitioning was used to partition the total contribution of one variable into its unique contribution and the portion of its contribution shared with the effect of other variables. Shared contribution among variables can occur when they are correlated, and can be negative.

We Hellinger-transformed our data prior to the analyses, so that RDA and variation partitioning preserve the Hellinger distance among sites, which is more appropriate for compositional data and give lower weights to rare species that were potentially more prone to sampling error (Legendre and Gallagher 2001). Kelp canopy biomass and chlorophyll *a* concentration were square-root transformed. All statistical analyses were performed in R 3.0.2 (R Core Team 2014), using the *rda* and *varpart* functions of the vegan package. The spatial correlogram was computed using the ncf package, adapted to accommodate overwater distances.

Results

Space, time and their interaction explained significant variation in the assemblages of fish and benthic invertebrates and algae ($P < 0.001$ for all terms; Table 1). However, space had a more prominent role in explaining variation in assemblages of fish ($R^2 = 0.560$) and of invertebrates and algae ($R^2 = 0.668$), compared to time ($R^2 = 0.047$ and $R^2 = 0.027$ for fish assemblages and invertebrate and algal assemblages respectively) and the space-time interaction ($R^2 = 0.128$ and $R^2 = 0.120$ respectively). Hence, the following results focus on the long-term spatial variation in both assemblages.

Different ecological drivers can operate across multiple spatial scales, leading to complex and multiscale patterns of variation in kelp forest communities. Here we found that both assemblages were primarily structured over the broad spatial scale, which explained 39% of the total variation in fish composition ($F_{2,83} = 27.828$; $P = 0.001$; Fig. 1b-c and 3), and 23% of the total variation in invertebrate and algal composition ($F_{2,71} = 11.942$; $P = 0.001$; Fig. 2a-b and 4). Two significant RDA axes captured different facets of the broad scale of variation. The first axis of the broad scale of variation in fish assemblages ($F_{1,83} = 39.918$; $P = 0.001$) captured an east to west gradient (Fig. 1b), while the second ($F_{1,83} = 15.739$; $P = 0.001$) captured a north to south gradient (Fig. 1c), which explained 28% and 11% of the total variation in fish composition, respectively. Both gradients were accurately explained by the combination of all five environmental variables, which collectively accounted for 98.7% ($P < 0.001$) and 90.1% ($P < 0.001$) of the variation in fish species composition along the first and second gradients, respectively. SST and kelp biomass explained a large fraction of the east to west gradient when analysed independently (Fig. 5a and Table S3). However, the individual contribution of each variable was to a large extent shared with the contribution of other variables (Fig. 5a). For instance, only 33.3% of the 96.67%, contribution of SST to the east to west gradient was unique, while the remaining fraction could also be attributable to other variables, notably kelp biomass and to a lesser extent to potential propagule delivery strength or chlorophyll-*a* concentration. In contrast, the explanation of the north to south gradient was less confounded. Potential propagule delivery strength was the best predictor of this gradient, explaining 70.0%, more than half of which was a unique contribution (Fig. 5b). Wave exposure also explained a large fraction of this gradient (36.6%), though most of its effect was shared with chlorophyll *a* concentration and potential propagule delivery strength (Fig. 5b). Kelp biomass and SST did not contribute to the explanation of the north to south gradient.

The broad scale of variation in assemblages of invertebrates and algae was very similar to the one characterized for fishes. Indeed, two significant RDA axes also characterized the broad scale variation in invertebrates and algae and similarly captured an east to west gradient ($F_{1,71} = 17.9184$; $P = 0.001$; Fig. 2a), and a north to south gradient in the species composition of invertebrates and algae ($F_{1,71} = 5.966$; $P = 0.001$; Fig. 1b), which explained 17.3% and 5.8% of the total variation respectively. Here again, the five environmental variables collectively accounted for most of the variation in the first ($R^2_a = 98.1\%$; $P < 0.001$) and second gradient ($R^2_a = 90.7\%$; $P < 0.001$). SST and kelp biomass were tightly related to the east to west gradient in invertebrate and algal species composition, explaining 95.6% and 63.3% of this gradient respectively, though most of their contributions were also confounded (Fig. 6a and Table S3). Indeed, only 27.7% and less than 1% of the contribution of SST and kelp biomass were unique contributions. Potential propagule delivery strength and wave exposure were tightly related to the north to south gradient, explaining 54.6% and 41.3% of the second gradient, half of the contribution of potential propagule delivery strength (32.8%) being unique (Fig. 6b and Table S3).

Eleven dbMEM variables were used to model the medium spatial scale (≥ 65 km to ~200 km) of variation. Among these, three variables (#1, 2 and 8) for fishes and four variables (#1, 2, 3 and 5) for invertebrates and algae were retained following forward selection to capture the medium scale of variation. The latter explained 11.6% ($F_{3,80} = 7.434$; $P = 0.001$) of the total variation in fishes (Fig. 3), and 15.5% ($F_{4,67} = 5.464$; $P = 0.001$) of the total variation in invertebrates and algae, after controlling for the

broad scale (Fig. 4). Only the first RDA axis of the medium scale of variation in fish species composition was significant (R^2_a = 0.105; P = 0.001; Fig. 1d). Variation along this significant axis captured additional non-linear variation in fish species composition occurring in the SBC, such as differences between the mainland coastline and the islands, and among the islands and sides of islands (Fig. 1d). For instance, both the northeast side of Santa Cruz Island and the southern side of Santa Rosa Island displayed similar scores along this RDA axis. All environmental variables taken together explained 62.9% of the medium scale variation in fish species composition (Fig. 5c). Taken independently, each variable had a unique contribution (Fig. 5c and Table S3), but potential strength delivery strength was the strongest contributor, explaining 42.6% of the variation in fish assemblages at this scale. Kelp biomass explained 11.3% of the variation at this scale, while all other variables had marginal contributions. By contrast, two RDA axes of the medium scale of variation in the assemblages of invertebrates and algae were significant (R^2_a = 0.109; P = 0.001 and R^2_a = 0.029; P = 0.001; Fig. 2c-d). All environmental variables collectively explained 37.4% of these two axes, with unique contributions by potential propagule delivery strength (15.3%) and kelp biomass (13.5%; Fig. 6c and Table S3).

The small scale (1-65 km) explained 10% of the total variation in fishes and 14% of the total variation in invertebrates and algae after controlling for the broad and medium scales (Fig. 3 & 4). Two RDA axis of the small scale of variation in the fish assemblages were significant (axis 1: R^2_a = 0.032; P = 0.003; axis 2: R^2_a = 0.026; P = 0.018) and captured very small-scale features of the spatial variation occurring within islands or along the mainland coastline (Fig. S3a-b). However, all five environmental variables collectively explained only 32.2% of this small-scale variation (Fig. 5d). Two RDA axis also explained the small scale of variation in invertebrate and algal assemblages (axis 1: R^2_a = 0.075; P = 0.001; axis 2: R^2_a = 0.022; P = 0.043; Fig. S3c-d), but were also poorly explained by the five environmental variables (R^2_a = 0.24; P = 0.001; Fig. 6d). Of the five environmental variables, chlorophyll-*a* concentration had the highest contribution to the small scale of variation in assemblages of fishes (11.5%), and in assemblages of invertebrates and algae (9.5%).

Finally, 13%, 19% and 67% of the broad, medium and small scales of variation in fish assemblages (Fig. 3), and 17%, 15% and 65% of the broad, medium and small scales of variation in invertebrate and algal assemblages (Fig. 4) were uniquely explained by our spatial variables. This corresponds to purely spatial effects attributable to either dispersal limitation or unmeasured variables. After controlling for broad, medium and small scale spatial variables, local scale effects of the five environmental variables explained ~1% of the total variation in the assemblages of fishes and assemblages of invertebrates and algae (Fig. 3 & 4). The spatial multivariate correlogram further supported the identification of these scales of variation in both assemblages (Fig. 7). The correlogram for fishes displayed significant positive spatial correlation at distances ranging from 0 to 33.7 km (95% confidence interval [28.4-40]), while the positive spatial correlation in invertebrates and algae was no longer significant for overwater distances > 31.3 km (95% confidence interval [24.6-40.7]). Beyond this distance, assemblages of fishes and assemblages of invertebrates and algae exhibited significant negative spatial correlation, suggesting that they were more dissimilar than expected by chance.

Discussion

Understanding the causes of spatial variation in natural communities can be extremely complex when the underlying processes vary across multiple scales. In coastal ecosystems, the importance of multiple biotic and abiotic processes (e.g., species interactions, dispersal and environmental forcing) in explaining the local composition of marine communities (Connell 1971; Dayton 1971) or their spatial structure (Paine 1974; Dayton et al. 1984; Underwood and Chapman 1996) has been emphasized repeatedly. However, understanding precisely how these processes interact to shape the structure of coastal marine communities at multiple scales remains a challenge. Many assume that spatial variation will mainly arise over very broad scales because of presumed high connectivity in marine systems (Tittensor et al. 2010; Magurran et

al. 2015), while some evidence suggests that spatial variablity can also occur over very small spatial scales (Hughes et al. 1999; Edwards 2004). In this study, we used spatially extensive data from sustained monitoring efforts in the SBC to show that the spatial variation in assemblages of mobile fishes and benthic invertebrates and algae can be complex and manifested at multiple spatial scales. Nonetheless, our results suggest that most of the spatial variation arises at broad spatial scales, especially for fishes. We found strong evidence that the broad scale of variation in the species composition of kelp forest communities in the SBC was driven by an SST gradient running from west to east, as well as by north to south variation in oceanographic connectivity. The major effect of SST is consistent with findings identifying ocean temperature as a predominant driver of marine communities globally (Tittensor et al. 2010) as well as in our study region (Blanchette et al. 2007, 2009). Previous studies have investigated determinants of spatial patterns in community composition in the SBC, focusing on specific aspects of the environmental variability (e.g. SST, physical disturbance), dispersal and oceanographic connectivity (Reed et al. 2000; Kinlan and Gaines 2003; Broitman and Kinlan 2006; Blanchette et al. 2009; Watson et al. 2011a). While these studies suggested that SST and oceanographic connectivity play a role in structuring reef communities, we found that they act at different spatial scales, and hence underlie two different gradients in the structure of kelp forest communities. Medium scale variation in both fishes and the benthic assemblage of invertebrates and algae was closely associated with heterogeneity in oceanographic connectivity within the SBC. Small scale variation in both types of assemblages was significant, yet we could not accurately attribute it to any of the environmental variables included in our model. It is often difficult to attribute small-scale variation to environmental variables, in part due to biotic interactions, or variation in environmental factors that is smaller than the scale at which they are measured (Borcard et al. 2004). In our study, variation among sites in depth, substrate characteristics and species interactions undoubtedly account for some of the unexplained variation in our analyses. Site differences in coastal topography and morphology might also explain additional variation, as they can influence nutrient flux, wave exposure and propagule delivery (Broitman and Kinlan 2006; Lester et al. 2007). Small- and medium-scale effects might be more important for assemblages of sessile invertebrates and algae, for which broad scale drivers explained less variability than was the case for fishes.

By taking into account multiple spatial scales, one can evaluate how the relative influence of stochastic processes (e.g., dispersal limitation) and species sorting vary in space. Previous contributions have shown how stochastic processes can be a dominant force structuring metacommunities at one scale, while environmental conditions can become predominant at either lager (Jones et al. 2006; Laliberté et al. 2009) or smaller (Declerck et al. 2011) spatial scales. In coastal marine systems, regional differences in climate are known to be important in structuring communities over large scales whereas a combination of factors including local stochastic perturbations, successional processes and ecological interactions can substantially shape communities at small scales. Among these processes, dispersal plays an important role in community organization and spatial variation. Species with short-distance dispersal tend to be primarily structured by spatial factors, reflecting the importance of stochastic processes such as dispersal limitation, whereas assemblages of species with a capacity for longer range dispersal tend to be more influenced by regional environmental conditions (Reed et al. 2000; Hubbell 2001; Flinn et al. 2010; Schroeter et al. 2015). Species examined in this study differed greatly in the dispersal abilities of their propagules and adult forms (Shanks et al. 2003). The fact that we found very similar spatial scales of variation in assemblages of fishes and of benthic invertebrates and algae, despite differences in their mobility (e.g. Reed et al. 2000; Kinlan and Gaines 2003), suggests that environmental conditions prevail over dispersal limitation in controlling the structure of these communities. The lack of a strong pure spatial effect at any of the spatial scales studied further suggests that adult mobility and limitations on propagule dispersal do not play an overwhelming role in structuring the kelp forest metacommunity in the SBC. Our study considered the importance of adult mobility by contrasting mobile fish with benthic invertebrates and algae. Although these groups likely differ on average in larval dispersal potential (with

fish having potential for longer distance dispersal), pelagic larval duration is not known for many of these species. Our results suggest that dispersal limitation does not play a large role at the community level, but it is a potentially important predictor of the spatial distribution of the most dispersal limited species (Reed et al. 2000; Simons et al. 2016). A more detailed analysis of the roles of pelagic larval duration and the size of source populations on the spatial structure of rocky reef communities would be valuable especially for short dispersers.

Paradoxically, potential propagule delivery strength explained a large portion of the north to south gradient and of the medium scale variability in communities despite the overall lack of evidence for dispersal limitation effects. This suggests that dispersal limitation does not prevent species from tracking their environmental optima but that connectivity is still an important driver of metacommunity dynamics within the SBC. The effect of potential propagule delivery was due in large part to greater particle connectivity of the mainland sites compared with the island sites (Mitarai et al. 2009; Siegel et al. in review). Here, we used the mean transit time for Lagrangian particles between two sites to represent the physical connection between kelp forest habitats. In that respect, our index captured structural aspects related to the spatial configuration of the habitat in relation to oceanographic connectivity (Grober-Dunsmore et al. 2009; Olds et al. 2016). Integrating propagule dispersal abilities and movement of adult organisms among habitats has provided critical information on the demographic connectivity of single species (Watson et al. 2011b). Demographic connectivity has been shown to be a key process controlling metapopulations of giant kelp, which depend on the number of propagules available for dispersal as well as physical transport of those propagules (Castorani et al. 2015). However, further work is needed to fully understand the role connectivity in controlling kelp forest metacommunities that consist of a diverse array of species with different dispersal abilities (Melià et al. 2016; Magris et al. 2016).

Surprisingly, giant kelp biomass was not a good predictor of spatial community structure in SBC, especially at medium and small scales. Kelp biomass was an important predictor only at the largest scale (east to west gradient), although its influence was confounded with that of other variables, notably wave height and SST, which are known to influence kelp biomass in the region (Cavanaugh et al. 2011; Bell et al. 2015). Giant kelp is regarded as an important foundation species that provides habitat and food for many of the studied species (Graham 2004, Koenigs et al. 2015, Miller et al. 2015). However, fish are mobile; the home range of many species certainly exceeds the extent of a kelp forest, and metapopulations may shift between sites as a response to resource limitation or other factors. Alternatively, it could be that our definition of kelp canopy biomass does not capture the ecologically relevant aspects of giant kelp for fishes, and the relationship between kelp and its communities, or that kelp is more important for the temporal variability of rocky reef communities as patches of kelp come and go. Overall it suggests that kelp may not be as strong of a factor sorting fish species relative to SST and oceanographic connectivity.

This study benefited from the existence of four long-term programs designed to monitor temperate rocky reef ecosystems in the SBC. These data provided the large number of spatial replicates needed for multiscale analysis, and underscore the importance of long-term surveys (Hughes et al. 2017). Despite similar goals and sampling protocols, sampling effort varied across programs. To minimize potential resulting bias, we focused on the 74 species that could be appropriately compared across programs and standardized abundances to give less weight to rare species. Yet, we cannot disregard the possibility that combining data from several monitoring programs led to sampling bias that resulted in underestimating compositional changes at small scales, where changes can be more subtle. In addition, it is worth noting that grouping spatial variables into subsets to model different scales is a somewhat arbitrary process, as there is no general rule for defining broad, medium or small scales. Yet this represents the most effective way to dissect the spatial structure of natural communities (Borcard and Legendre 2002).

Islands and the mainland coastline (our spatial clusters) were the most parsimonious spatial predictors of assemblage structure. Indeed, the spatial structure of both fishes and the benthic assemblage of invertebrates and algae matched quite well with regional geomorphological structure. The maximum distances between any two intra-island sites for San Nicolas, Santa Barbara, Anacapa and San Miguel islands did not exceed 17 km, far below the 34 km and 31 km scale of variability characterized in the correlogram for the fish assemblages and the invertebrate and algal assemblage, respectively. In contrast, the larger Santa Cruz Island, Santa Rosa Island, and the mainland coast had sites that were separated by distances greater than these thresholds. Other processes in addition to connectivity likely contributed to the observed differences between island and mainland communities; for example, water clarity is typically greater around the islands than the mainland due to the lack of land-derived sediment around the islands (Ebeling et al. 1980; Henderikx Freitas et al. 2017). Declining water quality can have complex effect on coastal communities, as it can both directly influence recruitment, growth or survival of species as well as indirectly modulate species interactions (Airoldi 2003).

While some marine species can be primarily controlled by processes occurring over small spatial scales (e.g. Hughes et al. 1999; Edwards 2004), our results highlight the prominent role of large-scale environmental drivers on the spatial structure of kelp forest communities. The presence of the two broad-scale gradients indicate that SST and oceanographic connectivity act at scales even broader than the entire SBC, suggesting that the effects of global warming on kelp forest community structure should be evaluated at very large spatial scales (Wernberg et al. 2016; Krumhansl et al. 2016; Vergés et al. 2016). The coastal waters of California, and the Santa Barbara Channel in particular, has been the focus of recent conservation efforts designed to protect its valuable marine resources from human exploitation (Saarman and Carr 2013). Our results suggest that much of the variability in the kelp forest communities of the SBC is induced by environmental heterogeneity occurring over broad scales that may conflict with conservation efforts aimed at smaller local scales. 26 sites were located within marine protected areas, yet after accounting for the strong environmental gradients, the protection status of these sites did not explain the spatial variation of species composition in kelp communities. This suggests that existing reserves may not be significantly structuring communities in space, even though their effect on harvested species is indisputable (Caselle et al. 2015). A better understanding of how natural processes and anthropogenic factors interact across different spatial scales to affect biodiversity is needed to improve the effectiveness of conservation efforts. Ideally, the spatial scales at which conservation policies are developed should be matched with the scales at which relevant ecological processes occur (Socolar et al. 2016) and planned in ways that consider future changes in climate, which are predicted to greatly affect coastal marine communities including kelp forests (Wang et al. 2015; Coleman et al. 2017).

Acknowledgements

The authors are thankful to Daniel Borcard for stimulating discussions on multiscale analyses and for providing useful comments on an early version of this manuscript. This research was supported by the National Aeronautics and Space Administration Biodiversity and Ecological Forecasting program (NASA Grant NNX14AR62A), the Bureau of Ocean and Energy Management Ecosystem Studies program (BOEM award MC15AC00006) and NOAA in support of the Santa Barbara Channel Biodiversity Observation Network and the US National Science Foundation in support of the Santa Barbara Coastal Long Term Ecological Research program.

References

Airoldi L. 2003. The effects of sedimentation on rocky coast assemblages. Oceanography and Marine Biology 41:161–236.

Anderson MJ, Crist TO, Chase JM, et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14:19–28.

Bell TW, Cavanaugh KC, Reed DC, Siegel DA. 2015. Geographical variability in the controls of giant kelp biomass dynamics. Journal of Biogeography 42:2010–2021.

Bell TW, Cavanaugh KC, Siegel DA. 2017. SBC LTER: Time series of quarterly NetCDF files of kelp biomass in the canopy from Landsat 5, 7 and 8, 1984 - 2016 (ongoing). St. Barbar. Coast. LTER.

Bjørnstad ON, Falck W. 2001. Nonparametric spatial covariance functions: Estimation and testing. Environmental and Ecological Statistics 8:53–70.

Blanchet FG, Legendre P, Borcard D. 2008. Forward selection of explanatory variables. Ecology 89:2623–2632.

Blanchette CA, Helmuth B, Gaines SD. 2007. Spatial patterns of growth in the mussel, Mytilus californianus, across a major oceanographic and biogeographic boundary at Point Conception, California, USA. Journal of Experimental Marine Biology and Ecology 340:126–148.

Blanchette CA, Raimondi PT, Broitman BR. 2009. Spatial patterns of intertidal community structure across the California Channel Islands and link to ocean temperature. Proceedings of the 7th California Islands Symposium:161–173.

Borcard D, Legendre P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecoloical Modelling 153:51–68.

Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85:1826–1832.

Borcard D, Legendre P, Drapeau P. 1992. Partialling out the spatial component of ecological variation. Ecology 73:1045–1055.

Broitman BR, Kinlan BP. 2006. Spatial scales of benthic and pelagic producer biomass in a coastal upwelling ecosystem. Marine Ecology Progress Series 327:15–25.

Byrnes JEK, Reed DC, Cardinale BJ, et al. 2011. Climate-driven increases in storm frequency simplify kelp forest food webs. Global Change Biology 17:2513–2524.

Caselle JE, Rassweiler A, Hamilton SL, Warner RR. 2015. Recovery trajectories of kelp forest animals are rapid yet spatially variable across a network of temperate marine protected areas. Scientific Reports 5:14102.

Castorani MCN, Reed DC, Alberto F, et al. 2015. Connectivity structures local population dynamics: A long-term empirical test in a large metapopulation system. Ecology 96:3141–3152.

Cavanaugh KC, Kendall BE, Siegel DA, et al. 2013. Synchrony in dynamics of giant kelp forests is driven by both local recruitment and regional environmental controls. Ecology 94:499–509.

Cavanaugh KC, Siegel DA, Reed DC, Dennison PE. 2011. Environmental controls of giant-kelp biomass in the Santa Barbara Channel, California. Marine Ecology Progress Series 429:1–17.

Chase JM. 2010. Stochastic community assembly causes higher biodiversity in more productive environments. Science 328(5984):1388–1391.

Chase JM, Kraft NJB, Smith KG, et al. 2011. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. Ecosphere 2(2):1-11.

Chase JM, Myers JA. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions Royal Society B 366:2351–2363.

Coleman MA, Cetina-Heredia P, Roughan M, et al. 2017. Anticipating changes to future connectivity within a network of marine protected areas. Global Change Biology 23:3533–3542.

Connell JH. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In: Den Boer PJ, Gradwell GR (eds) Dynamics of Populations. PUDOC, pp 298–312.

Cottenie K. 2005. Integrating environmental and spatial processes in ecological community dynamics. Ecology Letters 8:1175–1182.

Dayton PK. 1971. Competition, disturbance, and community organization: the provision and subsequent utilization of space in a rocky intertidal community. Ecological Monographs 41:351–389.

Dayton PK, Currie V, Gerrodette T, et al. 1984. Patch dynamics and stability of some California kelp communities. Ecological Monographs 54:254–289.

Declerck SAJ, Coronel JS, Legendre P, Brendonck L. 2011. Scale dependency of processes structuring metacommunities of cladocerans in temporary pools of High-Andes wetlands. Ecography 34:296–305.

Dong CM, Idica EY, McWilliams JC. 2009. Circulation and multiple-scale variability in the Southern California Bight. Progress in Oceanography 82:168-190.

Dornelas M, Gotelli NJ, McGill B, et al. 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science 344:296–299.

Dray S, Legendre P, Peres-Neto PR. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196:483–493.

Dray S, Pélissier R, Couteron P, et al. 2012. Community ecology in the age of multivariate multiscale spatial analysis. Ecological Monographs82:257–275.

Dungan JL, Perry JN, Dale MRT, et al. 2002. A balanced view of scale in spatial statistical analysis. Ecography 25:626–640.

Ebeling AW, Larson TJ, Alevizon WS. 1980. Habitat groups and island-mainland distribution of kelp-bed fishes off Santa Barbara, California. *In*: Power, D.M. (ed.). *Multidisciplinary symposium on the California Islands*. Santa Barbara Museum of Natural History, Santa Barbara, California, USA

Edwards MS. 2004. Estimating scale-dependency in disturbance impacts: El Niños and giant kelp forests in the northeast Pacific. Oecologia 138:436–447.

Flinn KM, Gouhier TC, Lechowicz MJ, Waterway MJ. 2010. The role of dispersal in shaping plant community composition of wetlands within an old-growth forest. Journal of Ecology 98:1292–1299.

Fraschetti S, Terlizzi A, Benedetti-Cecchi L. 2005. Patterns of distribution of marine assemblages from rocky shores: evidence of relevant scales of variation. Marine Ecology Progress Series 296:13–29.

García Molinos J, Halpern BS, Schoeman DS, et al. 2016. Climate velocity and the future global redistribution of marine biodiversity. Nature Climate Change 6:83–88.

Graham HM. 2004. Effects of local deforestation on the diversity and structure of Southern California giant kelp forest food webs. Ecosystems 7:341–357.

Grober-Dunsmore R, Pittman SJ, Caldow C, et al. 2009. A Landscape Ecology Approach for the Study of Ecological Connectivity Across Tropical Marine Seascapes. In: Nagelkerken I (ed) Ecological Connectivity among Tropical Coastal Ecosystems. Springer Netherlands, Dordrecht, pp 493–530

Harms S, Winant CD. 1998. Characteristic patterns of the circulation in the Santa Barbara Channel. Journal Geophysical Research 103:3041–3065.

Henderikx Freitas F, Siegel DA, Maritorena S, Fields E. 2017. Satellite assessment of particulate matter and phytoplankton variations in the Santa Barbara Channel and its surrounding waters: Role of surface waves. Journal of Geophysical Research Oceans 122:355–371.

Hubbell SP. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, N.J.

Hughes TP, Baird AH, Dinsdale EA, et al. 1999. Patterns of recruitment and abundance of corals along the Great Barrier Reef. Nature 397:59–63.

Hughes BB, Beas-Luna R, Barner AK, et al. 2017. Long-Term Studies Contribute Disproportionately to Ecology and Policy. Bioscience 67:271–281.

Kenner MC, Estes JA, Tinker MT, et al. 2013. A multi-decade time series of kelp forest community structure at San Nicolas Island, California (USA). Ecology 94:2654.

Kinlan BP, Gaines SD. 2003. Propagule dispersal in marine and terrestrial environments: A community perspective. Ecology 84:2007–2020.

Koenigs C, Miller RJ, Page HM. 2015. Top predators rely on carbon derived from giant kelp *Macrocystis pyrifera*. Marine Ecology Progress Series 537:1–8.

Krumhansl KA, Okamoto DK, Rassweiler A, et al. 2016. Global patterns of kelp forest change over the past half-century. Proceedings of the National Acadamy of Sciences 113:13785–13790.

Kushner DJ, Rassweiler A, McLaughlin JP, Lafferty KD. 2013. A multi-decade time series of kelp forest community structure at the California Channel Islands. Ecology 94:2655.

Laliberté E, Paquette A, Legendre P, Bouchard A. 2009. Assessing the scale-specific importance of niches and other spatial processes on beta diversity: a case study from a temperate forest. Oecologia 159:377–388.

Legendre P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74:1659–1673.

Legendre P, Anderson MJ. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. Ecological Monographs69:1–24.

Legendre P, Cáceres M De, Borcard D. 2010. Community surveys through space and time: testing the space–time interaction in the absence of replication. Ecology 91:262–272.

Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

Legendre P, Legendre L. 2012. Numerical ecology, Third Edit. Elsevier.

Legendre P, Oksanen J, ter Braak CJF. 2011. Testing the significance of canonical axes in redundancy analysis. Methods in Ecology and Evolution 2:269–277.

Leibold M a., Holyoak M, Mouquet N, et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

Leibold MA, McPeek MA. 2006. Coexistence of the niche and neutral perspectives in community ecology. Ecology 87:1399–410.

Lester SE, Gaines SD, Kinlan BP. 2007. Reproduction on the edge: large-scale patterns of individual performance in a marine invertebrate. Ecology 88:2229–2239.

Levin SA. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73:1943–1967.

MacNeil MA, Graham NAJ, Polunin NVC, et al. 2009. Hierarchical drivers of reef-fish metacommunity structure. Ecology 90:252–264.

Magris RA, Treml EA, Pressey RL, Weeks R. 2016. Integrating multiple species connectivity and habitat quality into conservation planning for coral reefs. Ecography 39:649–664.

Magurran AE, Dornelas M, Moyes F, et al. 2015. Rapid biotic homogenization of marine fish assemblages. Nature Communications 6:8405.

McGill BJ, Dornelas M, Gotelli NJ, Magurran AE. 2015. Fifteen forms of biodiversity trend in the Anthropocene. Trends in Ecology and Evolution 30:104–113.

Melià P, Schiavina M, Rossetto M, et al. 2016. Looking for hotspots of marine metacommunity connectivity: a methodological framework. Scientific Reports 6:23705.

Menge BA, Gouhier TC, Hacker SD, et al. 2015. Are meta-ecosystems organized hierarchically? A model and test in rocky intertidal habitats. Ecological Monographs85:213–233.

Menge BA, Olson AM. 1990. Role of scale and environmental factors in regulation of community structure. Trends in Ecology and Evolution 5:52–57.

Meynard CN, Lavergne S, Boulangeat I, et al. 2013. Disentangling the drivers of metacommunity structure across spatial scales. Journal of Biogeography 40:1560–1571.

Miller RJ, Page HM, Reed DC. 2015. Trophic versus structural effects of a marine foundation species, giant kelp *Macrocystis pyrifera*). Oecologia 179:1199–1209.

Mitarai S, Siegel DA, Watson JR, et al. 2009. Quantifying connectivity in the coastal ocean with application to the Southern California Bight. Journal of Geophysical Research Ocean 114:C10026.

Newbold T, Hudson LN, Hill SLL, et al. 2015. Global effects of land use on local terrestrial biodiversity. Nature 520:45–50.

Olds AD, Connolly RM, Pitt KA, et al. 2016. Quantifying the conservation value of seascape connectivity: a global synthesis. Global Ecology and Biogeography 25:3–15.

Paine RT. 1974. Intertidal community structure. Oecologia 15:93–120.

Peres-Neto PR, Legendre P, Dray S, Borcard D. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 87:2614–2625.

Reed DC. 2016. SBC LTER: Reef: Kelp Forest Community Dynamics: Fish abundance. Santa Barbara Coastal LTER.

Reed DC, Raimondi PT, Carr MH, Goldwasser L. 2000. The role of dispersal and disturbance in determining spatial heterogeneity in sedentary organisms. Ecology 81:2011–2026.

Reed DC, Rassweiler A, Arkema KK. 2008. Biomass rather than growth rate determines variation in net primary production by giant kelp. Ecology 89:2493–2505.

Ryberg WA, Fitzgerald LA. 2015. Landscape composition, not connectivity, determines metacommunity structure across multiple scales. Ecography 39:932–941.

Saarman ET, Carr MH. 2013. The California Marine Life Protection Act: A balance of top down and bottom up governance in MPA planning. Marine Policy 41:41–49.

Sax DF, Gaines SD, Brown JH. 2002. Species invasions exceed extinctions on islands worldwide: a comparative study of plants and birds. The American Naturalist 160:766–783.

Schiel DR, Foster MS. 2015. The biology and ecology of giant kelp forests, 1st edn. University of California Press.

Schneider DC. 2001. The Rise of the Concept of Scale in Ecology. Bioscience 51:545–553.

Schroeter SC, Reed DC, Raimondi PT. 2015. Effects of reef physical structure on development of benthic reef community: A large-scale artificial reef experiment. Marine Ecology Progress Series 540:43–55.

Shanks AL, Grantham BA, Carr MH. 2003. Propagule dispersal distance and the size and spacing of marine reserves. Ecological Applications 13:S159–S169.

Siegel DA, Watson JR, Simons RD, et al. in review. Characterizing particle transit time metrics in a coastal ocean network.

Simons RD, Siegel DA, Brown KS. 2013. Model sensitivity and robustness in the estimation of larval transport: A study of particle tracking parameters. Journal of Marine Systems 119–120:19–29.

Simons RD, Page HM, Zaleski S, et al. 2016. The effects of anthropogenic structures on habitat connectivity and the potential spread of non-native invertebrate species in the offshore environment. PLoS One 11:e0152261.

Socolar JB, Gilroy JJ, Kunin WE, Edwards DP. 2016. How should beta-diversity inform biodiversity conservation? Trends in Ecology and Evolution 31:67–80.

Team RC. 2014. R: A language and environment for statistical computing.

Terlizzi A, Anderson MJ, Fraschetti S, Benedetti-Cecchi L. 2007. Scales of spatial variation in Mediterranean subtidal sessile assemblages at different depths. Marine Ecology Progress Series 332:25–39.

Tittensor DP, Mora C, Jetz W, et al. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466:1098–1101.

Tzanopoulos J, Mouttet R, Letourneau A, et al. 2013. Scale sensitivity of drivers of environmental change across Europe. Global Environmental Change 23:167–178.

Underwood AJ, Chapman MG. 1996. Scales of spatial patterns of distribution of intertidal invertebrates. Oecologia 107:212–224.

Vellend M, Srivastava DS, Anderson KM, et al. 2014. Assessing the relative importance of neutral stochasticity in ecological communities. Oikos 123:1420–1430.

Vergés A, Doropoulos C, Malcolm HA, et al. 2016. Long-term empirical evidence of ocean warming leading to tropicalization of fish communities, increased herbivory, and loss of kelp. Proceedings of the National Academy of Sciences 113:13791–13796.

Vilà M, Espinar JL, Hejda M, et al. 2011. Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters 14:702–708.

Wang D, Gouhier TC, Menge BA, Ganguly AR. 2015. Intensification and spatial homogenization of coastal upwelling under climate change. Nature 518:390–394.

Watson JR, Hays CG, Raimondi PT, et al. 2011a. Currents connecting communities: nearshore community similarity and ocean circulation. Ecology 92:1193–1200.

Watson JR, Siegel DA, Kendall BE, et al. 2011b. Identifying critical regions in small-world marine metapopulations. Proceedings of the National Academy of Sciences 108:E907–E913.

Wernberg T, Bennett S, Babcock RC, et al. 2016. Climate-driven regime shift of a temperate marine ecosystem. Science 353:169 LP-172.

Yeager LA, Layman CA, Allgeier JE. 2011. Effects of habitat heterogeneity at multiple spatial scales on fish community assembly. Oecologia 167:157–168.

Table 1. Spatio-temporal variation in fish assemblages and in assemblages of invertebrates and algae over a ten-year period (2005-2014) in the Santa Barbara Channel. The variance explained by each component (R^2) was tested using a multivariate analysis of variance by canonical RDA. *P* is the probability value.

	Fish assemblages				Invertebrate and alga assemblages			
Source	**d.f.**	R^2	*F*	*P*	**d.f.**	R^2	*F*	*P*
Space	85	0.56	13.849	0.001	73	0.668	24.041	0.001
Time	9	0.047	11.059	0.001	9	0.027	7.894	0.001
Space ×time	215	0.128	1.253	0.001	215	0.120	1.463	0.001
Residuals	466				391			

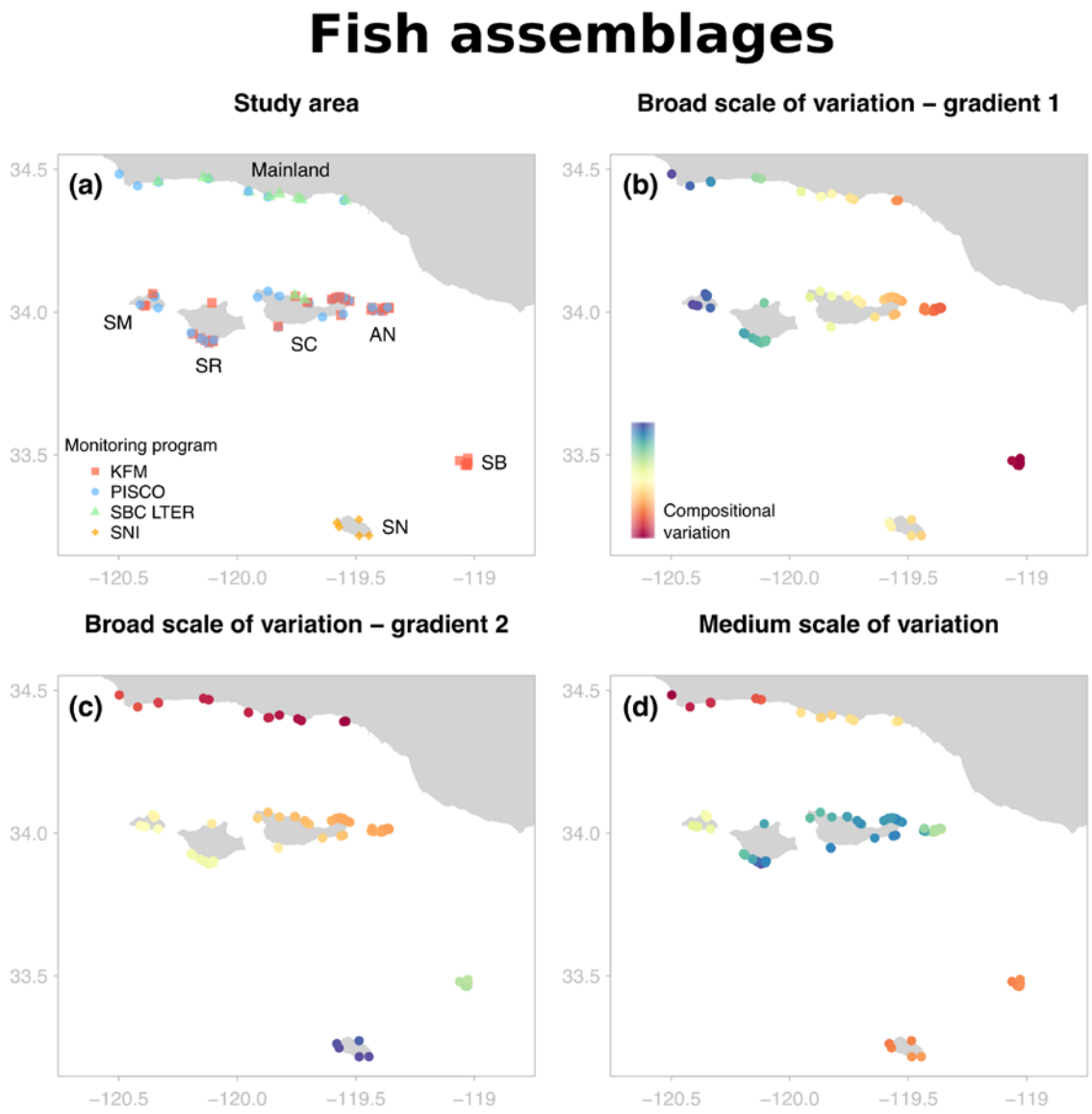


Fig. 1. Spatial variation in fish assemblages at different scales. (a) Study area showing the Santa Barbara Channel. Islands correspond to San Miguel (SM), Santa Rosa (SR), Santa Cruz (SC), Anacapa (AN), Santa Barbara (SB) and San Nicolas (SN). East to West gradient (b) and North to South gradient (c) of the broad scale of variation in fish assemblages. (d) Medium scale (MS) of variation in fish assemblages. Each symbol represents one of the 86 sites, and colour scales with the variation in species composition (*i.e.* sites displaying similar colour have similar species composition at that scale). The Small scale of variation is showed in Fig. S3.

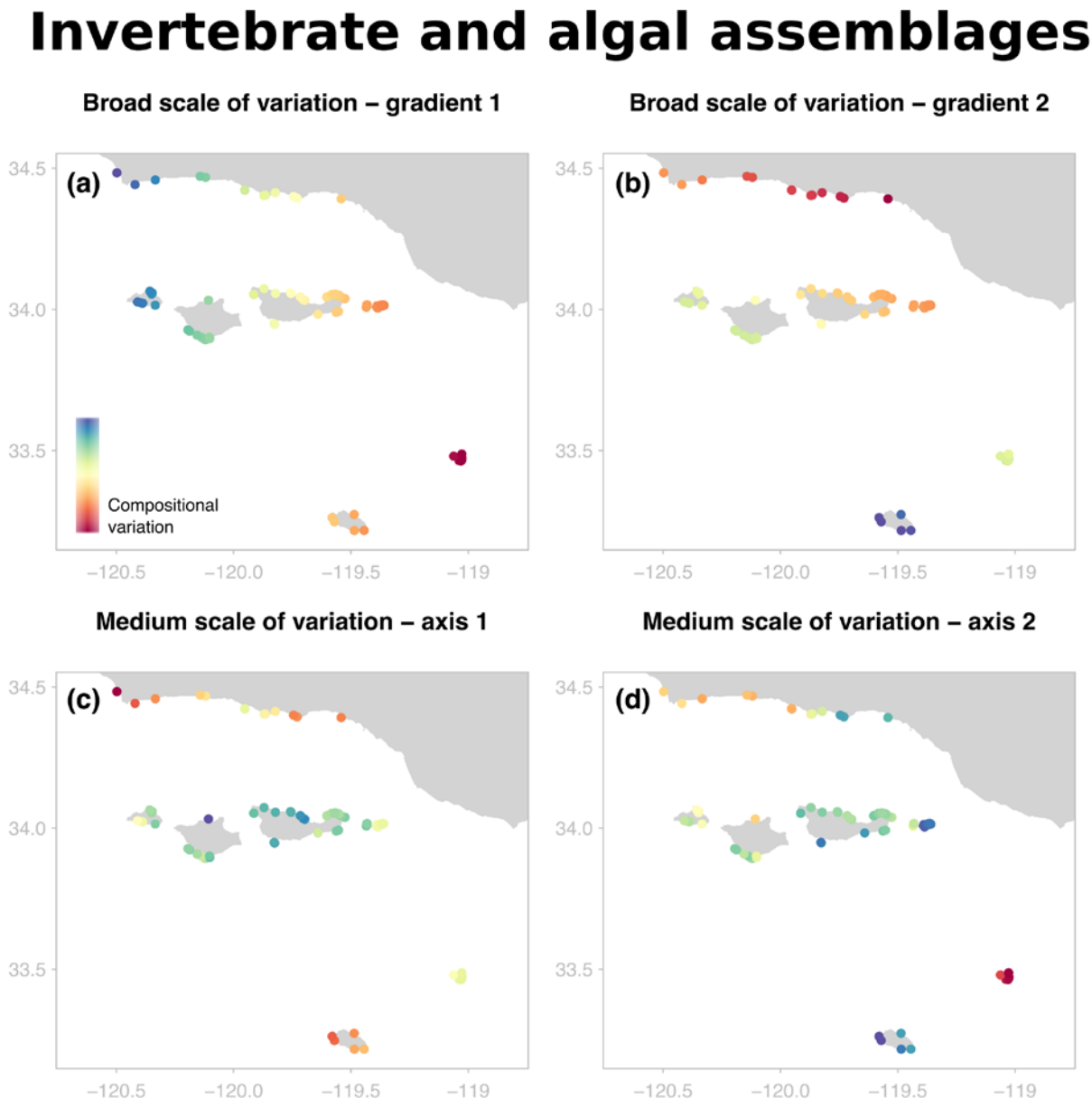


Fig. 2. Spatial variation in invertebrate and alga assemblages at different scales. East to West gradient (a) and North to South gradient (b) of the broad scale of variation in invertebrate and algal assemblages. First (c) and second (d) axis of the medium scale of variation in invertebrate and algal assemblages. Each symbol represents one of the 74 sites (resp. 74 for the invertebrate and algal assemblages), and colour scales with the variation in species composition (*i.e.* sites displaying similar colour have similar species composition at that scale). The Small scale of variation is showed in Fig. S3.

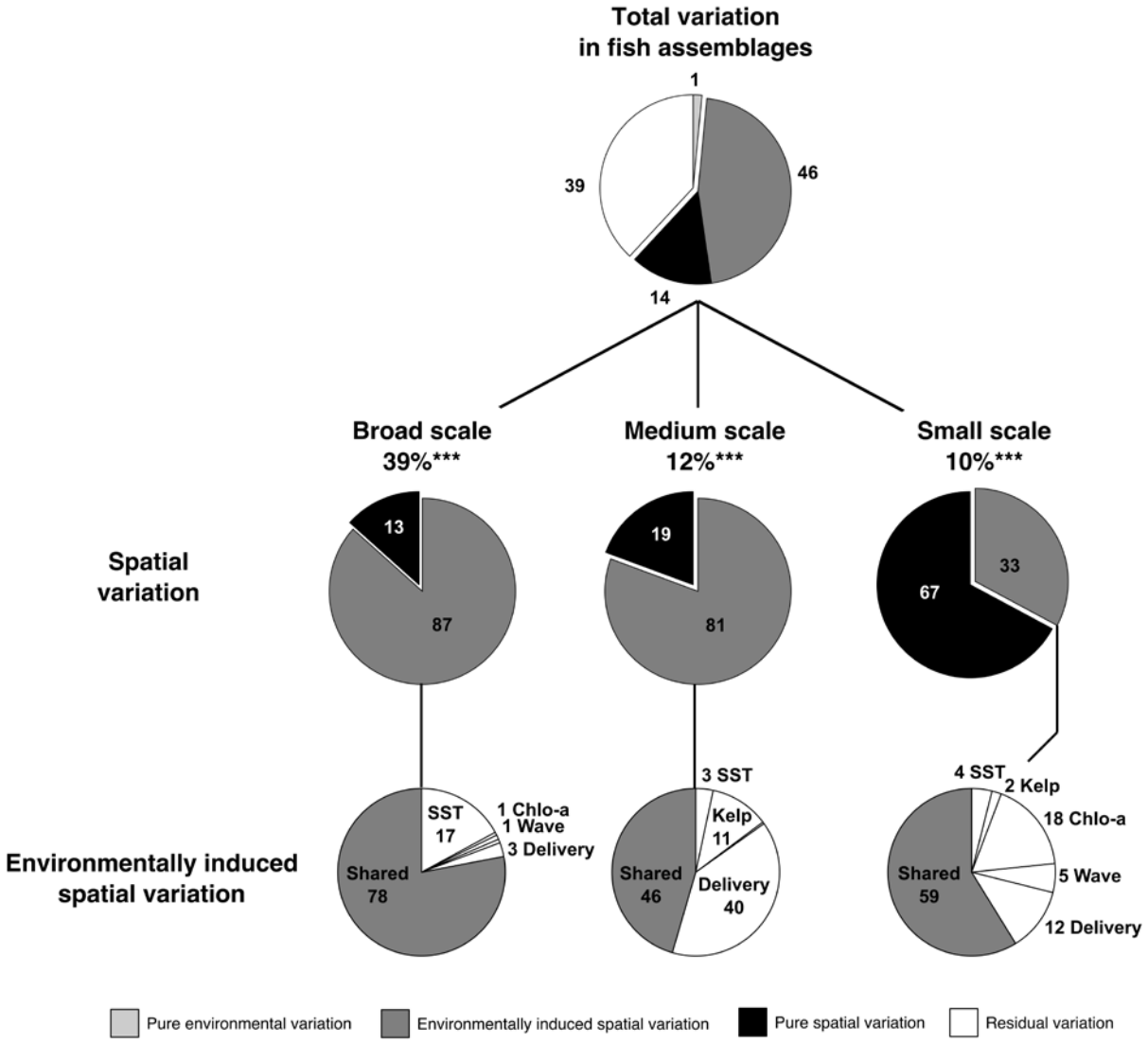


Fig. 3. Relative influence of spatial and environmental effects on the variation in fish assemblages at different spatial scales. The total variation is decomposed into purely spatial (black), environmentally induced spatial (dark grey), purely environmental (light grey) and residual (white) components. The total spatial variation (*i.e.* due to purely spatial effects and induced by the environment) is further decomposed into different spatial scales. Finally, the environmentally induced spatial variation at each scale is decomposed into the contributions of the five environmental variables. Their total contributions are divided into a unique contribution (white) and a shared contribution. The importance of each effect is given by the percentage outside or inside the chart and is calculated based on adjusted R^2.

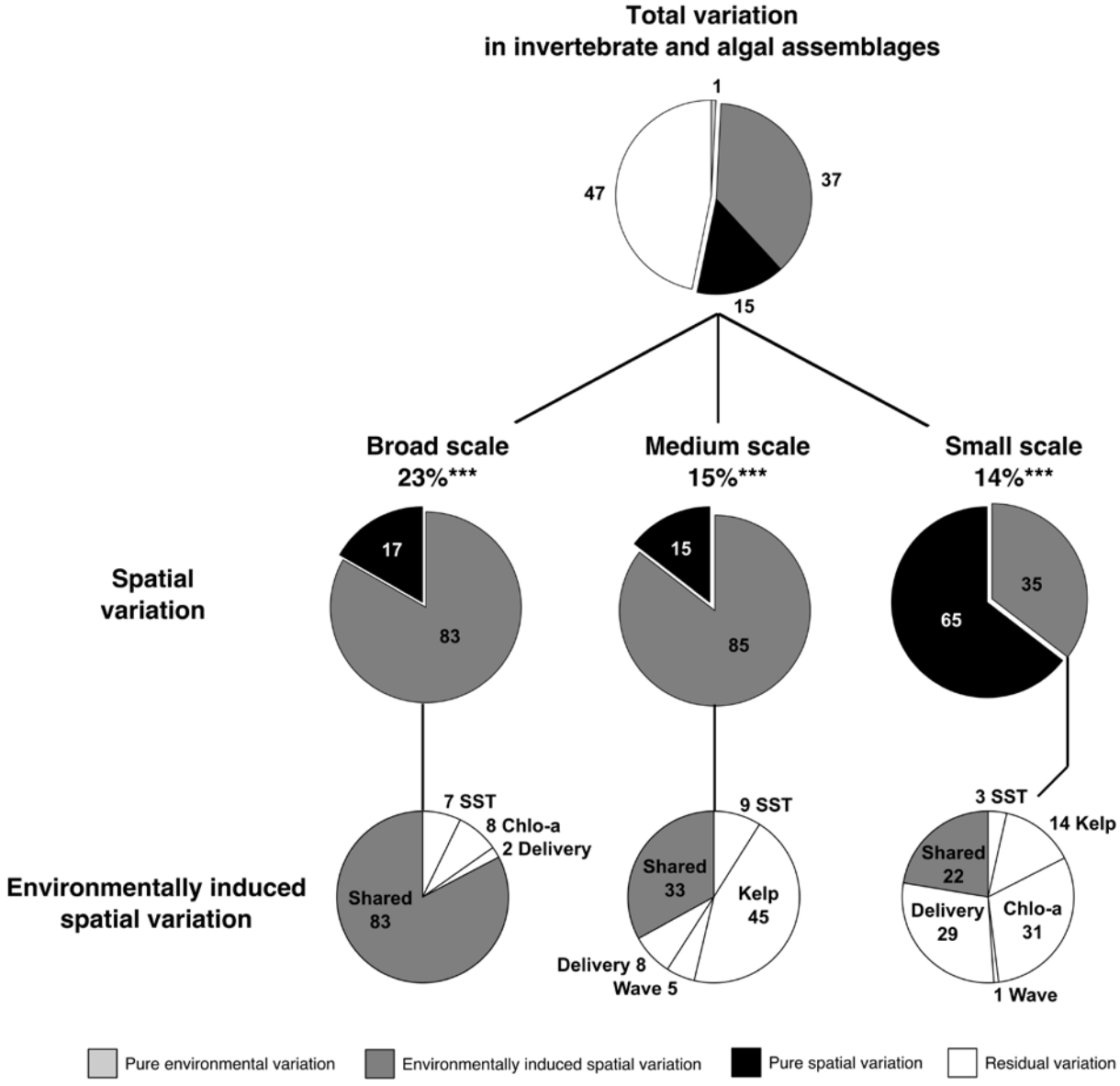


Fig. 4. Relative influence of spatial and environmental effects on the variation in invertebrate and algal assemblages. See Fig. 3 for legend description.

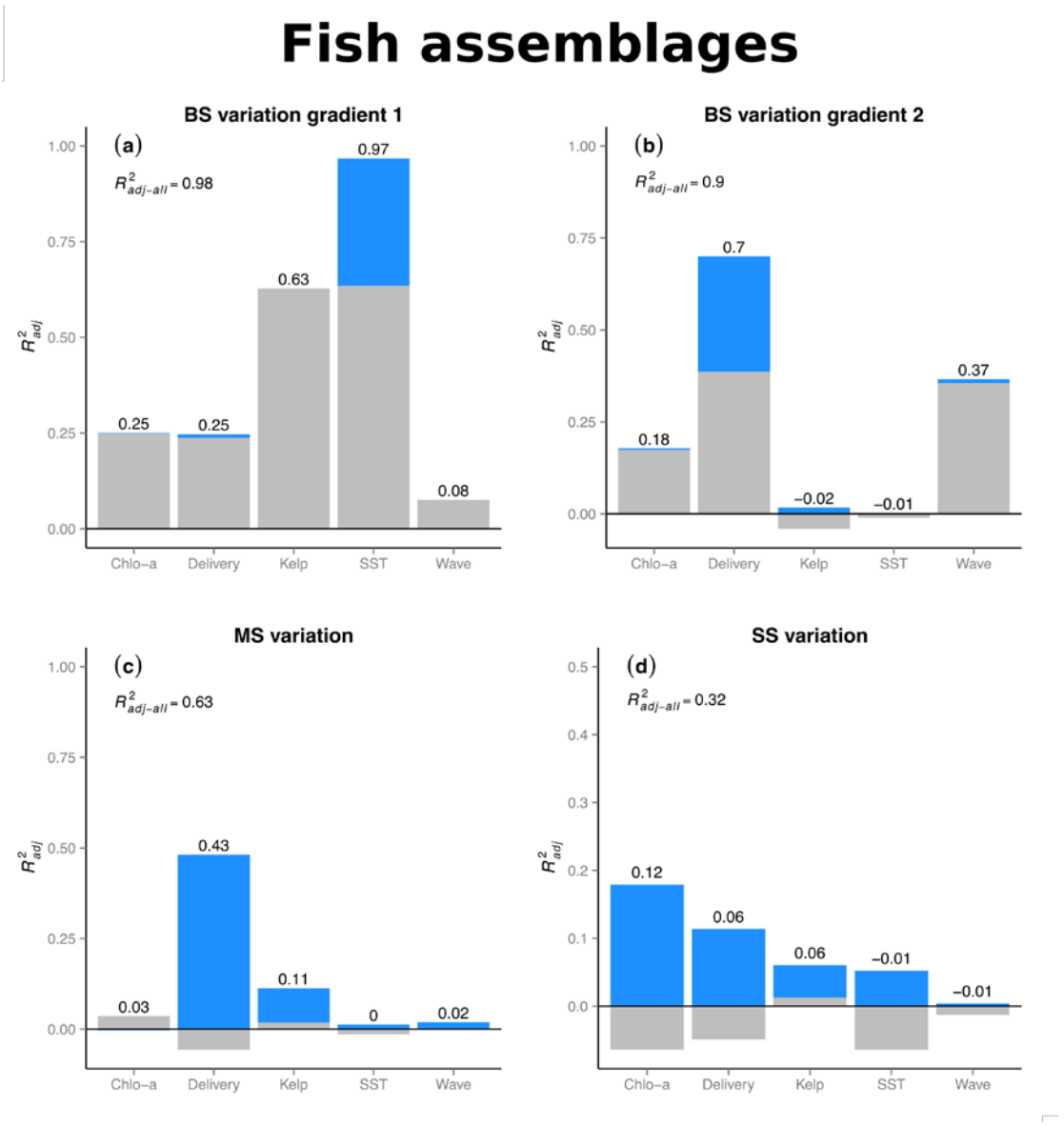


Fig. 5. Relationships between environmental variables and the variation in fish assemblages at different spatial scales. Each bar represents the contribution of an environmental variable to the variation in species composition at a given spatial scale. The total contribution of a variable is decomposed into a unique contribution (blue fraction) and a contribution that is shared with the other environmental variables (grey fraction). Negative shared contributions can emerge when variables are correlated, in which case its unique contribution can be higher than its total contribution. The number above each bar corresponds to the total contribution of each variable. The five environmental variables are chlorophyll-*a* concentration (Chlo-a), potential propagule delivery strength (Delivery), kelp biomass (Kelp), sea surface temperature (SST) and wave exposure (Wave). Note that the scale of the y-axis in (d) differs from (a-c).

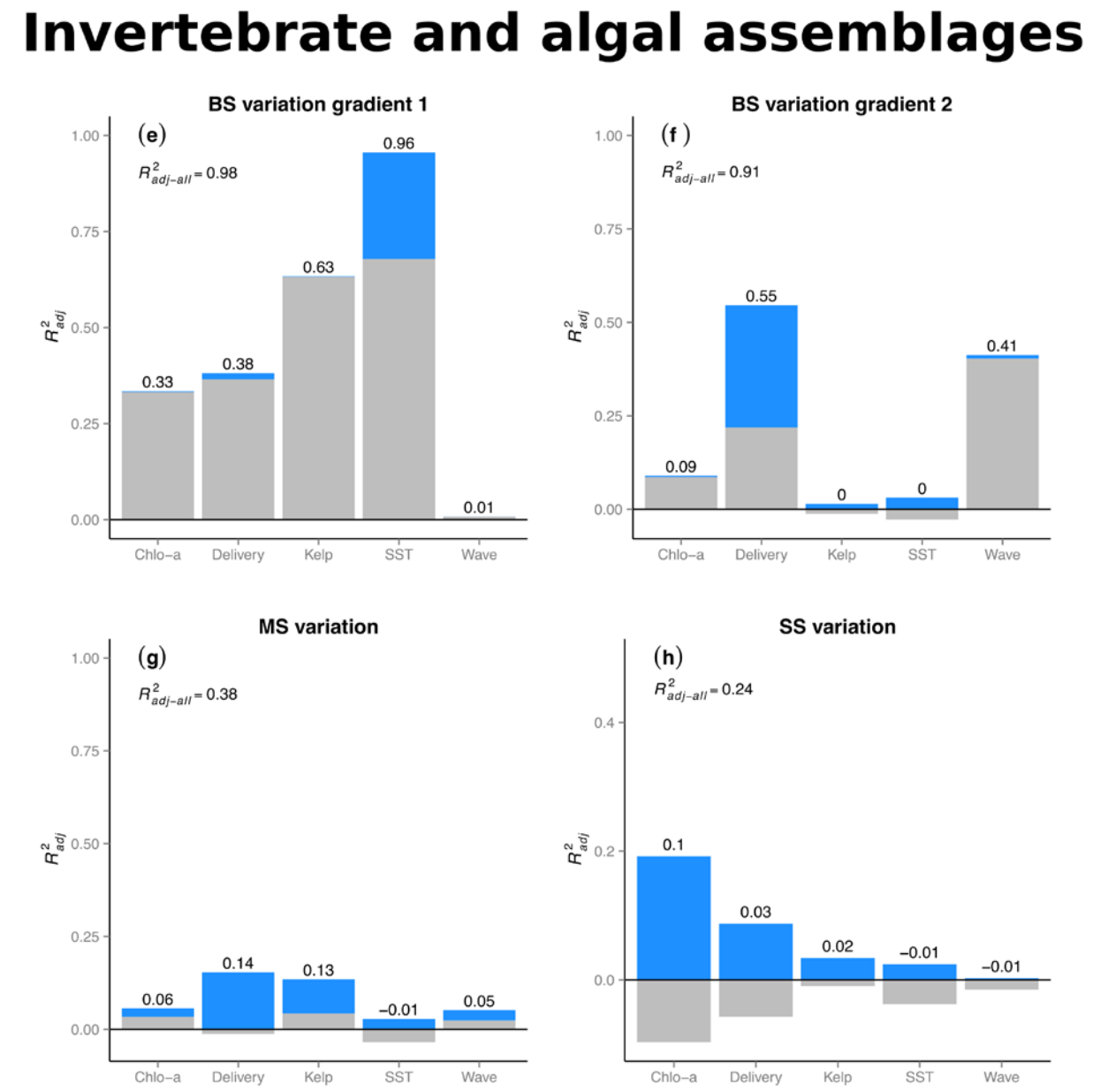


Fig. 6. Relationships between environmental variables and the variation in invertebrate and algal assemblages. See Fig. 5 for legend description.

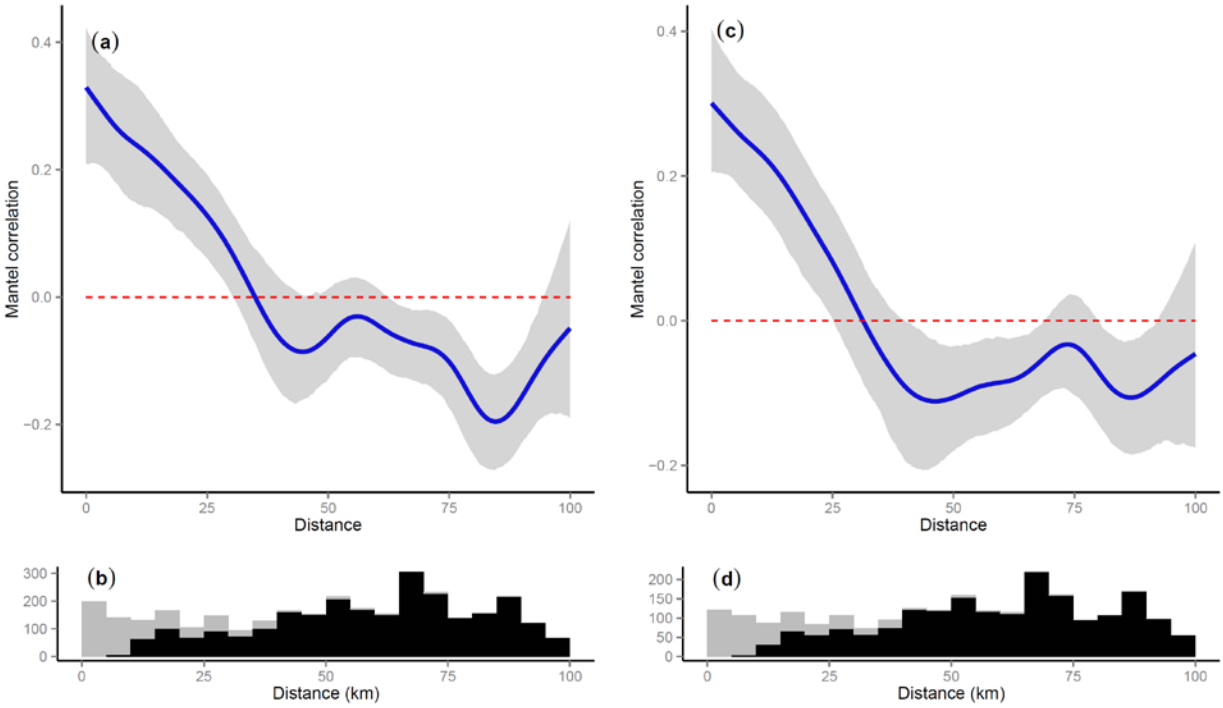


Fig. 7. Multivariate correlogram of fish assemblages (a-b) and of invertebrate and algal assemblages (c-d). The blue line corresponds to the mean spatial autocorrelation while the grey area represents its 95% confidence interval. The lower panels (b and d) illustrate the number of site pairs used for each distance class. The grey bars show site pairs that belong to the same spatial clusters (islands or mainland) while the black bars represent site pairs that belong to distinct spatial clusters.

Supplementary information can be found with the published version of this article:

https://doi.org/10.1007/s00442-017-3994-1

Department of the Interior (DOI)

The Department of the Interior protects and manages the Nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors the Nation’s trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities.

Bureau of Ocean Energy Management (BOEM)

The mission of the Bureau of Ocean Energy Management is to manage development of U.S. Outer Continental Shelf energy and mineral resources in an environmentally and economically responsible way.

BOEM Environmental Studies Program

The mission of the Environmental Studies Program is to provide the information needed to predict, assess, and manage impacts from offshore energy and marine mineral exploration, development, and production activities on human, marine, and coastal environments. The proposal, selection, research, review, collaboration, production, and dissemination of each of BOEM’s Environmental Studies follows the DOI Code of Scientific and Scholarly Conduct, in support of a culture of scientific and professional integrity, as set out in the DOI Departmental Manual (305 DM 3).